



OECD Journal

Journal of Business Cycle Measurement and Analysis

Editor-in-Chief: Michael Graff



OECD Journal

Journal of Business Cycle Measurement and Analysis

Editor-in-Chief: Michael Graff

Volume 2013/2

A Joint Publication of OECD and CIRET

The *OECD Journal* collection comprises four periodicals: *OECD Journal: Economic Studies*, *OECD Journal: Financial Market Trends*, *OECD Journal on Budgeting*, *OECD Journal: Journal of Business Cycle Measurement and Analysis*. Each periodical within *OECD Journal* retains its own ISSN, volume and issue numbering. The volume numbering of each journal corresponds to the calendar year. Volume 2013 of *Journal of Business Cycle Measurement and Analysis* contains two issues.



CIRET



Centre for International Research
on Economic Tendency Surveys

The Ciret Office is located at
KOF Swiss Economic Institute, ETH Zurich,
Switzerland

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Periodical: *OECD Journal: Journal of Business Cycle Measurement and Analysis*
ISSN 1995-2880 (print)
ISSN 1995-2899 (online)

<http://dx.doi.org/10.1787/jbcma-v2013-2-en>

Corrigenda to OECD publications may be found on line at: www.oecd.org/about/publishing/corrigenda.htm.

© OECD 2014

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.

Table of contents

General introduction: Short-term forecasting methods joint issue with Économie et Prévision	5
Short-term forecasting of French GDP growth using dynamic factor models by Marie Bessec and Catherine Doz	11
New indicators for tracking growth in real time by Troy D. Matheson	51
Dynamic factor models: A review of the literature by Karim Barhoumi, Olivier Darné and Laurent Ferrara	73
Constructing a conditional GDP fan chart with an application to French business survey data by Matthieu Cornec	109
Do business tendency surveys help in forecasting employment? A real-time evidence for Switzerland by Boriss Siliverstovs	129

Editors

Editor-in-Chief

Michael Graff, *ETH Zurich, Switzerland*

Associate Editor

Pierre-Alain Pionnier, *OECD, Paris, France*

Editorial Board

Chair

Martine Durand, *Chief Statistician, OECD, Paris, France*

Kai Carstensen, *Ifo Institute for Economic Research, Munich, Germany and University of Munich, Germany*

Marcelle Chauvet, *University of California Riverside, CA, USA*

Richard Curtin, *University of Michigan, Ann Arbor, MI, USA*

Horst Entorf, *University of Frankfurt, Germany*

Hélène Erkel-Rousse, *Ministry of Finance and Economy, Paris, France*

Roberto Golinelli, *University of Bologna, Italy*

Robert J. Gordon, *Northwestern University, Evanston, IL, USA*

Ulrich Heilemann, *University of Leipzig, Germany*

Kajal Lahiri, *State University of New York at Albany, NY, USA*

Marco Lippi, *University of Rome “La Sapienza”, Italy*

Marco Malgarini, *ISTAT, Rome, Italy*

Zbigniew Matkowski, *Warsaw School of Economics, Warszawa, Poland*

Gian Luigi Mazzi, *Eurostat, Luxembourg*

Ronny Nilsson, *former Associate Editor of the JBCMA, OECD, Paris, France*

Simon van Norden, *HEC Montréal, Quebec, Canada*

Karl Heinrich Oppenländer, *former President of IFO Institute, Munich, Germany*

Ataman Ozyildirim, *The Conference Board, New York, NY, USA*

Günter Poser, *former Editor-in-Chief of JBCMA, Technische Universität Darmstadt, Germany*

Gerd Ronning, *University of Tübingen, Germany*

Bernd Schips, *former Editor-in-Chief of JBCMA, ETH Zurich, Switzerland*

Jan-Egbert Sturm, *KOF, Swiss Institute for Business Cycle Research, ETH Zurich, Switzerland*

István P. Székely, *DG ECFIN, European Commission, Brussels, Belgium and Corvinus University, Budapest, Hungary*

András Vértés, *GKI Economic Research Co., Budapest, Hungary*

Martin Weale, *NIESR, National Institute of Economic and Social Research, London, UK*

General introduction: Short-term forecasting methods joint issue with *Économie et Prévision*

The result of an original partnership between two journals

This issue is a special issue in two respects. First, all contributions deal with short-term forecasting methods. Second and above all, this special issue results from an original and fruitful collaboration between two scientific journals: *Économie et Prévision* and the *Journal of Business Cycle Measurement and Analysis* (JBCMA).

Économie et Prévision (Economy and Forecasting) is a review that publishes studies in French by specialists in quantitative methods of applied economics, originating from a range of backgrounds (academic researchers as well as economists from French, non-French or international economic administrations).¹ One of the review's main aims is to foster scientific exchanges and technology transfers between economists from the academic world, and from the economic administration. Its editorial policy covers a wide range of topics including macro and microeconomic challenges, international issues and environmental studies. *Économie et Prévision* is one of the six main referenced French economic journals. Created in 1971, it has been published under the title *Économie et Prévision* since 1981. The review is sponsored by the Directorate General of the Treasury at the French Ministry of the Economy and Finance.²

Économie et Prévision is a more generalist publishing platform than the JBCMA. However, the two share a strong common interest in short-term forecasting methods as well as a similar peer review validation system. In addition, the present editor-in-chief of *Économie et Prévision*, Hélène Erkel-Rousse, has had close links with the Centre for International Research on Economic Tendency Surveys (CIRET) and the JBCMA for many years. She was a CIRET Council member from 2001 to 2008 and has been a member of the Editorial Board of the JBCMA since its creation in 2004. In 2011, the French Treasury became a CIRET institutional member. In order to celebrate this affiliation as well as to raise awareness of each journal to the readers of the other, it was decided to publish this joint issue.

This joint issue comprises a collection of papers submitted to either of the two journals, simultaneously released in English by the JBCMA and in French by *Économie et Prévision*.

The general topic chosen for the joint issue (short-term forecasting methods) and the soft data sources referred to in the empirical papers (business and consumer tendency surveys) are key fields of interest of both journals, as well as those of the organisations which support them (French Treasury, CIRET, KOF Swiss Institute and the OECD). Five papers were selected from recent papers submitted to either of the two journals, which illustrate various aspects of short-term forecasting methods while falling within the editorial scopes of both reviews. Two were submitted in French to *Économie et Prévision* (the paper by Marie Bessec and Catherine Doz, and that by Karim Barhoumi, Olivier Darné and Laurent Ferrara). Three were submitted in English to the JBCMA (those by Matthieu Cornec, Troy Matheson and Boriss Siliverstovs). The selection of these papers was based on mutual recognition of the peer-review process of the other review: the acceptance of a paper by one review resulted in automatic acceptance of the original version of the paper by the other. We owe our thanks to the anonymous referees who gave valuable advice on previous versions.

The translation of the original papers into the other language was a crucial issue that deserved close attention. The translated versions of the original papers were checked, corrected and validated by Hélène Erkel-Rousse with input from the translators and authors. We would like to thank Karim Barhoumi, Olivier Darné and Laurent Ferrara for providing an English version of their paper as well as Matthieu Cornec for delivering a French translation of his. We are also grateful to the OECD for sponsoring the translation of the papers by Troy Matheson and Boriss Siliverstovs into French, as well as to the French Treasury for financing the translation of the paper by Marie Bessec and Catherine Doz into English. Special recognition is owed to Jonathan Mandelbaum for his translation of the Bessec-Doz paper into English and for his precious support during the checking and finalisation of the English version of the paper. We also thank all the authors for their helpful contributions to the solving of the translation problems that inevitably occurred. Last but not least, we are indebted to Gyorgy Gyomai and Pierre-Alain Pionnier at the OECD for their support in making this joint issue possible.

Shedding light on key aspects of new developments in short-term forecasting methods

Official estimates of macroeconomic aggregates are usually released with significant delays: at least several weeks for flash estimates, notably more for less provisional data. Meanwhile, policy-makers, central bankers, business leaders and other economic agents require timely economic assessments of both the current state of the economy,³ and its future developments to make appropriate decisions according to the economic context.

The quality of these assessments is therefore of primary importance. This quality strongly depends on the suitability and accuracy of the forecasting methods⁴ used, as well as on the timeliness, reliability and appropriateness of the data to which these methods are applied. Very different kinds of data in terms of frequency, length, degree of timeliness and nature (qualitative data derived from business and consumer tendency surveys; quantitative data relating to the real economy – production, employment, etc., as well as monetary, financial and international data) constitute huge datasets from which the most relevant pieces of information for forecasting have to be extracted through optimal data selection and/or combination.

Unsurprisingly, a flourishing literature is devoted to short-term forecasting methods. This literature consists of both theoretical contributions and applied work in which the

comparative performance of competing forecasting techniques is assessed using given data sets of various kinds relating to different periods and countries. For example, since the creation of the European Monetary Union (EMU), empirical applications to European countries and to the euro area as a whole have been booming. This rapid development stems from the increasing needs of central bankers and policy-makers for timely assessments of economic conditions in the euro area for the monitoring of the EMU.

This issue by no means intends to give a comprehensive view of recent developments in short-term forecasting methods. It proposes, however, to give the reader a valuable insight into several important technical aspects in the field.

First, a number of techniques based on factor modelling have been developed in the past decade to combine numerous data of different kinds, degrees of timeliness, frequencies and lengths. Most notably, these techniques deal with the asynchronous timing of data releases and enable one to extract the information content of weekly or monthly releases for tracking quarterly variables. Factor models have quickly become both major objects of research and widespread operational tools, which are used on a regular basis. The first three papers of this special issue shed light on the latest generation of these models, namely dynamic factor models.

This branch of literature on short-term forecasting, however, only constitutes a partial response to societies' growing needs for timely economic assessments. In fact, efficient decision making does not only require predictions of the current and future general "locations" or "central tendencies" of the major economic variables (referred to as *point-forecasts* in the literature). Decision-makers also need to rely on valuable assessments of the level of uncertainty attached to these predictions; risk evaluation being a crucial issue for them. Since predictions in the form of point forecasts do not convey any guidance as to their likely accuracy, complementary forecasting tools are needed to fulfil this role, together with related techniques enabling short-term analysts to assess the accuracy of these tools. Interval forecasts, density forecasts and their illustrations (*fan charts*) are the best known tools of this kind. They have increasingly been used, notably in central banks, to describe the uncertainty inherent to any point forecast.⁵ Two papers in this issue provide a picture of these developments. The paper by Matthieu Cornec focuses specifically on the empirical assessment of forecast uncertainty, while the paper by Boriss Siliverstovs resorts to point-forecast and density-forecast techniques as complementary tools for comparing the forecast performances of different models.

A third crucial issue in short-term forecasting concerns model instability. The link between variables of interest and their leading indicators may change over time, either permanently in the case of structural breaks, or temporarily due to notable temporary shocks, or otherwise in a way that remains to be fully understood during highly unstable periods such as that of the recent crisis. This may cause forecast failures when the forecasting models are not flexible enough. This kind of problem concerns both point forecasting and density forecasting, a crisis being liable to affect not only the levels of economic variables but also their variability. The paper by Matthieu Cornec addresses this problem and suggests a possible way of solving it in the case of density forecasting.

Lastly, the wide range of techniques and approaches used in this issue does not only illustrate the richness of the short-term forecasters' toolkit resulting from the vitality of the research in the field. It also enables the reader to tackle some of the other numerous methodological issues which are intensively discussed in the literature. For instance, both

the Classical and the Bayesian approaches are represented in the papers by Troy Matheson and Boriss Siliverstovs. The question of whether one approach outperforms the other is raised through comparisons of the predictive performances of competing forecasting tools derived from the two approaches. These two papers also refer to the other highly debated issue of the performance of combined (point or density) forecasts versus that of single forecasts based on the “single-best model” approach.

Summary of the papers in this joint issue

The first three contributions in this issue illustrate the usefulness of dynamic factor models for short-term forecasting (Bessec and Doz; Matheson; Barhoumi, Darné and Ferrara). As highlighted above, this set of models proves to be particularly well adapted to dealing with big datasets consisting of heterogeneous time series in terms of nature, frequency and length. It therefore enables short-term forecasters to take into account numerous data of different kinds within a common framework to make their assessments of the state and future developments of the economy. Factor models are based on the hypothesis that the dynamics of a big set of economic variables can be summarised by that of a small set of underlying common factors. If this hypothesis holds, the estimation of these factors enables one to solve (notably) the dimension problem caused by big datasets, with limited information loss. The factors can then be used to assess the current and close future evolutions of macroeconomic aggregates of interest, such as GDP or industrial production growth.

The contributions by both Marie Bessec and Catherine Doz, and Troy Matheson apply this kind of model to two different contexts: that faced by a short-term forecaster working in a national economic organisation (Bessec and Doz); and that of a short-term forecaster having to deal with numerous heterogeneous countries, as is the case in international organisations (Matheson).

More precisely, Bessec and Doz study the performance of dynamic factor models in forecasting French GDP growth over the previous, current and following quarters. Their factors are extracted from a large data set of 93 variables including qualitative survey results and real, monetary, financial and international quantitative variables. An out-of-sample pseudo real-time assessment over the past decade shows that factor models generally provide a gain in accuracy relative to three common benchmark models (namely, a random walk with a constant, an auto-regressive model and a simple calibration model based on the INSEE⁶ index of business sentiment in France). Nevertheless, the forecasts remain fragile before the start of the quarter when information is scarce. In this context, the authors show that the use of financial and international variables can improve forecasts at the start of forecasting cycles. Beyond interesting results, the paper provides enlightening methodological discussions where the authors compare several competing estimation methods of the factors and various forecasting techniques using several evaluation criteria.

The paper by Matheson gives a striking illustration of the challenge faced by short-term forecasters working in international organisations. The aim of the paper is to develop monthly indicators for tracking short-run trends in real GDP growth in 32 advanced and emerging-market economies (that is, very heterogeneous countries in terms of both data availability and reliability). A methodology adapted to this problem must be homogeneous enough to be both tractable in an operational working context and easily applied to an

expanding list of countries. Although the context differs noticeably from that of the Bessec and Doz paper, the methodology used is again based on dynamic factor models, which enables the author to deal with big datasets containing from 97 to 290 time series (depending on the country). Testing the forecast performance of indicators derived from dynamic factor models within a quasi-real-time experiment, Matheson finds that they generally produce good real GDP growth forecasts compared to both a range of benchmark models and to pooled forecasts derived from different model combinations.

If dynamic factor models constitute a powerful tool for short-term forecasting, they can also be applied to many other contexts in which economists need to summarise numerous heterogeneous data. The theoretical research aiming to improve the estimation techniques and performances of these models is therefore booming, as is the empirical literature using this kind of model.

Karim Barhoumi, Olivier Darné and Laurent Ferrara provide a comprehensive survey of the recent literature dealing with dynamic factor models. After defining the main kinds of factor models, their survey gives valuable insight into both the parameter estimation methods and the statistical tests available for choosing the number of factors for these models. Several empirical applications of dynamic factor models are dealt with, including the construction of economic outlook indicators and macroeconomic forecasts, but also macroeconomic and monetary policy analyses. In doing so, the authors provide a reference paper on dynamic factor models.

While the first two contributions to this issue deal with point forecasting, the last two papers either focus on the issue of uncertainty assessment (Cornec) or combine point and density forecasting techniques to compare the predictive performances of different models (Siliverstovs). The paper by Cornec also addresses the issue of how to deal with increasing uncertainty during crises.

Mathieu Cornec reviews the main existing tools used to describe uncertainty. In particular, he describes and discusses the methodologies used by the French statistical institute INSEE and the Central Bank of England to derive density forecasts, while highlighting their respective drawbacks. For instance, the INSEE fan chart is *unconditional*, meaning that if the INSEE methodology is strictly applied, the magnitude of the uncertainty displayed is the same whatever the economic outlook, which is inconsistent with the increasing variability of economic variables when crises occur. The author therefore proposes a new method based on conditional quantile regression to represent uncertainty in real-time, whose main characteristic features are to be conditional upon the economic outlook, non-parametric and reproducible. He also builds a forecasting risk index associated with his new fan chart to measure the intrinsic difficulty of the forecasting exercise. Based on balances of opinion derived from business tendency surveys carried out by INSEE, the new GDP fan chart turns out to efficiently capture the growth stall during the 2008 crisis on a real-time basis. Meanwhile, the forecasting risk index increases substantially, showing signs of growing uncertainty.

Finally, the paper by Boriss Siliverstovs investigates the usefulness of business tendency surveys collected at the KOF Swiss Economic Institute and aggregated in the form of the KOF Employment Indicator (KOFEI) for the forecasting of employment in Switzerland over the current and one-quarter-ahead horizons. One of the originalities of the paper is that the author resorts to a Bayesian model averaging (BMA) framework instead of relying on a single-best model approach, thus aiming to take into account the part played by

model selection uncertainty in the forecasting process. Referring to a real-time dataset, Siliverstovs carries out an out-of-sample evaluation of the forecasting performance of autoregressive distributed lag (ARDL) models considering the KOFEI for the period 2004Q4-2010Q4. The author finds that inclusion of the KOFEI leads to a substantial improvement in prediction accuracy of both point and density forecasts compared to the performance of a benchmark autoregressive model. He also compares the predictive performances of the BMA procedure with those of the “single-best” ARDL model characterised by the highest posterior probability. They turn out to be rather similar, with a limited advantage in favour of the BMA approach.

In conclusion, we trust that this issue will make a substantial contribution to the exchange of knowledge and information on both theoretical and empirical developments in short-term forecasting methods, thus respecting the aims of the two journals which carried this joint project through together.

Hélène Erkel-Rousse (editor-in-chief, *Économie et Prévision*), Paris

Michael Graff (editor-in-chief, *JBCMA*), Zurich

November 2013

Notes

1. About 75% of the authors of the papers published in *Économie et Prévision* are academic researchers, the remaining 25% work in economic administrations and related institutions. More details on *Économie et Prévision* are available on the review's pages hosted on the French Treasury's website: www.tresor.economie.gouv.fr/2120_economie-et-prevision-in-english-brief-presentation-and-link-to-english-abstracts.
2. The opinions expressed and arguments employed in the papers published in *Économie et Prévision* do not necessarily reflect the official views of the Directorate General of the Treasury or of the French Ministry of the Economy and Finance.
3. Or, even the economic status of the previous quarter, in the months for which the related data are not yet published.
4. Across the text, the term “forecasting” is used to refer to “nowcasting” (the present state of the economy), “forecasting” (the economy's near future) and “backcasting” (the economy's potentially still unknown near past).
5. An interval forecast consists of an upper and lower bounds between which a future value of interest is expected to lie with a given probability. The interval defined by these two bounds is called a prediction interval. Density forecasting consists in estimating the entire probability distribution of a still unobserved current or future value of interest. Prediction intervals can naturally be derived from density forecasts. Fan charts result from the plotting on the same graph of several prediction intervals for different probabilities (from 10% to 90%) and consecutive future values of a given variable of interest. Prediction intervals get wider as forecast horizons become more distant, indicating increasing uncertainty. In other words, the intervals fan out, therefore the fan chart's denomination.
6. *Institut national de la statistique et des études économiques* (National Institute of Statistics and Economic Studies). The French statistical institute carries out business and consumer tendency surveys that are harmonised at the European level. Each quarter, it delivers short-term forecasts of the major French economic aggregates, notably based on calibration models involving leading indicators derived from these surveys.

Short-term forecasting of French GDP growth using dynamic factor models

by
Marie Bessec* and Catherine Doz**

In recent years, central banks and international organisations have been making ever greater use of factor models to forecast macroeconomic variables. We examine the performance of these models in forecasting French GDP growth over short horizons. The factors are extracted from a large data set of around one hundred variables including survey balances and real, financial, and international variables. An out-of-sample pseudo real-time evaluation over the past decade shows that factor models provide a gain in accuracy relative to the usual benchmarks. However, the forecasts remain inaccurate before the start of the quarter. We also show that the inclusion of international and financial variables can improve forecasts at the longest horizons.

Keywords: GDP forecast, factor models

JEL classification: C22, E32, E37

* Corresponding author: Marie Bessec, Place du Maréchal-de-Lattre-de-Tassigny, 75016 Paris, marie.bessec@dauphine.fr. LEDa-Université Paris Dauphine and Bank of France. At the time of writing, Marie Bessec was working at the Directorate General of the Treasury at the Ministry of the Economy, Finance and Industry.

** Paris School of Economics, Université Paris 1 Panthéon-Sorbonne.

We would like to thank A. Buissé, D. Denis, C. Dolignon, M. Forestier, T. Guyon, C. Mansuy, and N. Sormani, as well as two blind peer reviewers, for their observations and proposals. We also wish to thank L. Ferrara and all those involved in the Séminaire Fourgeaud in May 2011 for their many comments. Any mistakes and omissions are ours, and the views expressed in this paper do not necessarily reflect those of the Directorate General of the Treasury.

This article was translated from French by Jonathan Mandelbaum, with financial support from the Directorate General of the Treasury at the French Ministry of the Economy and Finance.

Introduction

Factor models have received substantial coverage in the literature in recent years (for a review of the models, see Stock and Watson, 2010; Bai and Ng, 2008b). Central banks and other international organisations are using them increasingly for short-term forecasting of GDP. The models are used in static form (for example at the Federal Reserve [Fed], under the impulse of the studies by Stock and Watson, 1999, 2002a, 2002b) or in dynamic form (at the European Central Bank [ECB], following the studies by Doz, Giannone, and Reichlin, 2011, 2012; Giannone, Reichlin and Small, 2008; at the Bank of Italy with the Eurocoin indicator developed by Altissimo et al., 2001, 2010). In France, by contrast, factor models are still relatively seldom used.

Factor models offer several advantages over classic tools. First, they can incorporate information provided by a large set of variables and summarise it in a small set of factors, which will then serve as explanatory variables in a standard regression model. Second, factor models can be adjusted if observations are missing at the end of a period. This is a valuable property for the short-term economic analyst, who is constrained by the availability of short-term indicators (release times are fairly short for balances of opinion¹ in business and consumer surveys and for financial variables, longer for real variables such as the industrial production index and manufactured-goods consumption). When one uses a factor model, one does not need to develop auxiliary models to predict missing observations or to use different models depending on the month of the quarter in which the forecast is prepared, that is, depending on the information available to the analyst.

Our article describes an application of these methods to the forecasting of French GDP growth in the previous,² current and following quarters. We use a database of about one hundred variables such as survey variables, real indicators, monetary and financial variables, and international indicators. An out-of-sample assessment shows that the quality of the forecasts supplied by our factor models is satisfactory, although longer-term forecasts are fragile. Some of the improvements suggested in our conclusion are worth exploring for the sake of improving accuracy.

Our study is organised as follows. Section 1 reviews the methods used for the short-term forecasting of GDP in various French and international institutions. We focus on the use of factor models in some of these organisations. Section 2 describes the factor models in their static and dynamic forms, as well as the associated estimation and forecasting methods. Section 3 presents the data used in our study and examines the forecasting performance of factor models tested on their sample base and on an out-of-sample basis.

1. Forecasting methods used in French and European organisations

This section presents an overview of the methods used to prepare short-term forecasts of GDP growth in the main French institutions and in selected international organisations (we provide a summary in Table A.1). Several organisations have developed

tools based on factor models, which they apply in parallel with other approaches, and different estimation methods are employed for these models.

1.1. French organisations

The French *Institut National de la Statistique et des Études Économiques* (INSEE, National Institute of Statistics and Economic Studies) forecasting units, the Directorate General of the Treasury, and the Bank of France use two complementary approaches to prepare short-term forecasts of French GDP growth: the first relies on macro-sectoral models; and, the second involves GDP calibration tools based on survey data.

Macro-sectoral models are used in two stages. The Directorate General of the Treasury prepares forecasts of certain supply components (production in the manufacturing industry, the food industry and the energy sector) and demand components (household consumption, general-government consumption and investment in construction). These forecasts are based on simple calibration models (or “bridge” models). The explanatory variables are survey balances of opinion (mainly from INSEE surveys), replaced by real quantitative indicators as they become available during the quarter. Most of the calibration models have been constructed to reproduce the methodology of the quarterly accounts or by using the General-to-Specific (GETS) algorithm,³ which selects the explanatory variables based on the most general specifications, via a gradual elimination of the least manifestly contributing variables. At a second stage, the macro-sectoral supply and demand forecasts are made consistent within an accounting framework that comprises several activity sectors.⁴

The Bank of France’s OPTIM model⁵ breaks down economic activity on the supply side for the current- and next-quarter forecasts.⁶ The first stage therefore, consists in preparing separate forecasts of the main economic sectors (agrifoods, manufactured goods, energy, construction and market services). Next, a GDP growth forecast is deduced by weighting the sectoral forecasts by values estimated with linear regressions over a fairly short period. Simple calibrations are used to forecast production components. The calibrations are identical regardless of the position in the quarter. Missing values are forecast by an autoregressive moving average (ARMA) model. Nevertheless, two concurrent equations are examined, with or without the industrial production index (IPI) of the sub-sector concerned (except for the forecast of service production, where the IPI of the manufacturing industry is used). The other explanatory variables of these equations are the balances of opinion of the tendency surveys by the Bank of France, INSEE and the European Commission, plus real quantitative indicators. The equations are constructed in two stages using the GETS algorithm. The algorithm is first applied to each source (for example, to the opinion balances of the service-sector survey, then to the balances in the manufacturing-industry survey, and so on), then to all these preselected variables together.⁷

In parallel with the use of these models, GDP calibrations based on survey data are used to verify the overall forecast. Erkel-Rousse and Minodier (2009), following Dubois and Michaux (2006), developed vector autoregressive (VAR) models for each month in the quarter and univariate calibrations of quarterly GDP growth per month and for each forecast horizon. The aim is to take into account the change in the total information set during the quarter. In univariate calibrations, explanatory variables are selected from opinion balances of INSEE surveys in the manufacturing, industry and services using the GETS algorithm.⁸ The Directorate General of the Treasury also calibrates GDP growth rates,

using two types of calibrations. The first (Bessec, 2010) are based exclusively on opinion balances of INSEE surveys (in manufacturing, services and the building sector). These calibrations are also prepared month by month and for each forecast horizon. Explanatory variables are selected using a block variant of the GETS method (Hendry and Krolzig, 2005). The second type of calibration is simpler and correlates GDP growth with the composite purchasing managers index (PMI).

The Bank of France also uses calibrations linking GDP to survey data (ISMA indicator)⁹ for the current quarter and following quarter.¹⁰ However, these calibrations rely on factor models. The information from Bank of France surveys of manufacturing and services is summarised by means of a principal component analysis (PCA). Factors are selected by means of an original GETS-based method. More specifically, the factor model is applied in several stages. First, a PCA is performed on the monthly opinion balances of each sector or sub-sector. The first factors of each PCA are selected and converted to quarterly values.¹¹ The GETS algorithm is then used to choose these preselected factors (and their lags) as explanatory variables of current GDP growth (for the current-quarter forecast) or future GDP growth (for the next-quarter forecast).

More recently, the Bank of France has conducted two studies on the use of static- and dynamic-factor models in short-term economic forecasting. These studies have yielded a factor model used to forecast GDP growth in the euro area for the current or previous quarter. The first study, carried out with other European central banks (Rünstler et al., 2009), concerns the forecast of GDP growth in the euro area as a whole, and in a group of ten euro area member states including France, in the current, following and previous quarters taking indicator release times into account. The second study (Barhoumi, Darné and Ferrara, 2010) concentrates on the next-quarter French GDP forecast as well as situations where the indicators are available for all three months of the quarter to be forecast. The explanatory variables used are survey variables, real variables, financial variables and price variables.

The two studies compare four methods for estimating factor models: the Stock and Watson method (1999, 2002a, 2002b); the two methods proposed by Doz, Giannone and Reichlin (2011, 2012) with a two-stage estimation (see Section 2) or a quasi-maximum likelihood estimation (QMLE); and, the method developed by Forni et al. (2000, 2005). Both studies find that factor models are superior to the benchmark methods, that is, a combination of bivariate calibrations, bivariate VARs and simple autoregressive (AR) processes. According to the first study, the models estimated with the Kalman filter appear to yield better results. However, the second study notes that the results obtained with this method do not differ significantly from those yielded by the simplest methods. The second study also suggests that there is no need to use a highly detailed database and that forecast quality is improved by incorporating more factors than recommended by the Bai and Ng (2002) criteria.

1.2. International organisations

1.2.1. European Central Bank

In the recent period, the ECB has been using two concurrent approaches to prepare short-term forecasts of euro area growth in the previous, current and following quarters.¹² Both approaches are used twice a month: mid-month after the release of real indicators such as the IPI; and then at the end of the month after the release of business and consumer tendency surveys and financial data.

The first approach rests on the combination of forecasts drawn from about ten standard calibrations (Rünstler and Sédillot, 2003; Diron, 2008). Several simple calibrations¹³ linking GDP growth to a small set of monthly indicators (real indicators, financial indicators and survey opinion balances) are estimated every month. The monthly indicators are forecast using simple autoregressive models for the missing months,¹⁴ then averaged to quarterly values in order to feed the calibrations. The final GDP growth forecast is obtained by simply averaging the forecasts of the different calibrations. Rünstler and Sédillot (2003) confine their choice of explanatory variables to real indicators and survey opinion balances, whereas Diron (2008) also includes financial variables.

The second approach is based on dynamic factor models introduced at the ECB (and at the Federal Reserve [Fed]) in keeping with the method presented by Giannone, Reichlin and Small (2008). Whereas the first approach relies on relatively few monthly indicators (up to 15 in Diron, 2008), the information set in the second approach comprises 85 monthly indicators, real indicators, financial indicators, and indicators derived from business and consumer tendency surveys. A Kalman filter is used to calculate missing factor observations due to the missing months of the monthly indicators. The factor model is estimated using the two-stage estimation method (PCA and Kalman filter) proposed by Doz, Giannone and Reichlin (2011). In this context, Bańbura and Rünstler (2011) measure the variables' contribution to forecasts and apply the results to the short-term GDP forecast for the euro zone.¹⁵

Angelini et al. (2011) compare the two approaches and conclude that factor models are superior for forecasting short-term euro area GDP growth. Angelini, Bańbura and Rünstler (2008) conducted a similar study but besides the GDP forecast, they look at the forecasts of GDP supply and demand components. In factor models, components are first forecast separately and then simultaneously by introducing a restriction (the accounting relationship between GDP and its components) in the space-state representation. This study also concludes in favour of factor models for GDP forecasting. In a study on growth forecasts for the euro area and six EU countries including France, for forecast horizons ranging from one to twelve quarters, Caggiano, Kapetanios and Labhard (2009) show by empirical means that forecasting accuracy is improved by: 1) pre-filtering the variables from which the factors are extracted using the criteria proposed by Boivin and Ng (2006); and, 2) combining the forecasts obtained from several factor models.

1.2.2. European Commission

The European Commission publishes bi-annual macroeconomic forecasts and sub-annual forecasts (usually in May and November) for EU member states, the euro area and the entire EU. The forecasts, which concern the current year and the following year, are obtained by means of an iterative process. They rely not on a centralised econometric model but on analyses by country teams. Each analysis is based to a varying degree on statistical methods and expert opinion. The forecasts are tested for consistency, notably in regard to trade flows. The euro area and EU aggregates are not forecast directly but obtained by aggregating the country forecasts. Since 2006, between these two forecasting cycles, the Commission has also published updated GDP and inflation forecasts for the current year only for seven of the leading euro area members. These interim forecasts are largely based on standard calibrations.¹⁶

1.2.3. *Organisation for Economic Co-operation and Development (OECD)*

Twice a year, in June and December, *OECD Economic Outlook* publishes annual and sub-annual forecasts over a horizon ranging from 18 months to two years. Between each main publication, the OECD also releases an interim assessment (generally in March and September). Within this framework, it prepares projections for a set of macroeconomic variables for each of the 34 member countries and selected non-OECD countries. The forecasts also rely on a combination of statistical methods and expert opinions. The international consistency of the projections is verified using a model developed by Pain et al. (2005).

In assessing the short-term situation, special attention is paid to calibration models constructed by Sédillot and Pain (2005) for six major OECD countries including France, and for the euro area. Standard calibration models link GDP growth to real indicators, financial indicators and survey opinion balances. Auxiliary models are used to forecast the missing months of monthly indicators, which are then made quarterly. The explanatory variables for GDP growth calibration and for the auxiliary models are selected by means of a semi-automatic procedure.¹⁷ The variables chosen for France are two INSEE survey variables: the recent production trend; and, the personal production outlook in manufacturing, plus two real variables: the growth rates of the IPI; and, household consumption of manufactured goods (for France, the introduction of real variables is useful for the current-quarter forecast when the indicators are known for at least the first two months of the quarter, but is useless for forecasting the next quarter).

In sum, factor models are now widely used in European institutions but are still relatively seldom used for short-term forecasts of French GDP growth. By applying these models to French GDP growth forecasts, our study therefore follows on the earlier research by Rünstler et al. (2009) and Barhoumi, Darné and Ferrara (2010). Unlike these two previous articles, we do not try to compare the forecasts derived from estimation methods that differ from the underlying factor model and we confine our choice to the two-stage estimation method proposed by Doz, Giannone and Reichlin (2011). By contrast, we study how the results are affected by the different methods for selecting the underlying factor model and by the various forecast construction methods.

2. Factor models and their use in forecasting

This section gives a concise description of factor models in their static form and their dynamic extension. We go on to discuss alternative methods for estimating the models. We conclude by reviewing the methods that can be used to construct a forecast based on the prior estimation of a factor model.

2.1. *Factor models*

2.1.1. *Static factor model*

Factor models are designed to supply a parsimonious representation of the information provided by a large set of variables when these are correlated. Factor models assume that the observed variables can be described in terms of a small set of latent, unobservable variables called “factors” or “common factors” and that these latent common factors are the source of the correlations between the observed variables. In the static framework, there are two types of factor models: 1) exact factor models, in which the factors explain the entire correlation between variables; and, 2) approximate factor models, which are suited to cases where the number of observed variables tends toward

infinity and where the factors explain most of the correlations between variables (the residual portion being negligible).

More formally, with N the number of variables studied, T the number of observations available for each variable and x_{it} the observation of variable i at instant t , the exact model with r factors $(f_{jt})_{j=1,\dots,r}$ is written as follows:

$$x_{it} = \mu_i + \lambda_{i1}f_{1t} + \lambda_{i2}f_{2t} + \dots + \lambda_{ir}f_{rt} + e_{it}, \quad (1)$$

for $i = 1, \dots, N$, $t = 1, \dots, T$, $r < N$, that is, in matrix form:

$$x_t = \mu + \Lambda f_t + e_t, \quad t = 1, \dots, T \quad (2)$$

with $x_t = (x_{1t}, \dots, x_{Nt})'$ and $e_t = (e_{1t}, \dots, e_{Nt})'$ N -dimensional vectors, $f_t = (f_{1t}, \dots, f_{rt})'$ a r -dimensional vector, Λ an (N, r) -dimension matrix and under the following assumptions: $E(e_t) = 0$, $E(f_t) = 0$, $E(e_t e_t') = D = \text{diag}(d_1, \dots, d_N)$, $E(f_t f_t') = I_r$, $E(f_t e_t') = 0 \quad \forall (t, \tau)$, $E(f_t f_\tau') = 0 \quad \forall (t, \tau)$, $t \neq \tau$, $E(e_t e_\tau') = 0 \quad \forall (t, \tau)$, $t \neq \tau$, I_r representing the r -dimensional identity matrix, and $(d_1, \dots, d_N)'$ a vector of N positive parameters to be estimated.

In what follows, we shall focus on the case where $\mu = 0$ and work with variables mean-centred beforehand.

When r is very small compared with N , the model does indeed yield a parsimonious representation of the covariances between x_{it} variables, as shown below.

In this static model, the r common factors are not auto-correlated. We can further assume without loss of generality, that they are not correlated with one another and have unit variance. The term e_{it} , called the specific or idiosyncratic component, represents the share of variable x_{it} that is not explained by the common factors. As the e_{it} variables are uncorrelated two by two, the entire correlation between observed variables is provided by the factors.

The factor weights (λ_{ij}) measure the covariances between the observed variables i and the common factors j . The variance of each variable can thus be written as:

$$V(x_{it}) = \sum_{j=1}^r \lambda_{ij}^2 + d_i.$$

The term λ_{ij}^2 represents the share of the variance of x_{it} explained by factor j .

The term $\sum_{j=1}^r \lambda_{ij}^2$ is the total share of the variance (communality) captured by the r factors. In addition, the variance-covariance matrix of the vector of observed variables is written as $V(x_t) = \Lambda \Lambda' + D$ and as D is diagonal, the covariances between the observed variables are explicitly expressed in terms of factor loadings. Thus, the variance-covariance matrix of x_t is expressed in terms of the $N(r+1)$ parameters of Λ and D instead of depending on $N(N+1)/2$ parameters if we do not assume the existence of a factor model. Note that the model is invariant to change of scale, so that decomposing the variance-covariance matrix of x_t is equivalent to decomposing its correlation matrix.

In the approximate static model, one no longer assumes that the idiosyncratic terms are uncorrelated two by two. It is merely assumed that in the correlation between the observed variables, the share due to the correlation between the idiosyncratic terms is negligible compared with the share due to the common factors. If one continues to write $E(e_t e_t') = D$ (with a non-diagonal matrix D), one assumes that when the number N of observed variables tends toward infinity, the matrix D remains bounded whereas the matrix $\Lambda \Lambda'$ is unbounded. Consequently, as $V(x_t) = \Lambda \Lambda' + D$, the share of the correlation between variables not explained by the factors can be regarded as negligible.

2.1.2. Dynamic factor models

Dynamic factor models aim to provide a parsimonious description of the common dynamics of the observed variables (or of the co-movements of the observed variables). These models generalise static models (exact or approximate) in two ways. First, the common factors are auto-correlated. Their dynamics are typically modelled in VAR form or in some cases, in vector autoregressive moving average (VARMA) form. Second, the observed variables can be influenced by the factor's contemporary values, but also by their lagged values. In both cases the model can be reduced, via suitable notation changes, to a form close to that of static factor models.

Let us begin by examining the framework of exact dynamic factor models. Within it, assuming the factor dynamics are correctly represented by a VAR(p) model and still using $x_t = (x_{1t}, \dots, x_{Nt})'$ to denote the vector of observed variables, one can define an initial class of models in which factors are included only via their contemporary values. These models have the following form:

$$x_t = \Lambda_0 f_t + e_t \quad (3)$$

$$f_t = \sum_{\ell=1}^p A_\ell f_{t-\ell} + \varepsilon_t \quad (4)$$

where (ε_t) is white noise and (e_t) is a process (which can be white noise or possess a dynamics) whose components are uncorrelated two by two ($E(e_{it}e_{jt}) = 0 \quad \forall t, \tau, i, j, i \neq j$) and are uncorrelated with the factors.

If f_t is a vector of size q and if we write $F_t = (f_t', \dots, f_{t-p+1}')$, it is straightforward to see that this type of model can also be written as:

$$x_t = \Lambda F_t + e_t \quad (5)$$

$$F_t = A F_{t-1} + \underbrace{B \varepsilon_t}_{\zeta_t} \quad (6)$$

with the matrixes Λ and A expressed in simple form as functions of the matrixes $\Lambda_0, A_1, \dots, A_p$ and $B = (I_q, 0 \dots 0)'$. The vector $F_t = (f_t', \dots, f_{t-p+1}')$ is a vector of size $r = pq$, and the model is defined as having r static factors (the static factors are the components of F_t) and q dynamic factors (the dynamic factors are the components of f_t).

We can also obtain the same type of formulation when the factor operates not only on a contemporary basis but also with its lags, that is, in the context of a model of the form:

$$x_t = \Lambda_0 f_t + \dots + \Lambda_s f_{t-s} + e_t \quad (7)$$

$$f_t = \sum_{\ell=1}^p A_\ell f_{t-\ell} + \varepsilon_t \quad (8)$$

Thus, the general form of an exact dynamic factor model is given by:

$$\begin{cases} x_t = \Lambda F_t + e_t & \text{with } V(e_t) = \Phi = \text{diag}(\varphi_1, \dots, \varphi_N) \\ F_t = A F_{t-1} + \underbrace{B \varepsilon_t}_{\zeta_t} & \text{with } V(\zeta_t) = \Sigma_\zeta \end{cases} \quad \begin{matrix} (9a) \\ (9b) \end{matrix}$$

where (ε_t) is white noise and (e_t) is a process (which can be white noise or possess a dynamic) whose components are uncorrelated two by two ($E(e_{it}e_{jt}) = 0, \quad \forall t, \tau, i, j, i \neq j$) and

are uncorrelated with the factors, B being a matrix (r, q) of rank q , with $q \leq r$. We accordingly state that r is the number of static factors in the model and q is the number of dynamic factors.

As with static models, we extend the scope of application of dynamic models by introducing approximate dynamic factor models when the number N of observable variables tends toward infinity. In this type of model, we allow the components of vector e_t to be correlated with one another, but we assume that the share of the observable variables' dynamics due to the idiosyncratic components is negligible by comparison with the factor-related share.

The equation system (9) has a form that puts it in the general category of space-state or state-measure models. Equation (9a), called measurement equation, describes the relationship between the observed variable x_t , called measurement variable, and the unobserved state variable F_t . Equation (9b), called state equation, describes how the latent variables are generated from their lags and from innovations. Annex A provides a description of the space-state representation and the estimation algorithms.

2.2. Estimation of a dynamic factor model

We conducted our study within the framework of approximate dynamic factor models, which are the standard choice for analysing macroeconomic data. Various methods for estimating these models have been proposed in the literature. For a full survey of the methods, see Bai and Ng (2008b), and Stock and Watson (2010).

The method most commonly used is principal component analysis (PCA), first proposed by Stock and Watson (2002a). This method is applied to a static factor model (or a dynamic factor model converted to static form using the procedure described in Section 2.1.2). Under the assumptions usually made in the specification of the approximate factor model, PCA¹⁸ is shown to yield convergent estimators of the model's parameters and an approximation of the factors that converges toward their true value when the number N of series studied and the number T of observations tend toward infinity.

However, other estimation methods have been proposed to allow factor dynamics to be taken into account. Forni et al. (2000, 2005) propose an estimation method based on the analysis of the spectral density of observations. Doz, Giannone and Reichlin (2011, 2012) have proposed a pseudo-maximum likelihood estimation method and a two-stage estimation method based on the Kalman filter.

The two-stage estimation method is fairly simple to implement. It has the added advantage of easily adjusting to missing values; one of the main problems faced by short-term analysts, as noted earlier. The two-stage method was used, for example by Giannone, Reichlin and Small (2008) to forecast US and euro area GDP, and by Angelini et al. (2008), and Bańbura and Rünstler (2011) to prepare a short-term forecast of euro area GDP.

It is important to stress here that PCA implementation requires a balanced data sample. This imposes a severe constraint on short-term forecasting. If we truncate the sample at the last date for which all the data are available, we deprive ourselves of a part of the existing information.

By contrast, with the two-stage method proposed by Doz, Giannone and Reichlin (2011), we can calculate the best approximations of factor values at each date, taking into account all the information available. Assuming normal disturbances, we know that the Kalman filter and smoother allow us to obtain for a given parameter value, the optimal

approximation of the latent variables on the basis of the full information available on the observable variables. We therefore find the two-stage method to be particularly well suited to short-term forecasting, and we shall accordingly choose it here.

With $\theta = (\Lambda, A_1, \dots, A_p, \Phi, \Sigma_\varepsilon, B)$ as the notation for the model parameters, the method comprises the following two stages:

1. Matrix Λ of equation (9a) is estimated by means of a PCA on the sub-sample of truncated data, that is, for the common period when the values of all the x_{it} variables are known. We also make a preliminary estimate of the factors at each sub-sample date. We use this initial estimate of factor values to estimate the parameters of equation (9b) with standard regression methods. Matrix B is estimated by means of a PCA on the estimated residuals $\hat{\zeta}_t$.
2. The factors are estimated for the entire estimation period, that is, including the instants when observations for some of the variables are missing. We can do this by applying the Kalman filter and smoother algorithms to the space-state representation (9).

At iteration t , depending on whether the variable i is available or not, we posit:

$$E(e_{it}^2) = \begin{cases} \varphi_i & \text{if } x_{it} \text{ is available} \\ +\infty & \text{if not} \end{cases}$$

in matrix R of the filter algorithm (see Annex A), with φ_i the variance of e_{it} , estimated on the truncated sample. In other words, if the value of variable i is not known at instant t , we replace the diagonal term on line i by an infinite term in matrix R of the filter algorithm. As the matrix is then inverted, this is tantamount to assigning a zero weight to variable i at iteration t . We deduce a new estimate of the factors for the entire period. By construction, this is the optimal approximation of the factors based on all the available information.

2.3. Use in forecasting

When, as in this circumstance the variable to forecast is a quarterly variable (noted y_t), whereas the data used and hence the factors as well have monthly values, the forecast initially requires the factors to be made quarterly. If $f_{j,t}^Q$ stands for the quarterly value of the j^{th} factor at date t given by the factor-model estimate, the forecast rests on a regression y_t of (here, quarterly GDP growth) on the $f_{j,t}^Q$ values.

At this point, two approaches are generally chosen. For our purposes we shall use both in order to compare their forecasting performance.

The first approach is suitable for static factor models, but can also be used for dynamic factor models since these can be represented in static form, as seen earlier. The approach consists in estimating the following model by means of ordinary least squares (OLSs):

$$y_{t+h} = c + \sum_{j=1}^r \delta_j f_{j,t}^Q + u_{t+h} \quad t = 1, \dots, T-h \quad (10a)$$

and then calculating the forecast for y_{T+h} at date T using the following equation:

$$\hat{y}_{T+h|T} = \hat{c} + \sum_{j=1}^r \hat{\delta}_j f_{j,T}^Q \quad (10b)$$

where c and $(\delta_j)_{j=1, \dots, r}$ denote the first-order parameters to be estimated, $\hat{c}$ and $(\hat{\delta}_j)_{j=1, \dots, r}$ their OLS estimators, and u the error term of the model.

To implement this approach in the present study, we use two equations, one for the current and previous quarter GDP forecast, the other for the next quarter forecast. More specifically, for the current and previous quarter forecast we estimate equation (10a) with $h = 0$, then calculate the forecast using equation (10b) for $h = 0$ and $h = -1$. For the next-quarter forecast, we use equations (10a) and (10b) with $h = 1$. This type of approach, introduced by Stock and Watson (2002a, 2002b) for static models, is used for example, by Boivin and Ng (2006), Rünstler et al. (2009), and Barhoumi, Darné and Ferrara (2010).

The second approach is specifically used for dynamic models. It relies on the estimate of factor dynamics obtained when estimating the factor model. If the factors confirm a model of the form: $f_t = \sum_{\ell=1}^p A_\ell f_{t-\ell} + \varepsilon_t$, we can obtain recursively a forecast $f_{T+h|T}$ of f_{T+h} at date T using the estimated values of the A_ℓ matrixes and the factors. We can then make the forecasts obtained quarterly and determine a forecast $f_{j,T+h|T}^Q$ of each quarterly factor. Next, we estimate by means of ordinary least squares the equation linking GDP growth to the quarterly $f_{j,t}^Q$ factors that are contemporary with it:

$$y_t = c + \sum_{j=1}^q \delta_j f_{j,t}^Q + u_t \quad t = 1, \dots, T, \quad (11a)$$

This equation coincides with equation (10a) when the latter is estimated for $h = 0$. We obtain the forecast y_{T+h} of at date T using the following equation:

$$\hat{y}_{T+h|T} = \hat{c} + \sum_{j=1}^q \hat{\delta}_j f_{j,T+h|T}^Q \quad (11b)$$

This type of approach is applied by Giannone, Reichlin and Small (2008), Angelini, Bańbura and Rünstler (2008), and Bańbura and Rünstler (2011).

In this study, we also set out to compare the forecasts obtained using these two approaches with those obtained using a third method, which in a sense is a compromise between the first and second. In this new approach, we use equations (10a) and (10b) but change the procedure when the forecasting date T is the first or second month of a quarter. More specifically, we use equation (10a) as before but applying a VAR representation to obtain a factor forecast for all three months of the quarter concerned. We then calculate the forecast by applying equation (10b) to the associated quarterly factor.

2.4. Choice of model specification

Until now, we have assumed that the number of factors and the order of the VAR on the factors are known. In practice, that is not the case. In our application, we have used two different methods to choose the model specification.

2.4.1. Information criteria

We begin by using the standard information criteria (Akaike information criterion [AIC] and Bayesian information criterion [BIC]) to choose the VAR order p in (4):

$$AIC(p) = \ln[\det(\hat{\Sigma}_\varepsilon)] + \frac{2N^2p}{T} \quad \text{and} \quad BIC(p) = \ln[\det(\hat{\Sigma}_\varepsilon)] + \frac{N^2p \ln(T)}{T}$$

where T is the number of observations, N the number of variables in the system (that is, the number of variables used to estimate the factors), $\hat{\Sigma}_\varepsilon$ the variance-covariance matrix of the residuals of the estimated VAR, $\det$ the determinant and $\ln$ the natural logarithm. We choose the order $p \in \{1, \dots, p_{\max}\}$ that minimises the information criterion considered.

Bai and Ng (2002, 2007) offer criteria for choosing the number of factors. In their 2002 article, they introduce an initial series of criteria suited to static factor models:

$$\begin{aligned} PC_1(r) &= V(r) + r \frac{N+T}{NT} \ln \left(\frac{NT}{N+T} \right) & IC_1(r) &= \ln(V(r)) + r \frac{N+T}{NT} \ln \left(\frac{NT}{N+T} \right) \\ PC_2(r) &= V(r) + r \frac{N+T}{NT} \ln C_{NT}^2 & IC_2(r) &= \ln(V(r)) + r \frac{N+T}{NT} \ln C_{NT}^2 \\ PC_3(r) &= V(r) + r \frac{\ln C_{NT}^2}{C_{NT}^2} & IC_3(r) &= \ln(V(r)) + r \frac{\ln C_{NT}^2}{C_{NT}^2} \end{aligned}$$

with $C_{NT}^2 = \min(N, T)$ and $V(r) = N^{-1} \sum_{i=1}^N \hat{\sigma}_i^2$, where $\hat{\sigma}_i^2 = \hat{e}_i' \hat{e}_i / T$, $\hat{e}_i = (\hat{e}_{i1}, \dots, \hat{e}_{iT})'$ is the vector of residuals of (1), the r static factors being estimated using PCA. The number of factors to choose, when applying one of these criteria, is the number $r \in \{1, \dots, r_{\max}\}$ that minimises the criterion concerned.

In their 2007 article, the same authors propose a second series of criteria to determine the number of dynamic factors q . Let $\hat{\Sigma}_{\hat{e}}$ be the variance-covariance matrix of the residuals of the VAR estimated on the r static factors and $c_1 \geq c_2 \geq \dots \geq c_r \geq 0$ the ordered eigenvalues of matrix $\hat{\Sigma}_{\hat{e}}$. The authors define the quantities:

$$D_{1,k} = \left(\frac{c_{k+1}^2}{\sum_{j=1}^r c_j^2} \right)^{1/2} \quad \text{and} \quad D_{2,k} = \left(\frac{\sum_{j=k+1}^r c_j^2}{\sum_{j=1}^r c_j^2} \right)^{1/2}$$

and associate them with two decision-making rules:

$$\kappa_1 = \{k : D_{1,k} < m / \min[N^{0.5-\delta}, T^{0.5-\delta}]\} \quad \text{and} \quad \kappa_2 = \{k : D_{2,k} < m / \min[N^{0.5-\delta}, T^{0.5-\delta}]\}$$

where m and δ are parameters whose values are to be calibrated. The authors choose $q_1 = \min\{q \in \kappa_1\}$ or $q_2 = \min\{q \in \kappa_2\}$. They opt for $\delta = 0.1$ and show through simulations that satisfactory results are obtained for $m = 1$.¹⁹

In practice, these criteria are used in three stages. First, we use one of the six PC or IC criteria (Bai and Ng, 2002) to determine the optimal number of factors $r^* \in \{1, \dots, r_{\max}\}$ in a static setting. Second, we estimate a VAR on these r^* factors and we choose the VAR order $p^* \in \{1, \dots, p_{\max}\}$ so as to minimise the standard AIC or BIC criterion. Third, we apply the Bai and Ng (2007) criteria to the variance-covariance matrix or the correlation matrix of the VAR(p^*) residuals to obtain the optimal number of dynamic factors q^* .

2.4.2. Root mean square error (RMSE) criterion

Several studies show that in practice, the use of the Bai and Ng criteria can entail the choice of too few factors, undermining forecast quality; see for example, Barhoumi, Darné and Ferrara (2010) for an application to the French GDP forecast, and Schumacher (2007) for an illustration concerning German GDP. A possible explanation is that the choice of factor-model specification is totally unrelated to the variable to be forecast.

Like Schumacher (2007), we therefore propose an alternative to the information criteria so as to compare the results obtained. The alternative consists in choosing the

number of factors that minimises the RMSE criterion in the GDP growth regression on the factors:

$$\text{RMSE} = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2} \quad \text{with} \quad \hat{y}_t = \hat{c} + \sum_{j=1}^r \hat{\delta}_j f_{j,t}^Q \quad t = 1, \dots, T$$

where $\hat{y}_t$ is the GDP growth rate estimated from equation (10b) or (11b), depending on the method used. The RMSE criterion also determines the choice of order p of the VAR process on the factors.

3. Use of dynamic factor models to forecast French GDP growth

3.1. Data

We use a database of 93 variables (Table A.2). Like most studies of GDP forecasts derived from factor models (Table A.1), we have chosen four groups of variables:

- *survey balances*: The main balances of INSEE business tendency surveys used to construct the synthetic (or business climate) indicators in manufacturing, services, the building sector and the retail trade, plus the main balances of the consumer tendency survey;
- *real variables*: Household consumption of manufactured goods and its components, new car registrations, building starts and building permits, the industrial production index and its components, and labour market variables;
- *nominal variables* (monetary and financial): Interest rates, yield-curve slope, several stock market indexes, a market volatility index, monetary aggregates and price indexes;
- *international indicators*: Euro exchange rate against major currencies, and indicators for the German and US economies.

Many short-term analysts base their forecasts on survey variables and as they become available, on real variables: in particular the industrial production index, household consumption of manufactured goods, building starts, building permits and customs data for foreign trade (see Section 1). By contrast, the nominal and international variables considered here are less often used for calibration. However since the 2008-09 crisis, these variables have been the focus of special attention. In our simulations we shall see that there is indeed an advantage in including them, particularly for the next-quarter forecast.

The estimates reported here were prepared with the series published in early June 2010. For lack of available series, we have not studied the impact of data revisions on the results. The variables are seasonally adjusted²⁰ and most have monthly frequencies. We have converted the quarterly series to monthly series using a cubic spline function.²¹ Some financial variables published daily or weekly have been converted to monthly values by taking the last observation of the month. Lastly, one should bear in mind the timing differences in series releases. Some variables are available in the month to which they apply (survey opinion balances and financial variables) but real variables are released one or two months later.

We analysed the stationary properties of the variables. We applied several unit-root tests (ADF, Phillips Perron and KPSS) to all the series (in log form for real and financial series except interest rates, and labour-market series). The latest recession caused a steep jump in levels in most series at the end of the period. To guard against this effect, we performed the tests for the period 1980-2007. The results are summarised in Table A.3. With a few exceptions, the survey opinion balances are stationary, and the real and

financial variables are order-one integrated (except the yield-curve slope which is stationary in most tests). In light of these results, we differentiated the non-stationary variables to obtain stationary variables, then standardised all the variables.

3.2. In-sample analysis on a full-information basis

Our first analysis is based on a full information set, that is, without taking into account the indicators' publication calendar. In other words, the results described here are those that would be obtained a few days before the release of the quarterly accounts, when the available indicators notably include the industrial production index for all three months of the quarter.

We first report the model estimation results for the period from first-quarter 1990 to third-quarter 2009 (1990Q1-2009Q3) using all 93 variables listed in Table A.2. On the total sample and for a number of dynamic factors $q \leq 5$, and maximum lags $s=2$ and $p=3$, the Bai and Ng criteria (2002, 2007) make us select $q = 5$ and $s = 1$, and the BIC criterion make us choose order $p = 1$ for the VAR on the factors (we use the criterion IC_2 in static mode and the criteria κ_1 and κ_2 based on the variance-covariance matrix in the dynamic form).²² The model accordingly contains $r = q(s+1) = 10$ static factors.

We therefore begin by estimating the model associated with this specification, using the two stage method (PCA, followed by the Kalman filter and smoother) as proposed by Doz, Giannone and Reichlin (2011). We then estimate Equation (10a) under the two specifications used in the forecasting stage, that is, the specification associated with $h = 0$ (used for the current and previous quarter forecasts) and the specification associated with $h = 1$ (used for the next-quarter forecast).

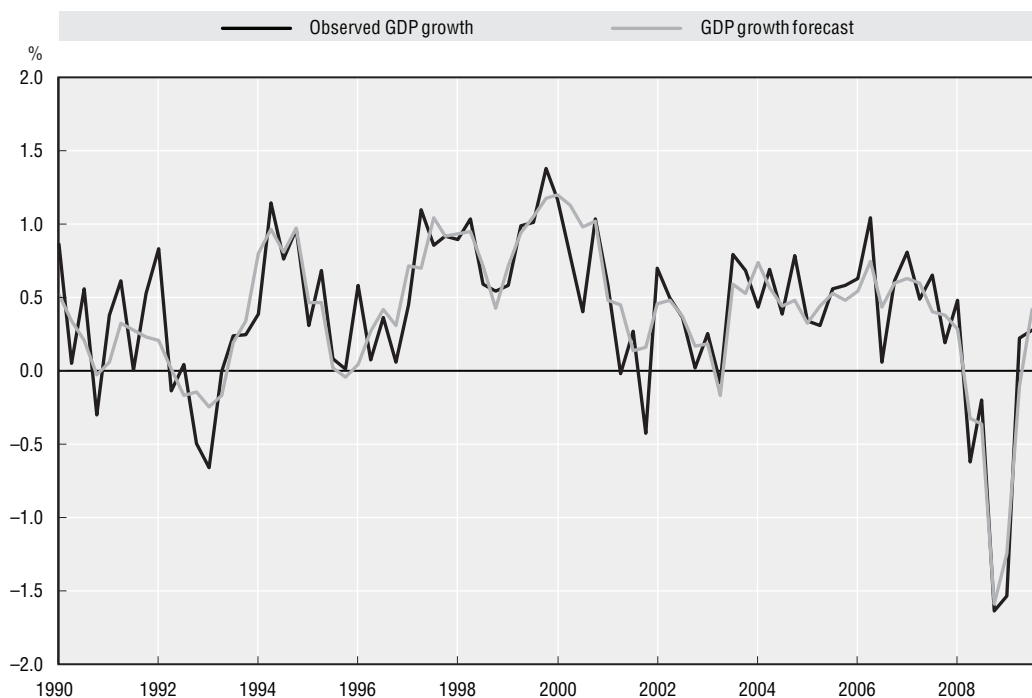
The results of these estimations are shown in Table A.4. After the removal of non-significant factors,²³ the two equations comprise five explanatory variables which are almost all identical. The adjustment seems to be of very high quality, with determination coefficients of 78% and 74% respectively. This result needs to be downplayed when $h = 0$ since the adjustment is based on the assumption that all indicators are known for the three months of the quarter to be estimated (by comparison, an autoregressive process to which we add the industrial production index available for the entire quarter yields a very high R^2 of approximately 0.65). The result is however, encouraging when $h = 1$.

Figure 1a and 1b plot the observed and adjusted GDP growth rates for the two equations. Both models reproduce GDP fluctuations rather accurately, although the 1993 and 2001 troughs are not well captured. By contrast, the adjustment quality is good during the latest recession in 2008-09.

Figure 2 gives the weights of each of the 93 variables (which we have standardised) in the construction of the six static factors ultimately chosen for the two models:

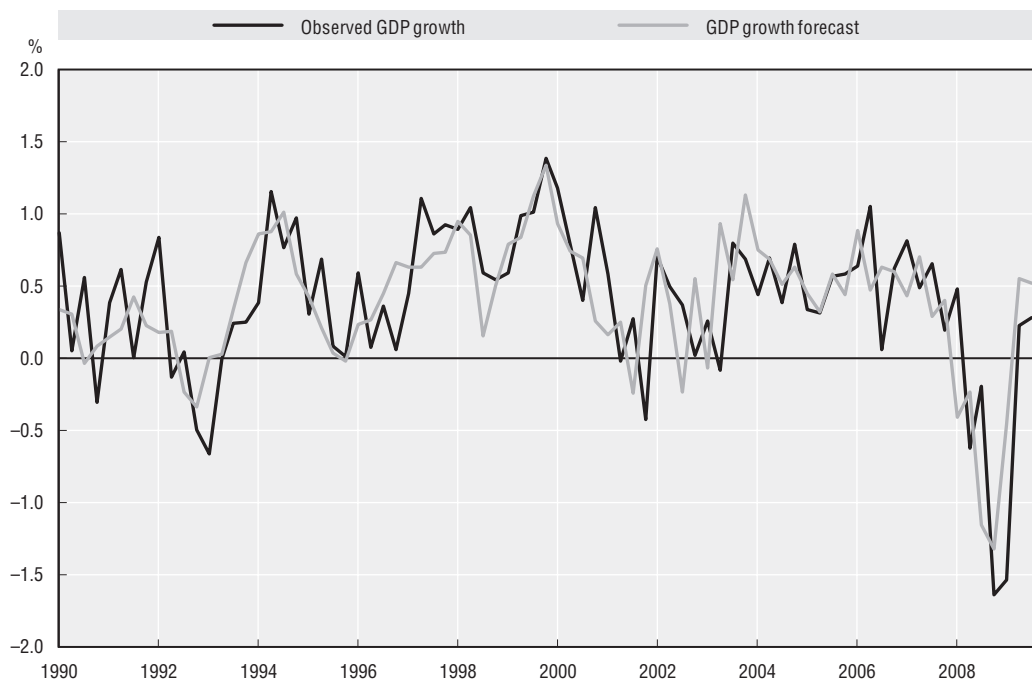
- Factor 1 mainly reflects the information on economic activity provided by survey variables for manufacturing, services, the building sector and retail trade.
- Factors 2 and 3 are of a more financial character, given the significant weight of stock market indexes in their construction.
- Factor 5 incorporates a major share of consumer survey opinion balances, loans and interest rates.
- Factor 9 is heavily dependent on industrial production indexes, interest rates and building permits.

Figure 1a. **Actual and adjusted GDP growth rate: Current quarter forecast**
Quarterly change in real GDP (wda-sa, %)



Note: Equation (10b) with $h = 0$.
Sources: See Table A.2.

Figure 1b. **Actual and adjusted GDP growth rate: Next quarter forecast**
Quarterly change in real GDP (wda-sa, %)



Note: Equation (10b) with $h = 1$.
Sources: See Table A.2.

- Factor 10 is mainly linked to the construction indicators building starts, building permits and interest rates (notably for real estate loans), and the consumer-survey opinion balance on intentions of making major purchases.

3.3. Out-of-sample analysis and pseudo-real-time analysis

3.3.1. Description of experimental framework

We now assess the quality of the forecasts in actual forecasting conditions apart from the non-incorporation of data revisions. We do so by evaluating out-of-sample forecasts, that is, outside the estimation period and allowing for the variables' release calendar.

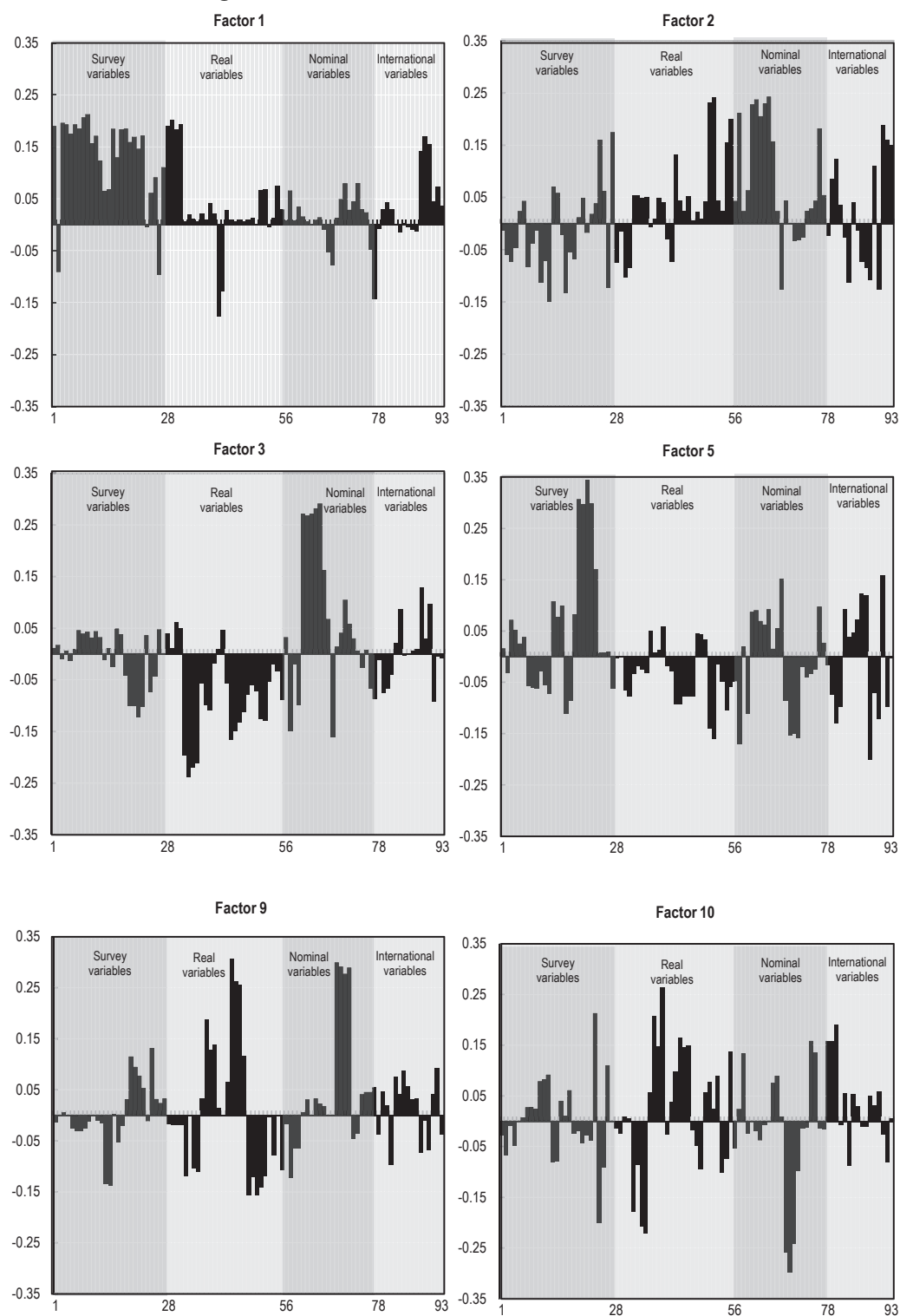
Given the differences in release calendars for the series, we assess the forecasting models in eight instants, starting with the first month of the previous quarter Q-1 (no information on quarter Q) up to the second month of the next quarter Q+1, when we have observations of indicators for all months of quarter Q. Figure 1a gives the chronology of GDP forecasting exercises for quarter Q.

The out-of-sample analysis is performed as follows, using a standard procedure. We remove the last 39 observations from the sample of 80, we estimate the models for the remaining period from 1990Q1 to 1999Q4, and we construct the forecasts over the various horizons and the associated forecast errors. We then increase the sample by one observation, re-estimate the models, and calculate the new forecasts and the related forecast errors. We increase the sample again by one observation, repeat the previous calculations and so on until we have processed the last 39 observations.

It is important to proceed in this manner because the in-sample and out-of-sample results can differ substantially. We must therefore give priority to the out-of-sample assessment which reflects actual forecasting conditions. The assessment however, differs from actual conditions in two ways. First, the exercise is not totally out of sample as the specification was chosen on the total sample (choice of order p of the VAR and of the number of factors). Moreover, as noted earlier this exercise is conducted in *pseudo real time*. We take account of the timing of series publications and we truncate each series so as to use only those values of the series that would have been available at the date when the forecast would have been calculated in an operational context. However, we do not allow for revisions to the series²⁴ (it is important to note that real quantitative indicators such as GDP, the IPI and household consumption of manufactured goods are heavily revised).

This exercise yields a series of forecast errors for each forecast horizon²⁵ to which we apply the standard assessment criteria. The root mean square forecast error (RMSFE) and mean absolute forecast error (MAFE) measure the dispersion of forecast errors and their mean amplitude. The Pesaran and Timmermann test (1992) evaluates the models' abilities to predict the direction of change in the series; in this case, the accelerations and decelerations in economic activity. At this level, we make sure that the proportion of directions of change forecast correctly exceeds 50%.

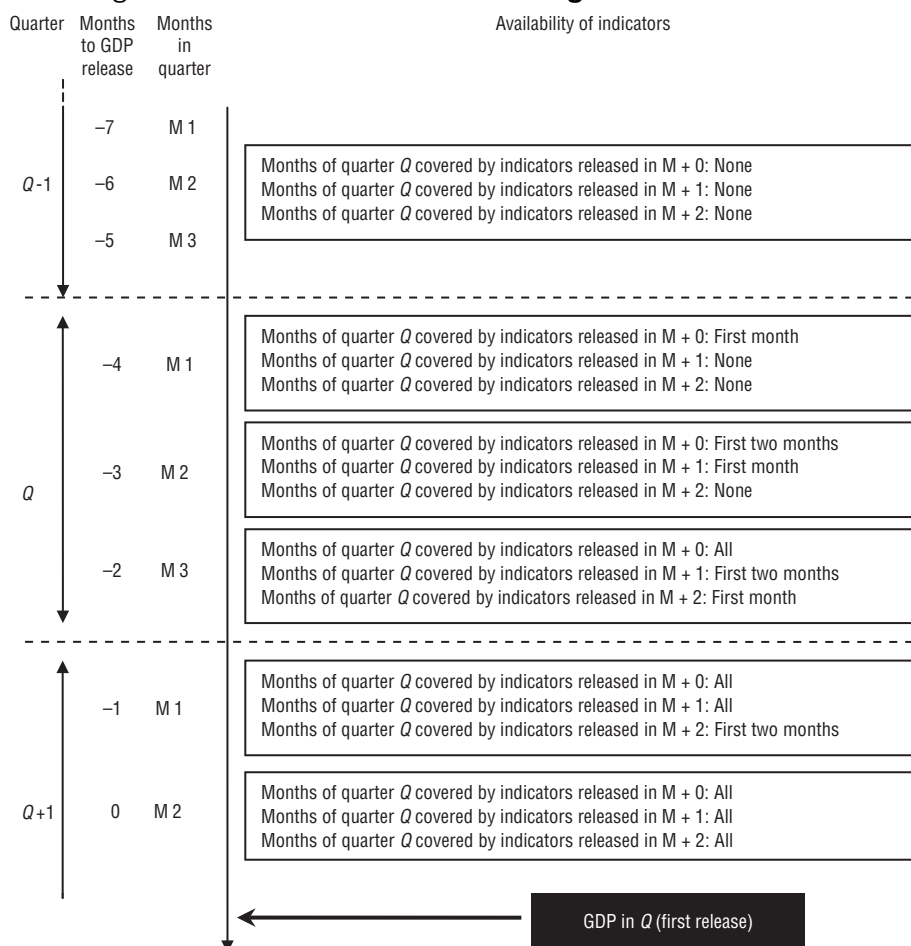
We conducted this exercise for three factor models. Model FM1 is the factor model obtained with the 93 variables. Model FM2 is built from the indicators usually tracked by short term analysts, that is, survey opinion balances and real variables. By comparing the two models, we can measure the contribution of the addition of the nominal and international blocks to the short term GDP growth forecast. We also examine a third model FM3, which uses all the variables except the real variables as these have the longest publication times and can therefore degrade the earliest forecasts. When the forecast

Figure 2. **Contributions of variables to factors**


Interpretation: The 93 variables considered are classified by category along the x-axes. The y-axes show the weights of each of the 93 variables in the construction of the six factors ultimately chosen.

Sources: See Table A.2.

Figure 3. Forecast calendar and change in information set



Notes: Q denotes the current quarter. "Indicators released in $M+0$ " are the monthly indicators released in month M of the quarter studied (e.g. industrial firms' assessments of their January activity are known at the end of January). "Indicators released in $M+1$ " are the monthly indicators released with a one-month delay (e.g. consumption of manufactured goods is known at the end of February). "Indicators released in $M+2$ " are those released with a two-month delay (e.g. the January IPI is published in early March). "Months of quarter Q covered by indicators released in $M+1$: first month/first two months" means that, in a given quarter $Q-1$, Q , or $Q+1$, the indicators released in $M+1$ are known for the first month or first two months of the quarter studied.

horizon is remote, the unavailability of a large number of series requires a more severe truncation of the sample to prepare the first-stage estimate (using PCA) since the latter is based on a balanced sample. When we add the real variables, the factors and their dynamics are therefore estimated with more data but over a shorter period.

To assess the sensitivity of results to the choice of the number of factors and lags, we also defined the specifications in two different ways: 1) on the basis of the information criteria; and, 2) on the basis of adjustment quality (RMSE criterion) for a maximum of five dynamic factors and maximum s and p lags of 2 and 3 respectively. We applied these criteria to the three variable sets corresponding to specifications FM1, FM2 and FM3 respectively, on the total sample.

As noted in Section 2.3, for each specification considered we compared the results obtained using three forecasting methods:

1. Use of equations (11a) and (11b), with the application of VAR model to forecast future factor values.

2. Use of equations (10a)²⁶ and (10b), without the extension of factors to the current quarter.
3. Use of equations (10a) and (10b) with the application of VAR model to extend factors to the current quarter.

By construction, methods 1 and 3 coincide for the current and previous quarter forecasts (that is, for a number of months before publication ranging from -4 to 0 in the diagram in Figure 3), since equation (10a) coincides with Equation (11a) when $h = 0$. Likewise, methods 2 and 3 coincide at the end of each quarter, since there is no point in extending the factors here; this happens when the number of months before publication is -5 or -2 . The three methods coincide when the quarter to forecast has passed (that is, for a number of months before publication ranging from -2 to 0).

As in most similar studies, we also compare the forecasts using factor models with those based on two simple benchmark specifications: a random walk with a constant (the GDP growth forecast is obtained as the mean of the process); and, an autoregressive process of order two (for this autoregressive specification, we chose two lags on the basis of the correlogram, the partial correlogram and standard diagnostic tests for the entire period). The forecasts over a horizon greater than unity are obtained by iteration of the resulting equation. We also examined a very simple calibration model constructed from the INSEE index of business sentiment in France. The additional equation is obtained by regressing contemporary GDP growth (or GDP growth one quarter ahead for the next quarter forecast) on the French business sentiment index in the forecast month in level terms and in quarterly variation terms in the previous month. In the last two forecasting exercises, we supplement the calibration model with the quarterly growth rate of the manufacturing production index (or its statistical overhang one month before GDP publication).

3.3.2. Results

The results of the assessments of the forecasts from first quarter 2000 to third quarter 2009 are shown in Tables A.5. For each model studied, we assess the forecast errors by calculating root mean square forecast errors (RMFSEs: Table A.5a) and mean absolute forecast errors (MAFEs: Table A.5b) and reporting the results of the Pesaran and Timmermann test (Table A.5c).

We give the results for the three forecasting methods listed previously, distinguishing between the forecast horizons so as to identify those for which the methods effectively diverge and those for which they coincide by construction.

For each proposed forecasting method, we study the forecasts obtained with the factor models (FM1, FM2, FM3) associated with the three data sub-sets presented above. We also choose two specifications for each factor model: the specification derived from the application of information criteria; and, the specification derived from the RMSE-based criterion.

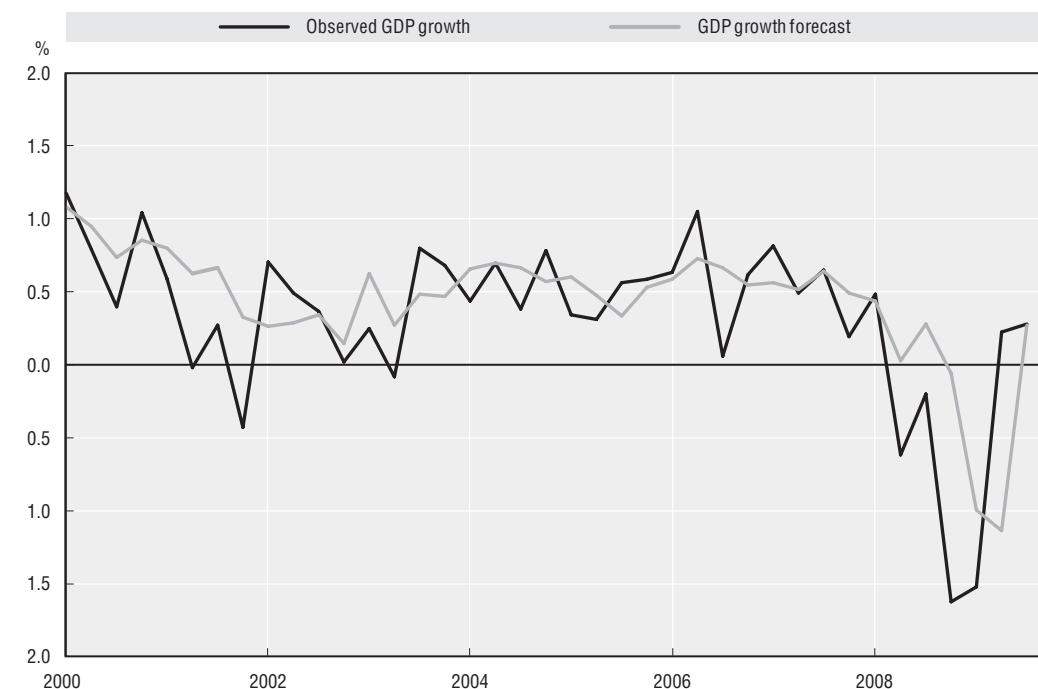
The comparison with the autoregressive process and the random walk favours the factor models, particularly in quarters Q and $Q+1$. In RMSFE terms, the gain is 22% relative to the autoregressive process in the first forecast (see line -7 ($Q - 1$) of Table A.5a, method 3) for the best factor-model specification (FM3 specified on information criteria). Generally speaking, however, the forecasting quality in quarter $Q-1$ is rather poor. The gain increases as one approaches the release of the quarterly accounts. A few days before their

publication, the gain ranges from 38% for the lowest-performing factor model to 60% for the best specification (see line 0 (Q+1) of Table A.5a). We also find a significant gain relative to the standard calibration model, ranging from 10% to 32% for the shortest horizons (see lines -2 (Q) to 0 (Q+1) of Table A.5a).²⁷

The quality of forecasts based on factor models improves almost continuously as the GDP release date approaches, that is, as the information set for the quarter to forecast increases. The absolute mean error falls from 0.32% in the best specification seven months before GDP release to just 0.15% a few days before. This result is illustrated in Figures 4a and 4b, which show the forecasts obtained recursively in the first month of quarter Q-1 (Figure 4a) and the forecasts a few days before the release of the quarterly accounts (Figure 4b) for the most effective specification over each of these horizons.

Figure 4a. **Actual and forecast GDP growth: Forecast in M-7**

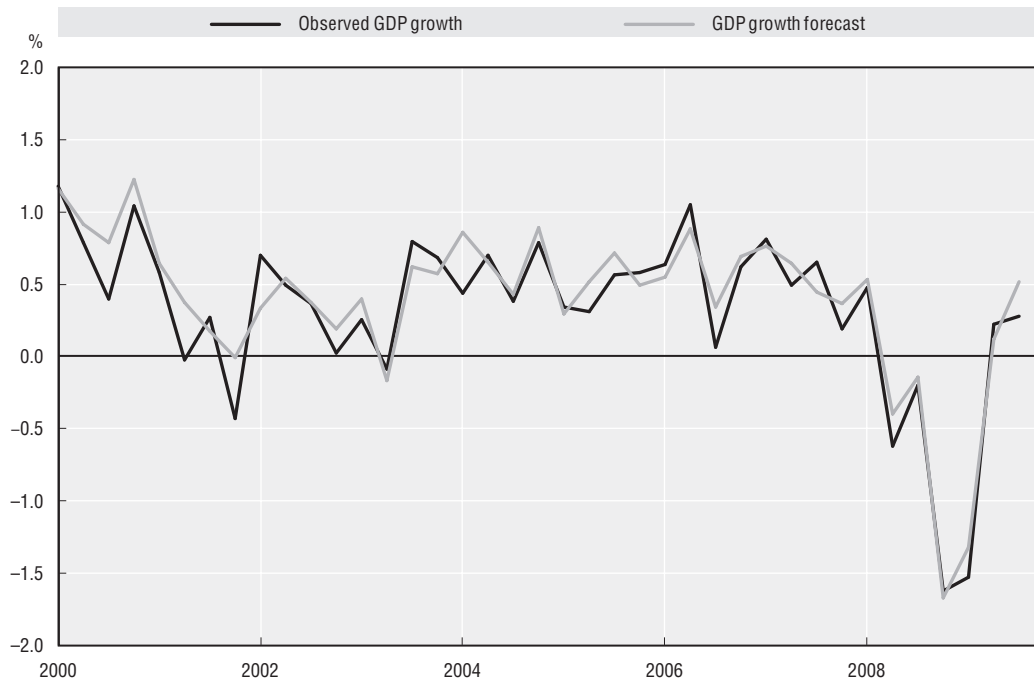
Quarterly change in real GDP (wda-sa, %), Method 3, model FM3



Sources: See Table A.2.

As regards the forecasting method, we find that for the forecasts prepared in the three months of quarter Q-1 and for each model studied, the most effective is method 3 using equation (10a) with $h = 1$ and extending the factors to the entire quarter Q-1 by means of the VAR. The result holds for the first two months of quarter Q (Equation [10a] is in this case, specified with $h = 0$, so methods 1 and 3 coincide). Note that the superiority of method 3 is verified by both the RMFSE and MAFE criteria. This result shows that the dynamic specification of the factors and hence the possibility of forecasting these factors make a tangible contribution to forecasting quality. However of the forecasting methods tested, method 3, which links the value to be forecast y_T directly to the factor values in quarter Q-1, outperforms method 1, which links y_T to contemporary factor values and so requires us to forecast the factors over a longer horizon.²⁸

Figure 4b. **Actual and forecast GDP growth: Forecast in M + 0**
 Quarterly change in real GDP (wda-sa, %), Methods 1, 2, and 3, model FM2



Sources: See Table A.2.

As regards the model-choice criteria, the results most often favour the information criteria, especially for the FM1 and FM3 models. The exception is in the one month or a few days before GDP release, when most models selected by RMSE minimisation yield equivalent or better results. However information criteria generally lead to a choice of more parsimonious specifications.²⁹ Our results therefore, do not match those of Barhoumi, Darné and Ferrara (2010). Our study however is not directly comparable with theirs, since we incorporate the changes in the entire information set during the quarter and our assessment period is different. In particular, for the longest forecast horizons, parsimonious specifications yield better results.

Concerning the sensitivity of results to the choice of variables, our first observation is that the inclusion of nominal variables (monetary and financial) and international variables improves forecast quality at the start of the period. Model FM3 (which uses all variables except real variables) yields the best results during the entire quarter Q-1 and the first two months of quarter Q. This finding (corroborated by the fact that FM1 produces better results than FM2 for the same forecast horizons) seems all the more valuable as nominal and international variables are often excluded from short-term forecasting tools. By contrast, the superiority of FM3 over FM1 for these horizons seems to indicate that the inclusion of real variables degrades the quality of the earliest forecasts provided by factor models. Closer to the GDP release date, the best model is FM2, that is, the one estimated from survey variables and real variables.

These findings consequently support the choice of short-term analysts who exclude real indicators such as industrial production at the start of the quarter (often it is not included in forecasting tools until the first two months of the quarter are published, that is, given the indicator's release calendar, not until the start of the first month of the next

quarter). In sum, our out-of-sample analysis suggests that one should build the factors in advance using the entire database except real data at the start of the forecasting cycle, then using real data and survey opinion balances as GDP release approaches.

Conclusion

Our study has examined the performance of a tool based on dynamic factor models for forecasting French GDP growth over short horizons. Such models allow the inclusion of information provided by a large variable set, summarised into a small set of factors. In their dynamic form, the models allow a time dependence of factors and a dependence of observed variables on contemporary and lagged factor values. If some indicator values are missing, we can adjust the associated estimation methods, avoiding the need for auxiliary models.

The out-of-sample and pseudo-real-time assessment of these methods generally shows the superiority of forecasts based on factor models over those obtained with standard benchmark processes. Nevertheless, the forecasts remain fragile before the start of the quarter, when information is scarce. In this context, we also show that the use of financial and international variables can improve forecasts at the start of forecasting cycles.

Several approaches could be explored for improving these results. The choice of different sets of initial variables seems to yield different forecast qualities. Ahead of factor construction, it might therefore be worth applying the variable selection methods recommended by Boivin and Ng (2006), and more recently Bai and Ng (2008a). The use of these methods by Caggiano et al. (2009) and Schumacher (2010) does show a gain for the GDP forecast, and Charpin's application on French data (2009) of the method proposed by Bai and Ng (2008a) yields encouraging results. Moreover, the introduction of non-linearities in the specification has thus far been relatively little explored in the context of factor models and could also be a major source of improved performance.

Notes

1. For a qualitative question in a business or consumer survey offering three possible categories of response (positive: "increasing" or "above average"; neutral: "stable" or "near average"; negative: "decreasing" or "below average"), the balance of opinion is the difference between the percentage of positive responses and the percentage of negative responses. Depending on the question, respondents' answers are weighted (by firm size for questions concerning an aspect of the business activity of the surveyed firm) or not weighted (for questions concerning the general business climate).
2. Recall that during the first half of a given quarter, the previous quarter GDP figure has not yet been released and is therefore still a forecast.
3. GEneral To Specific: The original version was proposed by Hoover and Perez (1999), and extended by Hendry and Krolzig (1999, 2001).
4. INSEE publishes its results quarterly in *Note de conjoncture* (Summaries of latest trends, March, June and December) and *Point de conjoncture* (Concise Report, October). The Directorate General of the Treasury does not disclose its sub-annual forecasts.
5. The model was developed by Irac and Sédillot (2002), and its current version at the time of this writing is presented by Barhoumi et al. (2012).
6. Demand is broken down as well (with forecasts of household consumption, general government consumption, investment, exports and imports) for the purposes of illustration however the breakdown is not used for the GDP forecast.

7. In the original version, Irac and Sédillot used principal component analyses (PCAs) rather than the GETS algorithm.
8. More recently, Minodier (2010) adapted these calibration models to forecast the initial results rather than the final figures of the quarterly accounts. For France, the author shows in both theoretical and empirical terms that when preparing the initial GDP growth forecast, it may be preferable to use the series of initial GDP results (that is, the series of initial GDP publications in the quarterly accounts, 45 days after the end of the quarter) rather than the final series (that is, the series available at the time the forecast is calculated) as the explained variable.
9. The ISMA indicator was introduced in 1997 and has since been modified. See Darné and Brunhes-Lesage (2007) for a description of the indicator in its current form.
10. The Bank of France publishes only the current quarter GDP forecasts at the same time as its monthly business survey of manufacturing and services.
11. At this level, several methods are considered: the average of the latest three months available; the average of the available months of the quarter only; and, the average of the three months of the quarter after extending the missing months by means of an autoregressive process.
12. The ECB also prepares medium term forecasts of economic activity for the current and following years from national forecasts by the network of national central banks. This process involves iteration between the national central banks and the ECB.
13. Diron (2008) uses eight calibration models (with fifteen explanatory variables in all) while Rünstler and Sédillot (2003) use seven (with eight explanatory variables). See Table A.1 for more details.
14. Rünstler and Sédillot (2003) use more elaborate specifications, notably a Bayesian VAR (BVAR) model.
15. They show on the basis of this statistic, that survey variables and financial variables contribute more than real indicators to the euro area GDP forecast, once the differences in the variables' release dates have been taken into account.
16. In parallel, the European Commission used a dynamic factor model for the short term forecasting of euro area GDP growth (Grenouilleau, 2004, 2006) but it abandoned the model in 2007.
17. The monthly indicators are first ranked according to the R^2 of the relationship linking GDP to its lags, and to the indicator and its lags. The best-performing indicators (and their lags) are then incorporated into a model linking GDP to its lags (up to four lags for the endogenous variable and the explanatory variables). An information criterion is applied to select the best combination of explanatory variables among all possible combinations.
18. In practice, the PCA is performed on the empirical correlation matrix of the observations, that is, on the standardised data, since the model is invariant by change of scale.
19. They also propose a variant based on the VAR residual correlation matrix. In this case, on the basis of their simulations, the authors recommend using $m = 1.25$ for the first criterion and $m = 2.25$ for the second.
20. The series for new construction (building starts and building permits) and the gross monetary aggregates supplied by the Ministry for Ecology and the Bank of France were seasonally adjusted by applying the X12-ARIMA method to the series in log form.
21. A data interpolation method based on a third-degree polynomial. More specifically, a cubic spline function is a piecewise function defined by polynomials of degree 3. This type of function is often used because of its effectiveness in solving interpolation problems. Note that there is very little difference in the results obtained when the series are converted to monthly values by linear interpolation.
22. The RMSE criterion leads us to choose a less parsimonious representation with $q = 5$, $s = 2$ and $p = 2$. Further on, we study the differences between these two specifications in regard to the out-of-sample results.
23. The t-statistics associated with the factor coefficients can be read in the standard manner. Stock and Watson (2002a), and Bai and Ng (2006) have shown that under certain conditions for N and T , one can use the factors estimated in an auxiliary regression without causing changes in the limit distributions of the estimated parameters. The most commonly used condition is the one given by Bai and Ng (2006), that is, $T \rightarrow +\infty$, $N \rightarrow +\infty$, $\sqrt{T}/N \rightarrow 0$. The orders of magnitude that we have for T and N are customary in this type of study, as is the order of magnitude of $\sqrt{T}/N$ (approximately 0.1 here).

24. We cannot conduct the exercise using the series values that were actually available at the forecast date since, as a rule only the revised values are available.
25. Similarly, the forecast error is calculated from the GDP value currently available. This value differs from the one provided by the “initial results,” which is probably the forecaster’s true target.
26. Recall that Equation (10a) is estimated for two values of h : to forecast the current or previous quarter, we estimate the equation with $h = 0$; and, to forecast the next quarter, we estimate the equation with $h = 1$.
27. This exercise however, does not establish the superiority of factor models to calibrations or combinations of calibrations. We have not explored all the potential explanatory variables in the information set used to estimate either the factor model or the calibration model.
28. At the start of quarter Q-1, we must forecast the factors over a six-month horizon. Consequently, the variance of the error in the forecast obtained with the VAR model becomes substantial.
29. For $h=0$, we specify $q=5$, $s=p=1$ using the information criteria rather than $q=5$, $s=2$, $p=1$ using the RMSE criterion for FM1, $q=4$, $s=p=1$ using the information criteria rather than $q=5$, $s=2$, and $p=3$ for FM3 using the RMSE criterion, but $q=4$, $s=2$, and $p=1$ for FM2 using both methods. For $h=1$, when we use two equations, the RMSE criterion leads us to specify $q=5$, $s=p=2$ for FM1, $q=5$, $s=2$, and $p=3$ for FM2 and FM3. The choice of (q, p, s) according to the information criteria is identical whatever h .

References

- Altissimo, F. et al. (2001), “A Real Time Coincident Indicator for the Euro Area Business Cycle”, CEPR Discussion Paper Series No. 3108, 46 pages.
- Altissimo, F. et al. (2010), “New Eurocoin: Tracking Economic Growth in Real Time”, *The Review of Economics and Statistics*, Vol. 92, No. 4, pp. 1024-1034.
- Angelini, E., M. Bańbura and G. Rünstler (2008), “Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model”, *ECB Working Paper* No. 953, October, 29 pages.
- Angelini, E. et al. (2011), “Short-Term Forecasts of Euro Area GDP Growth”, *Econometrics Journal*, Vol. 14, No. 1, C25-C44.
- Bai, J. and S. Ng (2002), “Determining the Number of Factors in Approximate Factor Models”, *Econometrica*, Vol. 70, No. 1, pp. 191-221.
- Bai, J. and S. Ng (2006), “Confidence Intervals for Diffusion Indexes Forecasts and Inference for Factor-Augmented Regression”, *Econometrica*, Vol. 74, No. 4, pp. 1133-1150.
- Bai, J. and S. Ng (2007), “Determining the Number of Primitive Shocks in Factor Models”, *Journal on Business and Economic Statistics*, Vol. 25, No. 1, pp. 52-60.
- Bai, J. and S. Ng (2008a), “Forecasting Economic Time Series Using Targeted Predictors”, *Journal of Econometrics*, Vol. 146, No. 2, pp. 304-317.
- Bai, J. and S. Ng (2008b), “Large Dimensional Factor Analysis”, *Foundations and Trends in Econometrics*, Vol. 3, No. 2, pp. 89-163.
- Bańbura, M. and G. Rünstler (2011), “A Look into the Factor Model Black Box – Publication Lags and the Role of Hard and Soft Data in Forecasting GDP”, *International Journal of Forecasting*, Vol. 27, No. 2, pp. 333-346.
- Barhoumi, K., O. Darné and L. Ferrara (2010), “Are Disaggregate Data Useful for Factor Analysis in Forecasting French GDP?”, *Journal of Forecasting*, Vol. 29, No. 1-2, pp. 132-144.
- Barhoumi, K. et al. (2012), “Monthly Forecasting of French GDP: A Revised Version of the OPTIM Model”, *Bulletin of Economic Research*, Vol. 64, No. 1, pp. 53-70.
- Bessec, M. (2010), “Étalonnages du taux de croissance du PIB français sur la base d’enquête de conjoncture”, *Économie et Prévision*, No. 193, pp. 77-99.
- Boivin, J. and S. Ng (2006), “Are More Data Always Better for Factor Analysis”, *Journal of Econometrics*, Vol. 132, No. 1, May, pp. 169-194.
- Caggiano, G., G. Kapetanios and V. Labhard (2009), “Are More Data Always Better for Factor Analysis? Results for the Euro Area, the Six Largest Euro Area Countries and the UK”, *Working Paper Series* No. 1051, European Central Bank.
- Charpin, F. (2009), “Estimation précoce de la croissance, de la régression LARS au modèle à facteurs”, *Revue de l’OFCE*, Vol. 2009/1, No. 108, pp. 31-48.

- Darné, O. and V. Brunhes-Lesage (2007), "L'indicateur synthétique mensuel d'activité (ISMA) : une révision", *Note d'Études et de Recherche de la Banque de France*, NER – E#171.
- Diron, M. (2008), "Short-Term Forecasts of Euro Area Real GDP Growth: An Assessment of Real-Time Performance Based on Vintage Data", *Journal of Forecasting*, Vol. 25, No. 5, pp. 371-390.
- Doz, C., D. Giannone and L. Reichlin (2011), "A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering", *Journal of Econometrics*, Vol. 164, No. 1, pp. 188-205.
- Doz, C., D. Giannone and L. Reichlin (2012), "A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models", *Review of Economics and Statistics*, Vol. 94, No. 4, pp. 1014-1024.
- Dubois, É. and E. Michaux (2006), "Étalonnage à l'aide d'enquêtes de conjoncture : de nouveaux résultats", *Économie et Prévision*, No. 172, pp. 11-28.
- Erkel-Rousse, H. and C. Minodier (2009), "Do Business Tendency Surveys in Industry and Services Help in Forecasting GDP Growth? A Real-Time Analysis on French Data", *INSEE Working Paper* No. G2009/03, *Direction des études et synthèses économiques*, 91 pages.
- Forni, M. et al. (2000), "The Generalized Dynamic Factor Model: Identification and Estimation", *Review of Economics and Statistics*, Vol. 82, No. 4, pp. 540-554.
- Forni, M. et al. (2005), "The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting", *Journal of the American Statistical Association*, Vol. 100, No. 471, pp. 830-840.
- Giannone, D., L. Reichlin and D. Small (2008), "Nowcasting GDP and Inflation: The Real-Time Informational Content of Macroeconomic Data Releases", *Journal of Monetary Economics*, Vol. 55, No. 4, pp. 665-676.
- Grenouilleau, D. (2004), "A Sorted Leading Indicator Dynamic (SLID) Factor Model for Short-Run Euro-Area GDP Forecasting", *European Economy*, European Commission DG-ECFIN Economic Paper, No. 219, ISSN 1725-3187, December, 46 pages.
- Grenouilleau, D. (2006), "The Stacked Leading Indicators Dynamic Factor Model: A Sensitivity Analysis of Forecast Accuracy Using Bootstrapping", *European Economy*, European Commission DG-ECFIN Economic Paper No. 249, ISSN 1725-3187, June, 63 pages.
- Hendry, D.F. and H.-M. Krolzig (1999), "Improving on Data Mining Reconsidered by K.D. Hoover and S.J. Perez", *Econometric Journal*, Vol. 2, No. 2, pp. 41-58.
- Hendry, D.F. and H.-M. Krolzig (2001), "Computer Automation of General-to-Specific Model Selection Procedures", *Journal of Economic Dynamic and Control*, Vol. 25, No. 6-7, pp. 831-866.
- Hendry, D.F. and H.-M. Krolzig (2005), "The Properties of Automatic GETS Modelling", *The Economic Journal*, Vol. 115, No. 502, pp. C32-C61.
- Hoover, K.D. and S.J. Perez (1999), "Data Mining Reconsidered: A General to Specific Approach to Specification Search", *Econometrics Journal*, Vol. 2, No. 2, pp. 167-191.
- Irac, D. and F. Sédillot (2002), "Short-Run Assessment of French Economic Activity Using OPTIM", *Note d'Études et de Recherche de la Banque de France*, No. 088, 36 pages.
- Minodier, C. (2010), "Avantages comparés des séries des premières valeurs publiées et des séries des valeurs révisées, un exercice de prévision en temps réel de la croissance trimestrielle du PIB en France", *INSEE Working Paper* No. G2010/01, *Direction des études et synthèses économiques*, 20 pages.
- Pain, N. et al. (2005), "The New International Trade Model", *OECD Economic Department Working Paper*, No. 440, 105 pages.
- Pesaran, M.H. and A. Timmermann (1992), "A Simple Non-Parametric Test of Predictive Performance", *Journal of Business and Economic Statistics*, Vol. 10, No. 4, pp. 461-465.
- Rünstler, G. et al. (2009), "Short-Term Forecasting of GDP Using Large Monthly Data Sets: A Pseudo Real-Time Forecast Evaluation Exercise", *Journal of Forecasting*, Vol. 28, No. 7, pp. 595-611.
- Rünstler, G. and F. Sédillot (2003), "Short-Term Estimates of Euro Area Real GDP by Means of Monthly Data", *Working Paper Series* No. 276, European Central Bank, 28 pages.
- Schumacher, C. (2007), "Forecasting German GDP Using Alternative Factor Models Based on Large Datasets", *Journal of Forecasting*, Vol. 26, No. 4, pp. 271-302.
- Schumacher, C. (2010), "Factor Forecasting Using International Targeted Predictors: The Case of German GDP", *Economic Letters*, Vol. 107, No. 2, pp. 95-98.

- Sédillot, F. and N. Pain (2005), "Indicator Models of Real GDP Growth in the Major OECD Economies", *OECD Economic Studies*, Vol. 40, No. 1, pp. 167-217.
- Stock, J. and M. Watson (1999), "Forecasting Inflation", *Journal of Monetary Economics*, Vol. 44, No. 2, pp. 293-335.
- Stock, J. and M. Watson (2002a), "Forecasting Using Principal Components from a Large Number of Predictors", *Journal of the American Statistical Association*, Vol. 97, No. 460, pp. 1167-1179.
- Stock, J. and M. Watson (2002b), "Macroeconomic Forecasting Using Diffusion Indexes", *Journal of Business and Economic Statistics*, Vol. 20, No. 2, pp. 147-162.
- Stock, J. and M. Watson (2010), "Dynamic Factor Models", in *Oxford Handbook of Economic Forecasting*, M.P. Clements and D.F. Hendry (ed.), Oxford University Press, Chapter 2.

ANNEX A

State-space models and Kalman filter

This appendix presents the state-space representation and the latent variable estimation procedure based on the Kalman filter and smoother. The filter gives the best linear approximation of the latent variable (in our application, the vector of common factors) at time t , given the values of the observable variables at dates up to t . In a second stage, the Kalman smoother gives a more accurate approximation of the latent variable at t , by taking into account the information contained in the whole sample of observable variables.

A.1.1. The state-space model

A state-space model for the multivariate observable process Y_t is given by equations (s) and (m):

$$\begin{cases} Z_{t+1} = AZ_t + \varepsilon_{t+1} & (s) \\ Y_t = CZ_t + \eta_t & (m) \end{cases} \quad \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \approx \text{WN} \left(0, \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix} \right)$$

with $EZ_0 = \mu$, $VZ_0 = P$ and where WN stands for *white noise*. Equation (s) is the state equation and equation (m) is the measurement equation.

The following definitions hold:

Y_t is the measurement variable at time t

Z_t is the state variable at time t

ε_t is the innovations vector at time t

η_t is the vector of measurement errors at time t

A is the transition matrix

C is the measurement matrix

CZ_t is the signal at time t .

Here, we consider the *canonical form* of the system, since we assume that:

$$E(\varepsilon_t \eta_s) = E(\varepsilon_t Z_0) = E(\eta_t Z_0) = 0 \quad \forall t, s = 1, \dots, T$$

Moreover, we assume that A and C do not depend on time.

A.1.2. Kalman filter

The Kalman filter gives, at each time t , an estimation of the state vector conditionally to the observations of Y up to time t , namely $Z_{t|t}^* = E(Z_t | Y_0, \dots, Y_t)$, $t = 1, \dots, T$. This algorithm consists of five computations iterated for $t = 1, \dots, T$:

$$\begin{cases} (a) & Z_{t|t}^* = Z_{t|t-1}^* + K_t (Y_t - CZ_{t|t-1}^*) \\ (b) & \Sigma_{t|t} = (I - K_t C) \Sigma_{t|t-1} \\ (c) & Z_{t+1|t}^* = AZ_{t|t}^* \\ (d) & \Sigma_{t+1|t} = A \Sigma_{t|t} A' + Q \\ (e) & K_t = \Sigma_{t|t-1} C' (C \Sigma_{t|t-1} C' + R)^{-1} \end{cases}$$

with $Z_{t|t}^* = E(Z_t | Y_0, \dots, Y_t)$ the current estimation of the state vector, $Z_{t+1|t}^* = E(Z_{t+1} | Y_0, \dots, Y_t)$ the best linear forecast of Z_{t+1} computed at t , and $\Sigma_{t+1|t} = V(Z_t - Z_{t+1|t}^*)$ the forecast error variance, A' and C' denoting the transpositions of matrices A and C respectively. K_t is the gain matrix. The first two equations update the estimate of Z_t (equation a) and the corresponding forecast error variance (equation b). In equation (a), the current estimate of the state vector is obtained by adding an updating term to the value estimated at time $t-1$. This updating term incorporates the new information contained in the observed value Y_t . This term is computed in such a way that the current forecast error variance is minimised and dependent in a simple way on the estimate of Z_t at time $t-1$, on the observed value Y_t , and on the gain matrix K_t given by (e). Equations (c) and (d) use the state equation to update the state variable estimate at time $t+1$.

It is necessary to know μ and P in order to initialise the algorithm, $Z_{1|0}^* = \mu$ and $\Sigma_{1|0} = P$. It is also crucial to have an *a priori* knowledge of Z_0 for the consistency of the process. When (Z_t) is stationary (as here), μ and P can be explicitly calculated from the parameters of the model.

A.1.3. The smoothing algorithm

In a second step, the smoothing algorithm provides a more accurate estimate of the state vector. Indeed, the best linear forecast is calculated conditionally to all the information (and not only to present and past information). The iterations are made in reverse order for $t = T-1$ to 1.

$$\begin{cases} Z_{t|T}^* = Z_{t|t}^* + F_t (Z_{t+1|T}^* - Z_{t+1|t}^*) & \text{with } F_t = \Sigma_{t|t} A' \Sigma_{t+1|t}^{-1} \\ \Sigma_{t|T} = \Sigma_{t|t} + F_t (\Sigma_{t+1|T} - \Sigma_{t+1|t}) F_t' \end{cases}$$

The algorithm is initialised at time T using the results of the filter.

All the previous calculations (filtering and smoothing) are made for given values of A , C , Q , and R . If these matrices are not known, we can calculate the model's log-likelihood under a Gaussian assumption, using the iterations of the filter, and we can estimate the unknown parameters by maximum likelihood. In our application, the parameters estimated in the first step of our 2-step procedure serve as A and C , and the estimated variance matrices of the associated residuals serve as Q and R .

ANNEX B

Statistical Tables

Table B.1. **Methods for forecasting French and EU GDP growth**

Institution	Country ¹	Horizon	Model	Explanatory variables	Reduction of the information set ²	References
Directorate General of the Treasury (France)	FR	$h = -1, 0, 1$	“Garde Fou” [literally: Safeguard] macro-sectoral model: tool for ensuring macro-sectoral consistency of forecasts of individual aggregates prepared by means of standard calibration methods	Survey variables: industry, services, retail trade, wholesale trade, and construction surveys by INSEE; Bank of France retail survey Real variables: Industrial Production Index (IPI), consumption of manufactured goods, new car registrations, building starts, building permits, power consumption, deliveries of fuel and oil, customs foreign-trade statistics, trade turnover, temperature, CNAM (National Health Insurance Fund) data on health expenditure, etc. Financial variable: €/€ exchange rate	GETS, information criteria or <i>a priori</i> choice of variables following the method used to prepare quarterly accounts	
	FR	$h = -1, 0, 1$	Calibration model (bridge model) for GDP growth	INSEE survey variables: Industry, services, and construction	GETS applied to blocks of variables	Bessec (2010)
	FR	$h = 0$	Calibration model for GDP growth	Composite Purchasing Managers' Index (PMI)	<i>A priori</i> choice	
INSEE	FR	$h = -1, 0, 1, 2$	Macro-sectoral model	Survey variables Real variables	GETS, information criteria or choice of variables following the method used to prepare quarterly accounts	Erkel-Rousse and Minodier (2009)
	FR	$h = -1, 0, 1, 2$	Calibration model for GDP growth	INSEE survey variables (industry and services)	GETS	

Table B.1. **Methods for forecasting French and EU GDP growth (cont.)**

Institution	Country ¹	Horizon	Model	Explanatory variables	Reduction of the information set ²	References
Bank of France	FR	$h = 0, 1$	Macro-sectoral model (OPTIM model)	INSEE (industry, services, construction, retail trade, and consumers -households-), Bank of France (industry, services, and retail trade) and European Commission survey variables Real variables: IPI, consumption of manufactured goods, prices, power consumption, building starts, customs foreign-trade statistics Financial variable: €/€ exchange rate	GETS	Irac and Sédillot (2002) Barhoumi et al. (2012)
	FR	$h = 0, 1$	Calibration for GDP growth (ISMA model)	Bank of France survey opinion balances in industry and main production and service sectors	PCA then GETS	Darné and Brunhes-Lesage (2007)
	FR	$h = 1$	Factor model for GDP growth	Survey variables: industry and services (source: Bank of France), retail trade and households (<i>Source</i> : INSEE). Real variables: IPI, consumption of manufactured goods, new car registrations, sales of industrial vehicles, building starts, imports and exports Financial variables: French stock market index, short- and long-term interest rates, mortgage interest rates, oil price, consumer prices	Dynamic factor model estimated using 4 alternative methods (SW, DGR 2011, DGR 2012, FHLR) Choice of number of factors: Bai and Ng (2002, 2007) or <i>a priori</i> choice	Barhoumi, Darné and Ferrara (2010)
	FR, BE, DE, IT, NL, PT, LT, HU, PL and EA	$h = -1, 0, 1$	Factor model for GDP growth	Survey, real and financial variables, prices	Dynamic factor model estimated using 3 alternative methods (SW, DGR 2011, FHLR) Choice of number of factors: RMSE criterion	Rüntler et al. (2009)
	FR, EA, US, DE, IT, UK, JP	$h = 0, 1$	Calibration model for GDP growth	INSEE survey variables: Industry and households Real variables: IPI, construction IPI, consumption of manufactured goods, retail sales, new car registrations, monthly exports and imports, unemployment rate Financial variables: Short- and long-term interest rates, real exchange rates	Pre-selection of indicators according to their explanatory power as measured by R^2 coefficient from bivariate regressions between GDP growth and each individual series Selection of best combination of indicators according to the Schwarz criterion	Sédillot and Pain (2005)
European Commission	EA	$h = 0, 1, 2$	Factor model for GDP growth (SLID model) (before 2007)	Survey variables: Industry, services, construction, retail trade, and households (source: European Commission), PMI Real variables: IPI, retail sales and wholesale trade, nominal monthly foreign trade figures, building starts, new car registrations, variables on labour market and start-ups, business failures Financial variables: Prices, monetary aggregates, stock market indexes, nominal exchange rates, term structure, implied volatility index of CAC40, DAX20, Eurostoxx50, and AEX These are sectoral variables for each country	Dynamic factor model estimated using PCA (SW) Choice of number of factors: information criteria	Grenouilleau (2004, 2006)

Table B.1. **Methods for forecasting French and EU GDP growth (cont.)**

Institution	Country ¹	Horizon	Model	Explanatory variables	Reduction of the information set ²	References
European Central Bank	EA	$h = -1, 0, 1$	Combination of calibration models for GDP growth	Survey variables: Industry, retail trade, households and services Real variables: IPI, construction IPI, retail sales and new car registrations Composite indexes: European Commission economic sentiment index (TSEN), CEPR EuroCOIN, and OECD advanced composite indicator for euro area Financial variables: Stock market index, spread, effective exchange rates in euro area in Diron (2008)	Combination of forecasts from standard calibration models with a selection of regressors relying on <i>Stepwise regressions</i> in Rünstler and Sédillot (2003) and relying on an algorithm which estimates all possible combinations in the set of candidate indicators and ranks them according to out-of-sample RMSE for some of the calibration equations in Diron (2008)	Diron (2008), Rünstler and Sédillot (2003)
	EA	$h = -1, 0, 1$	Factor model for GDP growth	Survey variables: Industry, households, construction, retail trade, and two US survey opinion balances on business and consumer expectations Real variables: IPI and sectoral IPIs, exports and imports within and outside the euro area, retail sales, new car registrations, employment, US indicators (IPI, retail sales, unemployment, employment) Financial variables: Nominal effective exchange rate, real effective exchange rates, bilateral exchange rates, commodity prices, gold price, oil price, European and US stock market indexes, European and US interest rates	Dynamic factor model estimated following DGR with a new measure of the weight of individual series in forecast Choice of number of factors: RMSE criterion	Bañbura and Rünstler (2011)
	EA	$h = -1, 0, 1$	Factor model for GDP growth	European Commission survey variables (industry, construction, retail trade, and consumers) Real variables: IPI and sectoral IPIs, employment and unemployment, exports and imports, retail sales and new car registrations Financial variables: Exchange rate, interest rates and stock market indexes, monetary aggregates, commodity prices International variables: Macroeconomic indicators	Dynamic factor model estimated following DGR and estimation of the contributions of variables to forecast following Bañbura and Rünstler (2011) Univariate and multivariate models (GDP and its components) with the balance-sheet identity constraint from the national accounts introduced in the state-space representation Choice of number of factors: Bai and Ng (2002) criteria, choice of number of lags in models: BIC criterion	Angelini, Bañbura and Rünstler (2008)
	FR and EA, BE, DE, UK, IT, NL, ES	$h = 1, \dots, 12$	Factor model for GDP growth	Survey variables Real variables: IPI and its components, new car registrations, consumption of manufactured goods, export, imports, employment and unemployment Financial variables: Monetary aggregates, exchange rates, stock market indexes and interest rates, Harmonised Index of Consumer Prices (HICP) and HICP components, oil price, gold price, producer price indexes International variables: US indicators (IPI, employment, prices, retail sales, etc.)	Dynamic factor model estimated following SW with pre-selection of explanatory variables following Boivin and Ng (2006) and alternative rules for factor selection and then, optionally, combination of forecasts using selected factor models	Caggiano, Kapetanios, and Labhard (2009)

Notes:

1. BE = Belgium, FR = France, EA = Euro area, DE = Germany, HU = Hungary, IT = Italy, JP = Japan, LT = Lithuania, NL = Netherlands, PL = Poland, PT = Portugal, ES = Spain, UK = United Kingdom, US = United States.
2. GETS = General to Specific algorithm, PCA = Principal Component Analysis, SW = Stock and Watson (2002a, 2002b), DGR = Doz, Giannone and Reichlin (2011, unless otherwise stated), FHLR = Forni, Hallin, Lippi and Reichlin (2000, 2005), RMSE = Root Mean Square Error, BIC = Bayesian Information Criterion.

Table B.2. Data source details

Type	Sector	Variable	Periodicity ¹	Start of the series	Timeliness	Source
Survey	Industry	Past production	M	Apr. 76	M + 0	a
	Industry	Inventories	M	Apr. 76	M + 0	a
	Industry	Overall order books	M	Apr. 76	M + 0	a
	Industry	Export order books	M	Apr. 76	M + 0	a
	Industry	Personal production outlook	M	Apr. 76	M + 0	a
	Industry	General production outlook	M	Apr. 76	M + 0	a
	Services	Past activity	Q/M	1988Q1-2000Q2 then 06/2000	M + 0	a
	Services	Expected activity	Q/M	1988Q1-2000Q2 then 06/2000	M + 0	a
	Services	Expected demand	Q/M	1988Q1-2000Q2 then 06/2000	M + 0	a
	Construction	Past activity	Q/M	1975Q1-1993Q3 then 09/1993	M + 0	a
	Construction	Expected activity	Q/M	1975Q1-1993Q3 then 09/1993	M + 0	a
	Construction	Past employment	Q/M	1975Q2-1993Q3 then 09/1993	M + 0	a
	Construction	Order books	Q/M	1975Q1-1993Q3 then 09/1993	M + 0	a
	Construction	Productive capacity utilisation rate	Q/M	1975Q2-1993Q3 then 09/1993	M + 0	a
	Retail trade	General business outlook	M	Jan. 91	M + 0	a
	Retail trade	Past sales	M	Jan. 91	M + 0	a
	Retail trade	Orders	M	Jan. 91	M + 0	a
	Retail trade	Expected workforce	M	Jan. 91	M + 0	a
	Consumers	Personal financial situation – recent changes	M	Jan. 70	M + 0	a
	Consumers	Personal financial situation – expected change	M	Jan. 70	M + 0	a
	Consumers	Standard of living in France – recent changes	M	Jan. 70	M + 0	a
	Consumers	Standard of living in France – expected change	M	Jan. 70	M + 0	a
	Consumers	Right time to make major purchases	M	Jan. 70	M + 0	a
	Manufacturing	Demand – recent changes	Q	1976Q2	Q+0	a
	Manufacturing	Demand – expected change	Q	1976Q2	Q+0	a
	Services	Past operating income	Q	1988Q1	Q+0	a
	Services	Expected operating income	Q	1988Q1	Q+0	a
Real	Households	New private vehicle registrations	M	Jan. 85	M + 1	a
	Households	Consumption of manufactured goods	M	Jan. 85	M + 1	a
	Households	Expenditure on durable goods	M	Jan. 85	M + 1	a
	Households	Expenditure on cars	M	Jan. 85	M + 1	a
	Households	Purchases of household durables	M	Jan. 85	M + 1	a
	Households	Consumption of textiles	M	Jan. 85	M + 1	a
	Households	Expenditure on other manufactured goods	M	Jan. 85	M + 1	a
	Households	Retail sales excluding cars	M	Jan. 95	M + 2	b
	Households	Unemployment rate	M	Jan. 83	M + 2	b
	Households	Unemployment rate for the under-25s	M	Jan. 83	M + 2	b
	Households	Job vacancies	M	Jan. 89	M + 2	c
	Construction	Total building starts	M	Jan. 94	M + 1	d
	Construction	Building starts (collective housing)	M	Jan. 94	M + 1	d
	Construction	Building starts (individual housing)	M	Jan. 94	M + 1	d
	Construction	Building starts (residential)	M	Jan. 94	M + 1	d
	Construction	Building permits (total)	M	Jan. 94	M + 1	d
	Construction	Building permits (collective housing)	M	Jan. 94	M + 1	d
	Construction	Building permits (individual housing)	M	Jan. 94	M + 1	d
	Industry	IPI – total industry (BE) ²	M	Jan. 90	M + 2	a
	Industry	IPI – manufacturing (CZ) ²	M	Jan. 90	M + 2	a
	Industry	IPI – agricultural and food industries (C1) ²	M	Jan. 90	M + 2	a
	Industry	IPI – coking and refining (C2) ²	M	Jan. 90	M + 2	a
	Industry	IPI – electrical and electronic equipment (C3) ²	M	Jan. 90	M + 2	a
	Industry	IPI – cars (CL1) ²	M	Jan. 90	M + 2	a
	Industry	IPI – transport equipment excluding cars (CL2) ²	M	Jan. 90	M + 2	a
	Industry	IPI – other manufactured products (C5) ²	M	Jan. 90	M + 2	a
	Industry	IPI – mining and quarrying, energy and water supply (DE)	M	Jan. 90	M + 2	a
	Industry	IPI – construction (F)	M	Jan. 90	M + 2	a

Table B.2. **Data source details** (cont.)

Type	Sector	Variable	Periodicity ¹	Start of the series	Timeliness	Source
Nominal	Stock market	CAC 40	M	Jul. 87	M + 0	e
	Stock market	SP 500	M	Jan. 80	M + 0	f
	Stock market	FTSE	M	Jan. 84	M + 0	g
	Stock market	DAX	M	Sep. 59	M + 0	h
	Stock market	Eurostoxx 50	M	Dec. 86	M + 0	i
	Stock market	Nikkei	M	Jan. 81	M + 0	j
	Stock market	Price Earnings Ratio (USA)	M	Jan. 54	M + 1	f
	Stock market	Volatility index (Vix)	M	Jan. 90	M + 0	k
	Money	M1	M	Jan. 80	M + 2	l
	Money	M2	M	Jan. 80	M + 2	l
	Money	M3	M	Jan. 80	M + 2	l
	Money	Loans (value)	M	Jan. 80	M + 2	l
	Interest rate	Mortgage interest rates	M	Feb. 78	M + 1	l
	Interest rate	3-month interest rate	M	Jan. 70	M + 1	m
	Interest rate	1-year interest rate	M	Jan. 60	M + 1	l
	Interest rate	10-year interest rate	M	Jan. 89	M + 0	n
	Interest rate	Term structure (France)	M	Jan. 90	M + 0	o
	Interest rate	Term structure (USA)	M	Jan. 90	M + 0	o
	Price	Gold	M	Jan. 79	M + 0	p
	Price	Oil	M	Jan. 80	M + 1	m
	Price	Commodities	M	Jan. 80	M + 1	m
	Price	Consumer Price Index (CPI)	M	Jan. 90	M + 1	a
International	Exchange rate	euro/dollar	M	Jan. 78	M + 0	q
	Exchange rate	euro/sterling	M	Dec. 77	M + 0	q
	Exchange rate	euro/yen	M	Dec. 78	M + 0	q
	Exchange rate	euro/yuan	M	Dec. 78	M + 0	q
	Exchange rate	Nominal effective exchange rate for the euro	M	Jan. 79	M + 2	m
	Exchange rate	Real effective exchange rate for the euro	M	Dec. 79	M + 2	m
	Germany	Assessment of present business situation (IFO)	M	Jan. 91	M + 0	r
	Germany	Business expectations, next 6 months (IFO)	M	Jan. 91	M + 0	r
	Germany	Current economic situation (ZEW)	M	Dec-91	M + 0	s
	Germany	Business expectations (ZEW)	M	Dec. 91	M + 0	s
	Germany	IPI – manufacturing	M	Jan. 91	M + 2	t
	USA	Retail sales (value)	M	Jan. 92	M + 1	u
	USA	IPI – manufacturing	M	Jan. 72	M + 1	v
	USA	Employment	M	Jan. 48	M + 1	w
	USA	Unemployment rate	M	Jan. 48	M + 1	w
	USA	Manufacturing PMI	M	Jan. 48	M + 1	x

Notes:

1. Q = quarterly; M = monthly.

2. IPI (C1) (or C2, C3, C5, in that order) represents the industrial production index in the following activities of the aggregated classification (NA): C1 = Manufacture of food products and beverages, C2 = Manufacture of coke and refined petroleum products, C3 = Electrical and electronic equipment; machine equipment, C5 = other manufacturing (excluding transport). CL1 = Manufacture of motor vehicles, trailers, and semi-trailers, CL2 = Manufacture of other transport equipment.

Sources: a) INSEE; b) Eurostat; c) OECD; d) French Ministry of Ecology; e) Nyse Euronext Paris; f) Standard & Poor's; g) FTSE; h) Frankfurt SE; i) Financial Times; j) OSE; k) CBOE; l) Bank of France; m) IMF; n) Daily press; o) Global insight; p) LBMA; q) ECB; r) IFO; s) ZEW; t) DESTATIS; u) Census Bureau; v) Federal Reserve Board; w) BLS; x) ISM.

Table B.3. Unit root test results

Type	Sector	Series	ADF ¹	PP ¹	KPSS ¹	Transformation ²
Survey	Industry	Past production	I(0)	I(0)	I(0)	-
	Industry	Inventories	I(0)	I(0)	I(0)	-
	Industry	Overall order books	I(0)	I(0)	I(0)	-
	Industry	Export order books	I(0)	I(0)	I(0)	-
	Industry	Personal production outlook	I(0)	I(0)	I(0)	-
	Industry	General production outlook	I(0)	I(0)	I(0)	-
	Services	Past activity	I(0)	I(0)	I(0)	-
	Services	Expected activity	I(0)	I(0)	I(0)	-
	Services	Expected demand	I(0)	I(0)	I(0)	-
	Construction	Past activity	I(0)	I(1)	I(0)	-
	Construction	Expected activity	I(0)	I(1)	I(1)	-
	Construction	Past employment	I(1)	I(1)	I(0)	-
	Construction	Order books	I(1)	I(1)	I(1)	1-L
	Construction	Productive capacity utilisation rate	I(1)	I(1)	I(1)	1-L
	Retail trade	General business outlook	I(1)	I(1)	I(0)	-
	Retail trade	Past sales	I(0)	I(0)	I(0)	-
	Retail trade	Orders	I(0)	I(0)	I(1)	-
	Retail trade	Expected workforce	I(0)	I(0)	I(0)	-
	Consumers	Personal financial situation – recent changes	I(0)	I(0)	I(0)	-
	Consumers	Personal financial situation – expected change	I(0)	I(0)	I(0)	-
	Consumers	Standard of living in France – recent changes	I(1)	I(1)	I(0)	-
	Consumers	Standard of living in France – expected change	I(0)	I(0)	I(0)	-
	Consumers	Right time to make major purchases	I(1)	I(1)	I(1)	1-L
	Industry	Demand – recent changes	I(0)	I(0)	I(0)	-
	Industry	Demand – expected change	I(0)	I(0)	I(0)	-
	Services	Past operating income	I(0)	I(0)	I(0)	-
	Services	Expected operating income	I(0)	I(0)	I(0)	-
Real	Households	New private vehicle registrations	I(0)	I(0)	I(0)	(1-L)log
	Households	Consumption of manufactured goods	I(1)	I(1)	I(1)	(1-L)log
	Households	Expenditure on durable goods	I(1)	I(1)	I(1)	(1-L)log
	Households	Expenditure on cars	I(0)	I(0)	I(1)	(1-L)log
	Households	Purchases of household durables	I(1)	I(1)	I(1)	(1-L)log
	Households	Consumption of textiles	I(1)	I(0)	I(1)	(1-L)log
	Households	Expenditure on other manufactured goods	I(1)	I(1)	I(1)	(1-L)log
	Households	Retail sales excluding cars	I(1)	I(1)	I(1)	(1-L)log
	Households	Unemployment rate	I(1)	I(1)	I(0)	1-L
	Households	Unemployment rate for the under-25s	I(1)	I(1)	I(0)	1-L
	Households	Job vacancies	I(1)	I(1)	I(1)	1-L
	Construction	Total building starts	I(0)	I(0)	I(1)	(1-L)log
	Construction	Building starts (collective housing)	I(1)	I(0)	I(1)	(1-L)log
	Construction	Building starts (individual housing)	I(1)	I(0)	I(1)	(1-L)log
	Construction	Building starts (residential)	I(0)	I(0)	I(1)	(1-L)log
	Construction	Building permits (total)	I(1)	I(0)	I(1)	(1-L)log
	Construction	Building permits (collective housing)	I(1)	I(0)	I(1)	(1-L)log
	Construction	Building permits (individual housing)	I(1)	I(1)	I(1)	(1-L)log
	Industry	IPI – total industry (BE)	I(1)	I(1)	I(1)	(1-L)log
	Industry	IPI – manufacturing (CZ)	I(1)	I(1)	I(1)	(1-L)log
	Industry	IPI – agricultural and food industries (C1)	I(1)	I(1)	I(1)	(1-L)log
	Industry	IPI – coking and refining (C2)	I(0)	I(0)	I(1)	(1-L)log
	Industry	IPI – electrical and electronic equipment (C3)	I(1)	I(1)	I(1)	(1-L)log
	Industry	IPI – cars (CL1)	I(1)	I(1)	I(1)	(1-L)log
	Industry	IPI – transport equipment excluding cars (CL2)	I(1)	I(0)	I(1)	(1-L)log
	Industry	IPI – other manufactured products (C5)	I(1)	I(1)	I(1)	(1-L)log
	Industry	IPI – mining and quarrying, energy and water supply (DE)	I(0)	I(0)	I(1)	(1-L)log
	Industry	IPI – construction (F)	I(1)	I(0)	I(1)	(1-L)log

Table B.3. **Unit root test results** (cont.)

Type	Sector	Series	ADF ¹	PP ¹	KPSS ¹	Transformation ²
Nominal	Stock exchange	CAC 40	I(1)	I(1)	I(1)	(1-L)log
	Stock exchange	SP 500	I(1)	I(1)	I(1)	(1-L)log
	Stock exchange	FTSE	I(1)	I(1)	I(1)	(1-L)log
	Stock exchange	DAX	I(1)	I(1)	I(1)	(1-L)log
	Stock exchange	Eurostoxx 50	I(1)	I(1)	I(1)	(1-L)log
	Stock exchange	Nikkei	I(1)	I(1)	I(1)	(1-L)log
	Stock exchange	Price Earnings Ratio (USA)	I(1)	I(1)	I(1)	(1-L)log
	Stock exchange	Volatility index (Vix)	I(0)	I(0)	I(0)	log
	Money	M1	I(1)	I(1)	I(1)	(1-L)log
	Money	M2	I(1)	I(1)	I(1)	(1-L)log
	Money	M3	I(1)	I(1)	I(1)	(1-L)log
	Money	Loans (value)	I(0)	I(0)	I(2)	(1-L)log
	Interest rate	Mortgage interest rates	I(1)	I(1)	I(1)	(1-L)
	Interest rate	3-month interest rate	I(1)	I(1)	I(1)	(1-L)
	Interest rate	1-year interest rate	I(1)	I(1)	I(1)	(1-L)
	Interest rate	10-year interest rate	I(1)	I(1)	I(1)	(1-L)
	Interest rate	Term structure (France)	I(0)	I(0)	I(0)	-
	Interest rate	Term structure (USA)	I(1)	I(1)	I(0)	-
	Price	Gold	I(1)	I(1)	I(1)	(1-L)log
	Price	Oil	I(1)	I(1)	I(1)	(1-L)log
	Price	Commodities	I(1)	I(1)	I(1)	(1-L)log
	Price	Consumer Price Index (CPI)	I(1)	I(1)	I(1)	(1-L)log
International	Exchange rate	euro/dollar	I(1)	I(1)	I(0)	(1-L)log
	Exchange rate	euro/sterling	I(1)	I(1)	I(0)	(1-L)log
	Exchange rate	euro/yen	I(0)	I(0)	I(1)	(1-L)log
	Exchange rate	euro/yuan	I(1)	I(1)	I(1)	(1-L)log
	Exchange rate	Nominal effective exchange rate for the euro	I(1)	I(1)	I(1)	(1-L)log
	Exchange rate	Real effective exchange rate for the euro	I(1)	I(1)	I(0)	(1-L)log
	Germany	Assessment of present business situation (IFO)	I(0)	I(0)	I(0)	-
	Germany	Business expectations, next 6 months (IFO)	I(0)	I(0)	I(0)	-
	Germany	Current economic situation (ZEW)	I(0)	I(0)	I(0)	-
	Germany	Business expectations (ZEW)	I(0)	I(0)	I(0)	-
	Germany	IPI – manufacturing	I(1)	I(1)	I(1)	(1-L)log
	USA	Retail sales (value)	I(1)	I(1)	I(1)	(1-L)log
	USA	IPI – manufacturing	I(1)	I(1)	I(1)	(1-L)log
	USA	Employment	I(1)	I(1)	I(1)	(1-L)
	USA	Unemployment rate	I(1)	I(1)	I(1)	(1-L)
	USA	Manufacturing PMI	I(0)	I(0)	I(0)	-

Notes:

1. The unit root tests are ADF (Augmented Dickey-Fuller), PP (Phillips-Perron) and KPSS (Kwiatkowski-Phillips-Schmidt-Shin) tests.

I(0) = stationary series, I(1) = series integrated of order 1.

2. L denotes the lag operator, log denotes the logarithm function.

Sources: See Table A.2.

Table B.4. **Estimation results for full sample**

Forecast quarter	$Q-1$ and Q	$Q+1$
Estimated equation	Equation (10a) with $h = 0$	Equation (10a) with $h = 1$
Intercept	0.41 (14.1)	0.41 (13.0)
F1	0.10 (13.4)	0.06 (7.5)
F2	0.11 (7.8)	0.14 (8.8)
F3	-0.06 (-2.1)	-
F5	-	-0.22 (-5.3)
F9	-0.36 (-2.8)	-
F10	-0.20 (-2.5)	0.40 (5.2)
R^2	0.78	0.74
	0.77	0.73
LM(4)	0.13	0.39
Heteroskedasticity	0.93	0.27
JB	0.87	0.64

Notes:

The T-statistics are given in parentheses to the right of the estimated coefficient. The last lines provide the P-values of the Breusch-Godfrey test for autocorrelation of order 4 (LM(4)), the Breusch-Pagan-Godfrey test for heteroskedasticity and the Jarque-Bera test for normality (JB). For each factor, the number corresponds to the number of the associated eigenvalue when the eigenvalues are ranked in decreasing order (only the significant factors are kept in the regressions).

Sources: See Table A.2.

Table B.5a. **Forecast evaluation: Root mean square forecast errors (RMSFE)**

Forecasts made one quarter ahead

Method 1: Factors forecast using a VAR model									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-7 (<i>Q</i> -1)	0.58	0.59	0.54	0.56	0.58	0.60	0.60	<i>0.51</i>	<i>0.51</i>
-6 (<i>Q</i> -1)	0.58	0.56	0.49	0.55	0.57	0.59	0.59	<i>0.47</i>	0.50
-5 (<i>Q</i> -1)	0.58	0.56	0.47	0.50	0.53	0.55	0.55	<i>0.41</i>	<i>0.41</i>
Method 2: Two forecast equations without extrapolating the factors over <i>Q</i>									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-7 (<i>Q</i> -1)	0.58	0.59	<i>0.54</i>	0.77	1.06	0.78	0.75	0.56	0.69
-6 (<i>Q</i> -1)	0.58	0.56	0.49	0.49	0.62	0.55	0.57	<i>0.46</i>	0.51
-5 (<i>Q</i> -1)	0.58	0.56	0.47	0.41	0.43	0.54	0.51	<i>0.39</i>	0.41
Method 3: Two forecast equations and extrapolation of the factors over <i>Q</i>									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-7 (<i>Q</i> -1)	0.58	0.59	0.54	0.54	0.58	0.56	0.51	<i>0.46</i>	0.51
-6 (<i>Q</i> -1)	0.58	0.56	0.49	0.52	0.55	0.54	0.54	<i>0.44</i>	0.50
-5 (<i>Q</i> -1)	0.58	0.56	0.47	0.41	0.43	0.54	0.51	<i>0.39</i>	0.41
N.B. methods 2 and 3 are equivalent in -5.									
Forecasts made during the quarter to be forecast at horizons where methods 1 and 3 are equivalent									
Methods 1 and 3									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-4 (<i>Q</i>)	0.58	0.56	0.43	0.47	0.51	0.50	0.50	<i>0.38</i>	<i>0.38</i>
-3 (<i>Q</i>)	0.58	0.47	0.40	0.40	0.47	0.41	0.41	<i>0.34</i>	0.37
Method 2									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-4 (<i>Q</i>)	0.58	0.56	0.43	0.68	0.80	0.73	0.73	<i>0.41</i>	0.47
-3 (<i>Q</i>)	0.58	0.47	0.40	0.42	0.50	0.45	0.45	<i>0.35</i>	0.41
Forecasts made during the quarter to be forecast at horizons where the three methods are equivalent									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-2 (<i>Q</i>)	0.58	0.47	0.38	0.39	0.46	<i>0.34</i>	<i>0.34</i>	0.37	0.41
-1 (<i>Q</i> +1)	0.58	0.47	0.28	0.24	0.24	<i>0.19</i>	<i>0.19</i>	0.29	0.23
0 (<i>Q</i> +1)	0.58	0.47	0.26	0.22	0.21	<i>0.19</i>	<i>0.19</i>	0.29	0.23

Notes:

FM1: Model with factors extracted from the 93 variables.

FM2: Model with factors extracted from survey and real variables.

FM3: Model with factors extracted from survey, nominal, and international variables.

Calibration model: Calibration model relating GDP growth to business climate in France and industrial production index.

Information criteria/RMSE: Factor model derived from the specification obtained using the information criteria/RMSE respectively.

In quarter *Q*-1, seven months before the first release of GDP for quarter *Q*, the factor model FM1 derived from the RMSE criterion yields forecasts of GDP growth in quarter *Q* whose RMSFE is equal to 0.58.

The lowest RMSFE obtained by each method for each horizon are shown in italics.

Sources: See Table A.2.

Table B.5b. **Forecast evaluation: Mean absolute forecast errors (MAFE) (cont.)**

Forecasts made one quarter ahead

Method 1: Factors forecast using a VAR model									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-7 (Q-1)	0.40	0.40	0.38	0.39	0.40	0.44	0.44	0.37	0.37
-6 (Q-1)	0.40	0.40	0.34	0.38	0.41	0.44	0.44	0.34	0.35
-5 (Q-1)	0.40	0.40	0.33	0.36	0.38	0.40	0.40	0.31	0.32
Method 2: Two forecast equations without extrapolating the factors over Q									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-7 (Q-1)	0.40	0.40	0.38	0.56	0.74	0.57	0.58	0.39	0.50
-6 (Q-1)	0.40	0.40	0.34	0.35	0.44	0.40	0.45	0.31	0.35
-5 (Q-1)	0.40	0.40	0.33	0.31	0.33	0.39	0.37	0.29	0.32
Method 3: Two forecast equations and extrapolation of the factors over Q									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-7 (Q-1)	0.40	0.40	0.38	0.39	0.45	0.41	0.39	0.32	0.36
-6 (Q-1)	0.40	0.40	0.34	0.36	0.40	0.40	0.40	0.30	0.34
-5 (Q-1)	0.40	0.40	0.33	0.31	0.33	0.39	0.37	0.29	0.32
N.B. methods 2 and 3 are equivalent in -5.									
Forecasts made during the quarter to be forecast at horizons where methods 1 and 3 are equivalent									
Methods 1 and 3									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-4 (Q)	0.40	0.40	0.32	0.34	0.38	0.37	0.37	0.29	0.29
-3 (Q)	0.40	0.35	0.29	0.30	0.36	0.30	0.30	0.26	0.28
Method 2									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-4 (Q)	0.40	0.40	0.32	0.45	0.57	0.52	0.52	0.33	0.33
-3 (Q)	0.40	0.35	0.29	0.32	0.36	0.33	0.33	0.28	0.31
Forecasts made during the quarter to be forecast at horizons where the three methods are equivalent									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1	FM1	FM2	FM2	FM3	FM3
				Information criteria	RMSE	Information criteria	RMSE	Information criteria	RMSE
-2 (Q)	0.40	0.35	0.29	0.29	0.33	0.28	0.28	0.28	0.30
-1 (Q+1)	0.40	0.35	0.22	0.18	0.19	0.15	0.15	0.22	0.18
0 (Q+1)	0.40	0.35	0.20	0.17	0.17	0.15	0.15	0.22	0.19

Notes:

FM1: Model with factors extracted from the 93 variables.

FM2: Model with factors extracted from survey and real variables.

FM3: Model with factors extracted from survey, nominal, and international variables.

Calibration model: Calibration model relating GDP growth to business climate in France and industrial production index.

Information criteria/RMSE: Factor model derived from the specification obtained using the information criteria/RMSE respectively.

In quarter Q-1, seven months before the first release of GDP for quarter Q, the factor model FM1 derived from the information criteria yields forecasts of GDP growth in quarter Q whose MAFE is equal to 0.39.

The lowest MAFE obtained by each method for each horizon are shown in italics.

Sources: See Table A.2.

Table B.5c. **Forecast evaluation: Pesaran and Timmermann test (cont.)**

Forecasts made one quarter ahead									
Method 1: Factors forecast using a VAR model									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1 Information criteria	FM1 RMSE	FM2 Information criteria	FM2 RMSE	FM3 Information criteria	FM3 RMSE
-7 (<i>Q</i> -1)	0.50 (0.50)	0.50 (0.50)	0.55 (0.26)	0.58 (0.16)	0.55 (0.26)	0.47 (0.63)	0.47 (0.63)	0.58 (0.16)	0.63 (0.05)
-6 (<i>Q</i> -1)	0.50 (0.50)	0.58 (0.16)	0.55 (0.26)	0.66 (0.02)	0.61 (0.09)	0.50 (0.50)	0.50 (0.50)	0.66 (0.02)	0.63 (0.05)
-5 (<i>Q</i> -1)	0.50 (0.50)	0.58 (0.16)	0.47 (0.63)	0.55 (0.26)	0.61 (0.09)	0.58 (0.16)	0.58 (0.16)	0.68 (0.01)	0.63 (0.05)
Method 2: Two forecast equations without extrapolating the factors over <i>Q</i>									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1 Information criteria	FM1 RMSE	FM2 Information criteria	FM2 RMSE	FM3 Information criteria	FM3 RMSE
-7 (<i>Q</i> -1)	0.50 (0.50)	0.50 (0.50)	0.55 (0.26)	0.66 (0.02)	0.53 (0.37)	0.68 (0.01)	0.53 (0.37)	0.79 (0.00)	0.66 (0.02)
-6 (<i>Q</i> -1)	0.50 (0.50)	0.58 (0.16)	0.55 (0.26)	0.68 (0.01)	0.61 (0.09)	0.60 (0.09)	0.47 (0.63)	0.76 (0.00)	0.68 (0.01)
-5 (<i>Q</i> -1)	0.50 (0.50)	0.58 (0.16)	0.47 (0.63)	0.63 (0.05)	0.61 (0.09)	0.50 (0.50)	0.47 (0.63)	0.71 (0.00)	0.66 (0.02)
Method 3: Two forecast equations and extrapolation of the factors over <i>Q</i>									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1 Information criteria	FM1 RMSE	FM2 Information criteria	FM2 RMSE	FM3 Information criteria	FM3 RMSE
-7 (<i>Q</i> -1)	0.50 (0.50)	0.50 (0.50)	0.55 (0.26)	0.58 (0.16)	0.53 (0.37)	0.58 (0.16)	0.53 (0.37)	0.71 (0.00)	0.53 (0.37)
-6 (<i>Q</i> -1)	0.50 (0.50)	0.58 (0.16)	0.55 (0.26)	0.68 (0.01)	0.55 (0.26)	0.53 (0.37)	0.37 (0.95)	0.71 (0.00)	0.58 (0.16)
-5 (<i>Q</i> -1)	0.50 (0.50)	0.58 (0.16)	0.47 (0.63)	0.63 (0.05)	0.61 (0.09)	0.50 (0.50)	0.47 (0.63)	0.71 (0.00)	0.66 (0.02)
N.B. Methods 2 and 3 are equivalent in -5.									
Forecasts made during the quarter to be forecast at horizons where methods 1 and 3 are equivalent									
Methods 1 and 3									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1 Information criteria	FM1 RMSE	FM2 Information criteria	FM2 RMSE	FM3 Information criteria	FM3 RMSE
-4 (<i>Q</i>)	0.50 (0.50)	0.58 (0.16)	0.58 (0.16)	0.55 (0.26)	0.50 (0.50)	0.47 (0.63)	0.47 (0.63)	0.61 (0.09)	0.63 (0.05)
-3 (<i>Q</i>)	0.50 (0.50)	0.32 (0.99)	0.50 (0.50)	0.58 (0.16)	0.66 (0.02)	0.53 (0.37)	0.53 (0.37)	0.58 (0.16)	0.61 (0.09)
Method 2									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1 Information criteria	FM1 RMSE	FM2 Information criteria	FM2 RMSE	FM3 Information criteria	FM3 RMSE
-4 (<i>Q</i>)	0.50 (0.50)	0.58 (0.16)	0.58 (0.16)	0.47 (0.63)	0.47 (0.63)	0.53 (0.37)	0.53 (0.37)	0.53 (0.37)	0.61 (0.09)
-3 (<i>Q</i>)	0.50 (0.50)	0.32 (0.99)	0.50 (0.50)	0.58 (0.16)	0.71 (0.00)	0.61 (0.09)	0.61 (0.09)	0.61 (0.09)	0.50 (0.50)
Forecasts made during the quarter to be forecast at horizons where the three methods are equivalent									
Months to GDP release	Random walk	AR(2)	Calibration model	FM1 Information criteria	FM1 RMSE	FM2 Information criteria	FM2 RMSE	FM3 Information criteria	FM3 RMSE
-2 (<i>Q</i>)	0.50 (0.50)	0.32 (0.99)	0.42 (0.84)	0.53 (0.37)	0.61 (0.09)	0.63 (0.05)	0.63 (0.05)	0.58 (0.16)	0.58 (0.16)
-1 (<i>Q</i> +1)	0.50 (0.50)	0.32 (0.99)	0.68 (0.01)	0.79 (0.00)	0.71 (0.00)	0.71 (0.00)	0.71 (0.00)	0.66 (0.02)	0.74 (0.00)
0 (<i>Q</i> +1)	0.50 (0.50)	0.33 (0.99)	0.76 (0.00)	0.82 (0.00)	0.79 (0.00)	0.82 (0.00)	0.82 (0.00)	0.63 (0.05)	0.71 (0.00)

Table B.5c. **Forecast evaluation: Pesaran and Timmermann test** (cont.)

Notes:

FM1: Model with factors extracted from the 93 variables.

FM2: Model with factors extracted from survey and real variables.

FM3: Model with factors extracted from survey, nominal, and international variables.

Calibration model: Calibration model relating GDP growth to business climate in France and industrial production index.

Information criteria/RMSE: Factor model derived from the specification obtained using the information criteria/RMSE respectively.

The Pesaran and Timmermann test is a nonparametric test that assesses the model's ability to forecast the direction of change in the variable under consideration. The table shows the proportion of times that the direction of change is correctly predicted and the P-value of the test in parentheses (if the P-value is below 5%, the proportion of correctly forecast directions exceeds 50% at the 5% significance level).

The lowest MAFE obtained by each method for each horizon are shown in italics.

Sources: See Table A.2.

New indicators for tracking growth in real time

by
Troy D. Matheson*

We develop monthly indicators for tracking short-run trends in real GDP growth in 32 advanced and emerging-market economies. We test the historical performance of our indicators and find that they do a good job at describing the business cycle. In a recursive out-of-sample forecasting exercise, we find that the indicators generally produce good real GDP growth forecasts relative to a range of time series models.

Keywords: Nowcasting, short-term forecasting, real-time data

JEL classification: C51, C53, E17

* Research Department, International Monetary Fund, Washington, DC, tmatheson@imf.org. The author would like to acknowledge comments from Jorg Decressin, Emil Stavrev, and Krishna Srinivasan. The paper outlines the methodology behind the Growth Tracker, which appears in the IMF *World Economic Outlook*. All errors and omissions are entirely my own and the views expressed in this paper do not necessarily reflect those of the IMF.

1. Introduction

Evaluating the current state of the business cycle is of crucial importance to policy makers and the general public alike. However, economic data are often noisy and available with a substantial lag. Determining the underlying state of an economy is thus very difficult in practice, requiring a mix of information gleaned from economic and statistical models and – perhaps most importantly – the expertise of economists. The importance of the real-time measurement of economic activity is reflected in the significant resources devoted to macroeconomic monitoring by policy-making institutions, and the large number of private firms providing economic analyses to clients eager to get a fix on the current state of the economy. Against this backdrop, we develop monthly growth indicators for 32 advanced and emerging-market economies that can utilise a wide range of economic information in real time. These indicators are currently used for tracking short-term trends in global growth in the *World Economic Outlook*.

The OECD and the Conference Board have a long history of publishing composite indicators,¹ while more recent attempts to capture the United States business cycle using dynamic factor models have been provided by Stock and Watson (1989), Mariano and Murasawa (2003), Aruoba et al. (2009), and Aruoba and Diebold (2010). Mariano and Murasawa (2003), Aruoba et al. (2009), and Aruoba and Diebold (2010) extend the dynamic factor model to incorporate data measured at different frequencies. The EuroCoin indicator represents one of the first serious attempts to publish an economic indicator that utilises a large panel of data in real time.² However, the EuroCoin indicator uses an approach to handle missing observations at the end of the sample that changes with the pattern of available data making the underlying model change over time.

Giannone et al. (2008) use the Kalman filter and a dynamic factor model to forecast real GDP growth in the presence of missing data at the end of the sample due to publication lags, the so-called “jagged edge”. This approach has been followed by Barhoumi et al. (2008) and Matheson (2010), amongst others, and extended to deal with an arbitrary pattern of missing data by Banbura and Modugno (2010). Similarly, Evans (2005) estimates real GDP at the daily frequency for the United States using different vintages of GDP, and Camacho and Perez-Quiros (2010) aim to estimate real GDP growth at the monthly frequency for the euro area by incorporating data on preliminary, advanced and final GDP releases.³ We use an approach that is similar in spirit to Giannone et al. (2008), but focus on estimating monthly real GDP growth indicators for a large number of countries.

The primary objective of this paper is to produce indicators for short-run trends in real GDP growth that describe the behaviour of economic activity at the monthly frequency, while utilising a wide range of economic time series in a timely fashion. Our approach does not explicitly model high-frequency data in the statistically optimal way, as suggested Aruoba and Diebold (2010). This choice was made to facilitate the use of the model in real time across a wide range of countries, where some countries have limited data of sufficient quality to produce reliable estimates of economic activity at higher frequencies than that which is published.

Moreover, we consider the computational cost of estimating a high-frequency model with a large number of economic indicators as currently being too high for our purposes.

We find that our indicators generally do a good job at describing the behaviour of real GDP growth for most countries considered. In a quasi-real-time forecasting experiment, we also find that the indicators produce good forecasting performance relative to a range of time series models. The indicators generally have good revision properties when applied in real time and, for those countries where historical revisions are particularly large, there is some hope for efficiency improvements with time.

The paper proceeds as follows. Section 2 outlines the methodology. Section 3 describes the data. Section 4 discusses the specification of the dynamic factor model and the historical fit of our growth indicators. Section 5 discusses a real-time forecasting exercise and presents the results. Section 6 evaluates the revision properties of the indicators and Section 7 concludes.

2. Methodology

2.1. Dynamic factor model

The monthly growth indicators are estimated using the dynamic factor model (DFM). The DFM is particularly useful in this context, because it can utilise a large number of economic time series in a timely fashion and it has been shown to produce reliable short-term forecasts. See, for example, Giannone et al. (2008), Barhoumi et al. (2008), and Matheson (2010).

The DFM assumes that monthly real GDP growth y_t can be decomposed into a common component X_t and an idiosyncratic component ε_t . The common component captures the bulk of the covariation between growth and a wide range of economic indicators, while the idiosyncratic component is assumed to mainly only affect growth:

$$y_t = \mu + \chi_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \psi) \quad (1)$$

where μ is a constant and $\chi_t = \Lambda F_t$, with $F_t = (F_{1t}, \dots, F_{rt})'$ and $\Lambda = (\lambda_1, \dots, \lambda_r)$, the time index t refers throughout the article to a given month and the prime denoting the transposed matrix. The common component is thus related to growth through a linear combination of a small handful of r static factors. The static factors themselves are, in turn, estimated using information from a potentially large panel of n economic indicators $X_t = (x_{1,t}, \dots, x_{n,t})$, where each indicator including monthly real GDP growth in X_t ⁴ has a factor representation:

$$x_{i,t} = \mu_i + \Lambda_i F_t + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim N(0, \psi_i) \quad (2)$$

The dynamics of the static factors are captured by the following vector autoregressive (VAR) process:

$$F_t = \sum_{s=1}^p \beta_s F_{t-s} + B v_t, \quad v_t \sim N(0, I_q) \quad (3)$$

where the β_s s are $r \times r$ matrices, p is the lag length of the process, B is an $r \times q$ matrix, and q is the number of underlying common shocks driving the economy. The number of static factors r is generally assumed to be large relative to the number of common shocks q in order to capture the dynamics of the economy. Together, Equations 2 and 3 form the state space system. This system constitutes the static factor representation of a model with q underlying dynamic factors (for more details, see Bai and Ng, 2007).

For the purposes of this paper, our indicator for monthly growth $\hat{y}_t$ is simply the component of growth estimated in Equation 1 (including the constant), after excluding the idiosyncratic component:

$$\hat{y}_t = \hat{\mu} + \hat{\lambda}_t \quad (4)$$

One of the key advantages of this framework is that the common component of growth can be estimated when some of the economic indicators have missing values at the end of the sample due to publication lags. This effectively means that the model can utilise all available information in real time.

2.2. Estimation

The estimation procedure follows Giannone et al. (2008) and begins with the panel of data X_t up to the last date when the balanced panel is available.⁵ The common factors F_t are then estimated from this balanced panel using principal components, and the factor loadings (Λ_i) and the covariance matrix of the idiosyncratic components ($\epsilon_{i,t}$) are estimated by regressing the variables on the estimated factors in Equation (2). The other parameters of the model are obtained by estimating a VAR (Equation 3) in the estimated factors.

The factors are then re-estimated using the Kalman filter by assuming that the errors are Gaussian, where for the unbalanced part of the panel, restrictions are imposed on the idiosyncratic components. These restrictions mean that the signal extraction process implicit in the Kalman filter will put no weight on the missing variables while computing the common factors at time t .⁶ See Giannone et al. (2008) for a more detailed description of how the procedure deals with missing observations.

This two-step estimation procedure (estimating the factors and parameters using principal components and OLS, and then re-estimating them using the Kalman filter) is simply the first step of the expectation-maximisation (EM) algorithm, and has been shown to produce consistent estimates by Doz et al. (2007). Banbura and Modugno (2010) adapt the EM algorithm to the case of an arbitrary pattern of missing data. This method is also discussed and utilised in Banbura et al. (2011).⁷

3. Data

Data selection is a crucial step in developing our monthly growth indicators. Choosing series that are too focused on particular sectors of the economy will bias the estimates, deteriorating the effectiveness of the DFM in estimating the underlying factors driving growth. Thus, for each country, we pay close attention to choosing data from a broad cross section of the economy.

Given poor data quality, particularly for some emerging countries, we employ a multi-step procedure for cleaning the data of outliers and missing observations. The vast majority of the series are measured at the monthly frequency, with the remaining series measured at the daily, weekly, quarterly and annual frequencies.⁸ All series are converted to the monthly frequency and, where required, they are transformed to be devoid of long-run trends (non-stationarity) prior to estimation of the DFM. Similar to Giannone et al. (2008), low frequency data are initially converted to the monthly frequency using linear interpolation. The data pre-filtering procedure is detailed in Annex A.

Broadly speaking, the data were chosen to cover the following categories:

- Activity (surveys) – includes Purchasing Managers' Index (PMI), consumer and business confidence, etc.
- Activity (hard data) – includes retail sales, industrial production, etc.
- Trade – includes exports, imports, exchange rates, etc.
- Financial conditions – includes interest rates, equity prices, credit conditions, etc.
- Employment and income – includes employment, wages, etc.
- Prices and costs – includes Producer Price Index (PPI), Consumer Price Index (CPI), inflation expectations, etc.

The implications of developments in key trading partners will be implicitly captured by the trade and survey data included for each country. However, each country's data set also includes eight key series for the USA, which are assumed to capture elements of the global business cycle not captured by the domestic indicators.⁹

Some information about the series used and their classifications can be found in Table 1. For most of the advanced economies the sample period begins in 1994, while the samples for many of the emerging-market economies begin later, due to a lack of available data and the presence of structural breaks. The number of series used also varies across countries depending on available data, ranging from 97 series for Kazakhstan to 290 for Sweden.

4. Specification and historical fit

4.1. Specification and general findings

Bai and Ng (2007) suggest a two-step procedure for determining the number of dynamic factors in factor models. The procedure relies on the fact that the $r \times 1$ matrix of innovations to the static factors (Bv_t in Equation 3) has rank equal to the number of dynamic factors q . The first step of the procedure requires the number of static factors r to be determined using information criteria described in Bai and Ng (2002). Then, once the number of static factors r is set, the rank of spectrum of the q dynamic factors is estimated using the eigenvalues of the residual covariance (or correlation matrix) of the VAR in the r static factors.

Unfortunately, we found that the Bai and Ng (2002) criteria generally produced too many factors, deteriorating the forecasting performance of the DFM. Likewise, the more ad-hoc approach used by Giannone et al. (2005) and Matheson (2010), where the number of factors is chosen to explain a certain percentage of the variation in a few key series, was not well suited to our multi-country setting because there is significant variation in the explanatory power of the factor model across countries. Instead, following Stock and Watson (2002), we choose the number of factors by minimising Schwarz's Bayesian information criterion (SBC).

Specifically, the number of common factors r is chosen by regressing interpolated monthly real GDP growth on the common factors for $r = 1, \dots, 8$; the number of factors is then that which minimises the SBC. The number of common shocks q is then chosen using information criteria described in Bai and Ng (2007).¹⁰ The number of lags of the factors p included in the model is determined using the SBC.

The specifications of the DFMs are displayed in Table 2. To get an idea of the quality of the growth indicators in describing the behaviour of quarterly real GDP growth over history, Table 2 also shows the percentage of the variation of quarterly growth explained by the indicators, R-squared and the proportion of time the indicators move in the same direction

Table 1. **Data description**

Country	Sample begins	Evaluation begins	Number of series in each category						Total
			Activity (surveys)	Activity (hard data)	Trade	Financial conditions	Employment and income	Prices and costs	
Americas									
Argentina	2003:01	2008:01	0	16	46	16	10	15	103
Brazil	1996:01	2001:01	17	31	56	22	10	12	148
Canada	1994:01	2000:01	19	57	38	12	17	18	161
Chile	2000:01	2005:01	9	29	53	30	12	17	150
Colombia	2000:01	2005:01	0	44	39	19	21	18	141
Dominican Republic	2000:01	2005:01	0	1	96	11	30	11	149
Ecuador	2000:01	2005:01	0	31	56	1	4	20	112
Mexico	2000:01	2005:01	20	33	33	10	17	16	129
Peru	2000:01	2005:01	0	48	24	18	14	20	124
United States	1994:01	2000:01	15	41	15	15	21	24	131
Uruguay	2001:01	2006:01	0	22	39	9	29	35	134
Venezuela	2004:04	2008:01	0	26	22	41	3	30	122
Europe									
France	1994:01	2000:01	60	28	20	17	24	39	188
Germany	1994:01	2000:01	58	31	39	18	26	15	187
Greece	2000:01	2005:01	33	41	26	19	19	32	170
Italy	1994:01	2000:01	55	32	23	22	12	29	173
Portugal	2000:01	2005:01	26	44	37	26	30	38	201
Russia	2000:01	2005:01	32	40	31	17	17	39	176
Spain	1994:01	2000:01	44	68	33	17	40	59	261
Sweden	1994:01	2000:01	59	60	66	14	42	49	290
United Kingdom	1994:01	2000:01	63	58	34	22	29	36	242
Euro area	1994:01	2000:01	20	27	17	17	6	29	116
Africa – Asia – Oceania									
Australia	1994:01	2000:01	32	37	42	8	20	32	171
China	2000:01	2006:01	23	82	29	7	34	17	192
India	2000:01	2007:01	32	25	36	18	4	12	127
Indonesia	2004:01	2008:01	3	24	41	12	3	24	107
Japan	1994:01	2000:01	30	39	22	9	7	6	113
Kazakhstan	2000:01	2005:01	0	10	51	12	5	19	97
Korea	2000:01	2005:01	37	49	42	22	20	30	200
Saudi Arabia	2000:01	2005:01	0	2	28	119	0	27	176
South Africa	1994:01	2000:01	24	58	45	23	14	27	191
Turkey	2002:01	2007:01	52	46	38	17	15	19	187

Note: yyyy:01 = January of year yyyy.

Source: Haver Analytics, www.haver.com/our_data.html, author's calculations.

as quarterly real GDP growth, concordance. To compute these statistics, the monthly growth indicators $\hat{y}_t$ have been estimated using all available data and converted to the quarterly frequency, that is $y_t^Q = y_t + y_{t-1} + y_{t-2}$ for all t corresponding to the third month of a given quarter). Clearly, the indicators will not fit the data as well in real-time; real-time revision properties are discussed in Section 6.

The indicators generally explain a sizable proportion of growth for the majority of countries, particularly for advanced countries. Because the growth indicators are estimates of the underlying, pervasive component of growth, their explanatory power tends not to be as great for emerging economies, where growth is generally more volatile and subject to larger idiosyncratic shocks. Nevertheless, the indicators do a good job at predicting the direction of quarterly real GDP growth over history, with all concordance statistics being above 50% – the proportion of time a coin toss would accurately predict the direction of a change in growth.

Table 2. **Factor model parameters and historical fit**

Country	Estimates for parameters r , q and p			Historical fit	
	r	q	p	R ² (%)	Concordance (%)
Americas					
Argentina	6	4	1	88	64
Brazil	4	2	1	62	66
Canada	2	2	1	71	61
Chile	1	1	3	49	68
Colombia	1	1	2	58	69
Dominican Republic	2	2	1	26	58
Ecuador	4	3	1	72	66
Mexico	1	1	3	60	54
Peru	5	3	1	70	63
United States	3	3	1	50	75
Uruguay	4	2	1	67	78
Venezuela	4	3	1	80	75
Europe					
France	4	3	2	86	85
Germany	3	2	2	90	69
Greece	4	3	2	82	80
Italy	5	3	1	55	61
Portugal	3	3	1	79	69
Russia	5	2	1	73	80
Spain	5	3	1	87	75
Sweden	2	2	2	84	73
United Kingdom	2	2	2	56	58
Euro area	3	1	1	65	57
Africa – Asia – Oceania					
Australia	1	1	3	63	66
China	3	3	1	50	67
India	6	4	1	59	81
Indonesia	3	3	3	42	76
Japan	4	3	2	85	78
Kazakhstan	6	4	2	69	85
Korea	1	1	3	34	60
Saudi Arabia	4	3	3	67	72
South Africa	6	4	1	59	68
Turkey	5	3	1	77	66
Average	4	3	2	67	69

Note: R²: percentage change in the quarterly growth rate of GDP explained by the growth indicator. Concordance statistic: the proportion of quarters in which the growth indicator is moving in the same direction as the quarterly growth rate of real GDP.

Source: Haver Analytics, www.haver.com/our_data.html, author's calculations.

Some of the growth indicators are subject to more short-run volatility than others. By construction, this volatility is pervasive across the series that went into constructing each indicator. It is also useful however, to consider an indicator that is both pervasive in the cross section and persistent over time. We thus introduce a smoothed monthly indicator y_t^* that removes the short-run volatility from the indicator estimated with the DFM. These smoothed indicators are simply centered seven-month-moving averages of the estimated monthly indicators $\hat{y}_t$:¹¹

$$\hat{y}_t = y_t^* + v_t \quad (5)$$

where v_t captures the short-run noise in the common component of growth. Note a centered seven-month-moving average requires forecasts of future observations at the end of the sample, making the quality of the growth “signal” in real time dependent on the quality of the DFM forecasts. These forecasts are assessed in Section 5.

The monthly growth indicators are displayed in Figure 1 along with interpolated monthly real GDP growth; dates beyond the collapse of Lehman Brothers near the beginning of the global financial crisis are shaded. The indicators generally do a good job at tracking trends in interpolated monthly GDP growth over time. All countries’ indicators fell markedly with the onset of the global financial crisis and have since recovered to around pre-crisis levels. As mentioned above, the indicators for the emerging economies produce reasonable estimates of the underlying trends in real GDP growth, despite the volatility inherent in these countries. Even for Saudi Arabia, where GDP is measured at the annual frequency, the growth indicators produce reasonable estimates of growth at the monthly frequency.

4.2. Smoothed indicators for tracking growth

By incorporating estimates of potential output growth, the behaviour of the smoothed growth indicators $\hat{y}_t$ can provide a great deal of information about the current state of the business cycle and the evolution of growth over time. The smoothed indicators are estimated with all available data and are subject to revision, particularly at the end of the sample period, where data on real GDP growth are not available. The indicators thus provide our “best guess” about the evolution of short-run trends in real GDP growth at the monthly frequency and can be easily updated as more information about the state of the economy arrives in real time. Their comparison with calculated estimates of potential growth can also provide valuable insight into the current state of the economy and the growth trends over time. As noted above, the real-time revision properties of the indicators are discussed in Section 6.

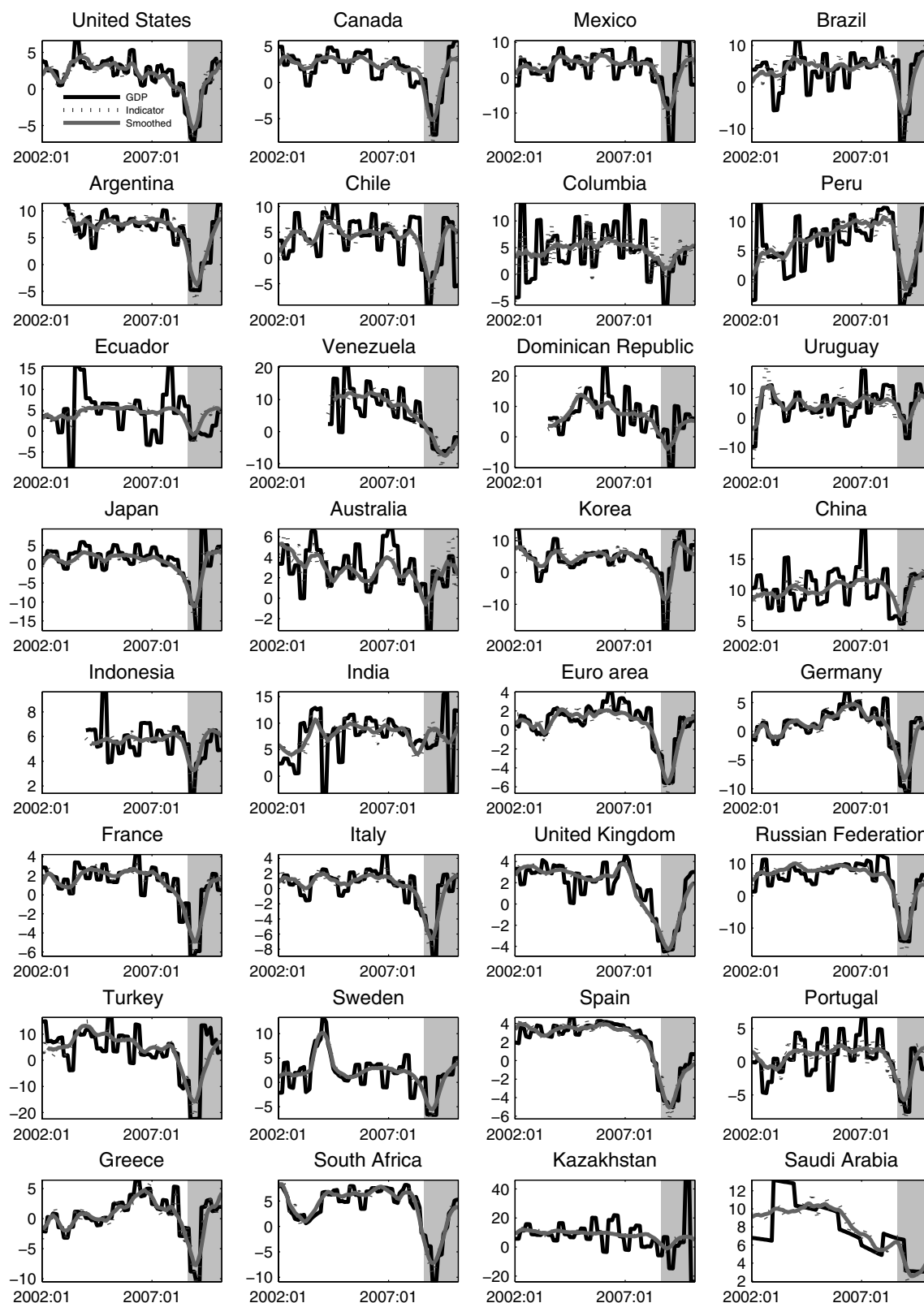
The heat map in Figure 2 displays information about growth for all of the countries for which we have indicators. The trends referred to in the heat map are the interpolated growth rates of potential output taken from *World Economic Outlook* (WEO) projections. The shading is based on the behaviour of the smoothed indicators relative to trend, with arrows indicating whether the indicators are increasing or decreasing: no shading with downward/upward arrows indicates growth above trend and moderating/rising; light grey with downward/upward arrows indicates growth below trend and moderating/rising; medium grey with upward arrow indicates contraction at a moderating rate; and, black indicates contraction at an increasing rate.

The heat map clearly shows the implications of the global financial crisis for growth. The effects of the crisis were seen across all 32 countries, but differed across regions. A contraction was evident in early 2008 for the United States before spreading to most other countries between January and the second quarter of 2008. The growth indicators suggest the US economy contracted from early 2008 to early 2009, a longer period than all other countries apart from Greece, where the effects of a sovereign debt crisis perpetuated the decline in activity.

The economies in the Western Hemisphere generally suffered a shorter contraction than the US, likewise for Korea and Australia. Meanwhile, activity in the Chinese, Indonesian and Indian economies slowed somewhat but growth remained positive throughout the crisis. The crisis was perhaps most keenly felt in Europe, with all major countries contracting from the middle of 2008 to the middle of 2009.

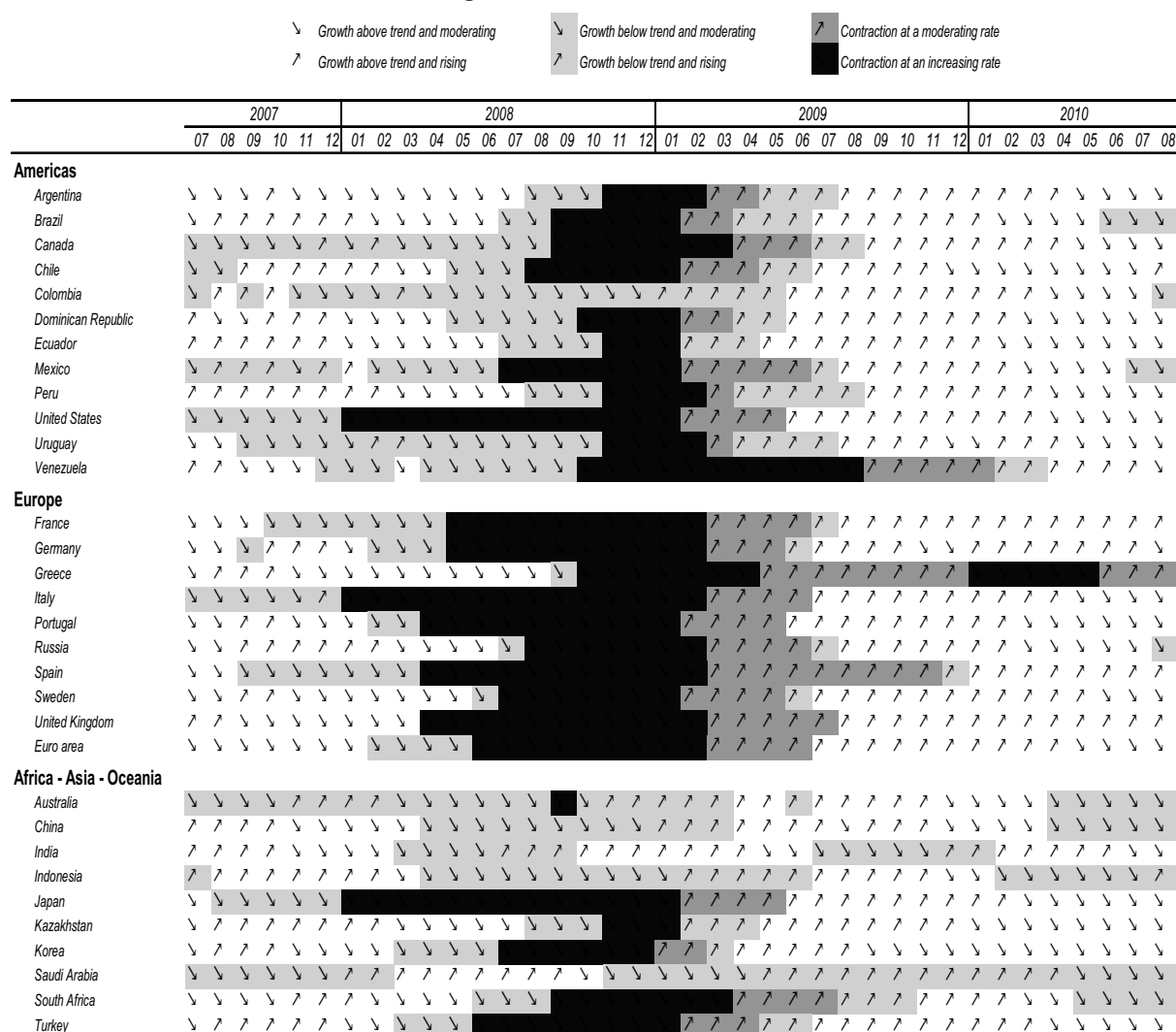
Figure 1. **Interpolated GDP growth and growth indicators**

Percentage, at an annual rate



Note: Shaded areas correspond to the period after the collapse of Lehman Brothers. yyyy:01 = January of the year yyyy.

Source: Haver Analytics, www.haver.com/our_data.html, author's calculations.

Figure 2. **Growth tracker**

1. The growth trackers are constructed using a large number of daily, monthly, and quarterly indicators and a dynamic factor model that incorporates all available data. The trackers are estimated and forecast at the monthly frequency. The classifications represented in the figure are based on the behaviour of a centred 7-month-moving average. The most recent estimates implicitly include forecasts and can change with the arrival of more data. Trend is potential output growth taken from GPM, where available. If not available, the trend is taken from the 'live' World Economic Outlook (WEO) database. If not available in the WEO database, the Hodrick-Prescott (HP) filter trend is used.
 2. Estimates based on data available on: 20/9/2010.
 3. Annex 1.2 of the October 2010 WEO describes the methodology in greater detail.
- Source: Haver Analytics and IMF staff estimates.

Following the downturn, growth recovered strongly for a period several months or more across most countries. Growth has since begun to moderate, first in parts of the western hemisphere and Asia, then in parts of Europe.

5. Quasi-real-time forecast evaluation

Assessing the underlying state of the economy is contingent on the behaviour of the data at hand and the model used to analyse the data. As such, to the extent new data differ from previous estimates produced by the indicators, they can be revised over both the historical and forecast periods. This may cause the indicators to produce some false

signals in real time. Thus, to evaluate how well the indicators perform in real time, we conduct a simulated real-time forecasting experiment.

5.1. The real-time problem

Within each quarter, contemporaneous values of key macroeconomic variables such as GDP are not available. Specifically, at the end of the last month t of each quarter, the data available is represented by the information set Ω_t^n , which includes the most recent data for n monthly time series. The forecaster's task is to project quarterly GDP growth y_{t+h}^Q for $h = 0, 3, 6, \dots, H$ based on the information set available at month t :

$$\hat{y}_{t+h}^Q = E[y_{t+h}^Q | \Omega_t^n], \quad h = 0, 3, 6, \dots, H \quad (6)$$

Assume Ω_t^n that composes of two blocks $[\Omega_t^{n1} \Omega_t^{n2}]$. The variables in Ω_t^{n2} , say industrial production, are released a month later than those in Ω_t^{n1} , say asset prices. This implies that variables in Ω_t^{n1} are available up to month t , while variables in Ω_t^{n2} is only available up month $t - 1$. Table 3 illustrates a stylised panel of data for different classes of variables. The forecaster needs to project on the basis of this unbalanced panel of data.

Table 3. **Stylised data panel for different classes of variable**

Month	Activity	Surveys	Asset prices	Foreign	GDP
$t-2$	X	X	X	X	0
$t-1$	X	X	X	X	0
t	0	X	X	0	0

Note: X indicates data is available at the end of the month, and O indicates data that is missing from the panel.

5.2. Quasi-real-time experiment

In our forecasting experiment, we aim to replicate the real-time application of the growth indicators as closely as possible. However, we do not have the vintages of data as they would have existed in real time. We also do not account for potential revisions to the seasonal factors that we estimate using all available data. Instead, we rely on data release dates recorded by Haver Analytics to compile quasi-real-time data sets; we manipulate the most recent vintage of data to match the publication lags that would have been seen in real time. These data sets mimic the data available at the beginning of the first month of each quarter of out-of-sample evaluation periods displayed in Table 1. For comparison, we also include a range of other forecasting models in the forecasting experiment, described below. In the experiment, we re-specify and re-estimate the models each time a forecast of quarterly real GDP growth is made.

Baseline quarterly autoregressive model (AR)

As a benchmark, we use an univariate AR model of order p for quarterly real GDP growth (y_t^Q):

$$y_t^Q = \gamma + \sum_{s=1}^p \beta_{3s} y_{t-3s}^Q + \varepsilon_t^Q \quad (7)$$

where $(\beta_{3s})_{s=1, \dots, p}$ are the coefficients of autoregressive terms, γ is a constant, ε_t^Q is a quarterly white noise term following a normal distribution of zero mean and variance σ_ε^2 ,

and where the lag length p is selected using the Schwarz Bayesian Criterion (SBC), the monthly index t refers exclusively to quarters (multiples of 3).

Pooled bridge equations (BE)

The bridge equation is perhaps the most widely used method for forecasting quarterly real GDP using monthly indicators.¹² Our bridge equation forecasts are constructed using the following four steps:

- (1) We consider the set of n monthly indicators x_t and forecast the individual indicators $x_{i,t}$ over the relevant horizon using a univariate AR(p) model:

$$x_{i,t} = c_i + \sum_{s=1}^{p_i} b_{i,s} x_{i,t-s} + U_{i,t}, \quad i = 1, \dots, n \quad (8)$$

- (2) Each indicator (including forecasts) is converted to the quarterly frequency $x_{i,t}^Q = x_{i,t} + x_{i,t-1} + x_{i,t-2}$, where t is the last month of a quarter (multiple of 3) and we estimate the following quarterly bridge equation:

$$y_t^Q = \gamma_i + \sum_{s=0}^{q_i} \beta_{i,3s} x_{i,t-3s}^Q + \varepsilon_{i,t}^Q \quad (9)$$

which relates quarterly GDP growth to the quarterly aggregate of the monthly indicator.¹³ The lag lengths p_i and q_i are determined using the SBC. The forecast of GDP growth is obtained by inserting the monthly indicator forecast from Equation 8 into 9.

- (3) We choose to select the set of 10 monthly indicators whose quarterly equivalents have the highest contemporaneous correlation with quarterly GDP growth.
- (4) The forecast for GDP growth is a weighted average of the 10 forecasts from the individual indicators, with the weights based on the inverse of the root mean squared forecast errors (RMSFE) of the individual indicators.

Note that the pooled bivariate VAR model described below also uses the 10 indicators selected in step 3 above. In preliminary work, we experimented with choosing the 5, 20, and 50 indicators most correlated with quarterly real GDP growth and found that forecasting performance generally deteriorated relative to the forecast based on 10 indicators. Choosing all of the available indicators for each country also deteriorated forecasting accuracy.

Pooled bivariate VARs (BV)

Similar to the bridge equation, the bivariate VAR model exploits the information content of monthly indicators. However, while the bridge equation relies on the autoregressive forecasts in step 1, it may be that information in real GDP growth itself can produce more efficient forecasts of the indicators and better forecasts of quarterly real GDP growth.

To capture some of the dynamics between each of the indicators and GDP, we estimate the following monthly bivariate VAR model on interpolated real GDP growth and each of the 10 indicators used in the bridge equations:

$$Z_{i,t} = \gamma_i + \sum_{s=1}^{p_i} \beta_{i,s} Z_{i,t-s} + \varepsilon_{i,t} \quad (10)$$

where $Z_{i,t} = [y_t, x_{i,t}]'$. Prior to estimation of each VAR, we interpolate all indicators measured at frequencies lower than monthly using the procedure proposed by Chow and Lin (1971); we follow Angelini et al. (2006) and use the monthly factors estimated using the DFM as regressors in the procedure. As with the other forecasting methods discussed, the lag length p_i of the VAR is determined using the SBC.¹⁴

Relative to the bridge equations, this methodology loses some information by using interpolated monthly real GDP growth, but it also may produce some efficiency gains by better capturing the dynamics between GDP growth and each indicator. We use the estimated VAR in Equation 10 to forecast the monthly real GDP growth rates, conditional on the latest monthly indicators available using the Kalman Filter. The forecast for quarterly real GDP growth is formed by weighting together the 10 bivariate VAR forecasts in the same way as the bridge equation forecast.

Bayesian VAR (BVAR)

One extension of the bivariate VAR is to include selection of potentially useful monthly indicators. Using the same notation as above, Z_t now includes a set of monthly indicators, as well as interpolated monthly real GDP growth:¹⁵

$$Z_t = \gamma + \sum_{s=1}^p \beta_s Z_{t-s} + \varepsilon_t \quad (11)$$

where the constant term γ is a $\kappa \times 1$ vector, β_s is a $\kappa \times \kappa$ autoregressive matrix, and ε_t is a $\kappa \times 1$ white noise process with covariance matrix Ψ . To overcome the “curse of dimensionality” problem, we estimate the VAR using Bayesian shrinkage methods by imposing prior beliefs on the parameters. In setting the prior distributions, we follow the procedure developed by Doan et al. (1984) and Litterman (1986). See Annex B for more details.

The vector Z_t of the BVAR contains interpolated monthly real GDP growth, industrial production, inflation, a real exchange rate, a short-term interest rate and equity prices.¹⁶ Following Banbura et al. (2010), the overall tightness of the prior μ_1 (which controls the overall concentration level of the a priori distribution of the elements $(\beta_s)_{l,j}$ of the matrix β_s around their mean, see Annex 2) is set such that the average R^2 across all equations of model (11) is fixed at 60% to avoid the problem of “over-fitting”. The BVAR contains six lags with λ (that controls the rate at which the standard deviation of the a priori distribution coefficients $(\beta_s)_{l,j}$ decreases with the magnitude of s delays in the model (11) – see Annex 2) set to 1, and the standard deviations of the parameters are taken from the estimated residuals of AR(6) processes. As with the bivariate VAR forecasts, the BVAR forecasts are made conditional on all available monthly data using the Kalman filter.

Pooled forecasts

There is a large literature showing that model combination tends to improve forecasting accuracy. As such, we also compute two pooled forecasts based on the forecasts described above. The first pooled forecast uses the recursively computed inverse RMSFEs of each forecast as weights (INVMSE) and the second is based on a simple average across forecasts (MEAN).

5.3. Forecasting results

The forecasting results for predicting the next GDP release – the nowcast – are displayed in Table 4.¹⁷ The right panel of the table contains the RMSFE of the AR

Table 4. **Forecast accuracy: Nowcast**

Country	Out-of-sample start	RMSFE relative to AR						RMSFE relative to AR						MEAN		
		AR (RMSFE)	DFM	BE	BV	BVAR	INVMSE	MEAN	AR	DFM	BE	BV	BVAR		INVMSE	
Americas																
Argentina	2008:01	4.79	1.05	1.18	0.90	1.01	1.01	0.97	3	6	7	1	5	4	2	
Brazil	2001:01	5.20	0.74	1.24	0.91	0.84	0.86	0.84	6	1	7	5	2	4	3	
Canada	2000:01	1.99	0.85	0.77	0.87	0.85	0.78	0.78	7	5	1	6	4	3	2	
Chile	2005:01	5.10	0.92	0.87	0.89	0.88	0.88	0.86	7	6	2	5	4	3	1	
Colombia	2005:01	5.18	0.87	0.80	0.95	1.03	0.83	0.81	6	4	1	5	7	3	2	
Dominican Republic	2005:01	7.55	0.81	0.96	0.95	0.91	0.89	0.88	7	1	6	5	4	3	2	
Ecuador	2005:01	4.75	0.95	0.95	0.95	0.99	0.92	0.89	7	5	3	4	6	2	1	
Mexico	2005:01	8.25	0.72	0.72	0.86	0.78	0.80	0.79	7	2	1	6	3	5	4	
Peru	2005:01	4.20	0.83	1.05	0.87	0.84	0.76	0.76	6	3	7	5	4	1	2	
United States	2000:01	2.79	0.69	0.79	0.89	0.73	0.73	0.76	7	1	5	6	2	3	4	
Uruguay	2006:01	6.03	0.72	0.81	0.88	0.86	0.84	0.82	7	1	2	6	5	4	3	
Venezuela	2008:01	11.46	0.66	0.75	0.83	0.99	0.80	0.76	7	1	2	5	6	4	3	
Europe																
France	2000:01	2.27	0.66	0.68	0.81	0.93	0.75	0.76	7	1	2	5	6	3	4	
Germany	2000:01	3.79	0.84	0.77	0.89	0.98	0.83	0.82	7	4	1	5	6	3	2	
Greece	2005:01	3.86	0.99	0.89	0.88	1.06	0.80	0.93	6	5	3	2	7	1	4	
Italy	2000:01	2.87	0.55	0.75	0.86	0.91	0.74	0.77	7	1	3	5	6	2	4	
Portugal	2005:01	3.64	0.81	0.62	0.85	1.15	0.8	0.81	6	4	1	5	7	2	3	
Russia	2005:01	9.41	0.39	1.37	0.80	0.78	0.54	0.59	6	1	7	5	4	2	3	
Spain	2000:01	1.82	0.85	1.10	1.20	1.63	1.44	1.10	2	1	3	7	6	5	3	
Sweden	2000:01	4.62	0.76	0.77	0.83	0.84	0.81	0.80	7	1	2	5	6	4	3	
United Kingdom	2000:01	2.17	0.81	0.95	0.85	0.99	0.89	0.87	7	1	5	2	6	4	3	
Euro area	2000:01	2.15	0.66	0.79	0.83	0.69	0.72	0.72	7	1	5	6	2	4	3	
Africa – Asia – Oceania																
Australia	2000:01	2.43	1.19	0.94	1.08	1.15	1.01	0.99	3	7	1	5	6	4	2	
China	2006:01	3.65	0.86	1.02	1.05	0.69	0.86	0.83	5	3	6	7	1	4	2	
India	2007:01	3.03	1.52	0.88	0.83	1.15	0.82	0.84	5	7	4	2	6	1	3	
Indonesia	2008:01	1.92	0.80	0.91	0.80	0.85	0.84	0.81	7	1	6	2	5	4	3	
Japan	2000:01	4.64	0.68	0.77	0.86	0.72	0.74	0.76	7	1	5	6	2	3	4	
Kazakhstan	2005:01	8.02	0.85	0.98	0.87	0.89	0.91	0.89	7	1	6	2	4	5	3	
Korea	2005:01	5.93	0.52	0.80	0.91	0.86	0.78	0.75	7	1	4	6	5	3	2	
Saudi Arabia	2005:01	2.09	0.98	1.15	1.36	0.81	0.85	0.85	5	4	6	7	1	2	3	
South Africa	2000:01	2.17	0.82	0.90	0.96	0.87	0.86	0.84	7	1	5	6	4	3	2	
Turkey	2007:01	15.19	0.70	0.82	0.81	0.73	0.79	0.77	7	1	6	5	2	4	3	
Average			0.81	0.90	0.91	0.92	0.84	0.83	6.2	2.6	3.9	4.8	4.5	3.2	2.8	

Notes:

AR = Baseline quarterly autoregressive model; BE = pooled bridge equations; BV = Pooled bivariate vector autoregressive (VAR); BVAR = Bayesian VAR forecast; INVMSE = weights based on the inverse of the root mean squared forecast errors (RMSFE); MEAN = simple average of combined forecasts.

A RMSFE relative to AR <1 is interpreted as a forecasting performance greater than or equal to the baseline forecast.

Ranks of forecast range from 1 (best forecast) to 7 (least best forecast), the classification criteria reflect the numerically ascending order of the RMSFE.

yyyy:01 = January of year yyyy.

Source: Haver Analytics, www.haver.com/our_data.html, author's calculations.

benchmark in predicting annualised quarterly real GDP growth and the RMSFE of the competing models relative to RMSFE of the AR, where a ratio less than one indicates that the model in question outperforms the AR in the sense that the predictions derived from it are assigned a mean squared error numerically lower than the baseline. The left panel ranks the seven competing models on the basis of RMSFE.

The more sophisticated models outperform the AR for all countries except Australia, Argentina, India, Saudi Arabia and Spain, and of these models, the DFM generally produces the most accurate forecasts. The average RMSFE ratio for the DFM across countries is 0.81, the lowest of the competing models. The forecasts based on model averages INVMSE and MEAN, are the next most accurate with average RMSFE ratios of 0.84 and 0.83, respectively.

The DFM ranks as the best model for just over half of the countries, with the model averages also generally ranking highly. Across all countries, the DFM ranks as the best forecasting model, followed by the simple average of the forecasts and the inverse-MSE-weighted average. It is noteworthy that the DFM generally outperforms the model combination methods presented here, given the relatively good performance of these types of forecasts shown in previous studies. Note however, that for some countries where the DFM performs particularly poorly (e.g. Chile, Ecuador, Greece, India or Saudi Arabia), the model combination methods tend to perform better.

Overall, our growth indicators generally show good forecasting performance relative to a range of models. This, combined with the usefulness of the indicators in describing the behaviour of economic activity over history, makes them a useful tool for evaluating growth in real time.

6. Revision properties

In the previous section, we evaluated how well our growth indicators predict GDP growth in real time. It is also worthwhile to consider the revision properties of the monthly smoothed indicators discussed in Section 4. Figure 3 displays the recursively estimated smoothed indicators (thick grey line), along with interpolated real GDP growth and the smoothed indicators estimated with all available data. For each country, the recursively estimated indicators are estimated using the data that would have been available at the beginning of every month of the out-of-sample period. The seven-month-moving averages are centered on the months in which the forecasts are made and implicitly include forecasts. The deviations in the dotted lines from the grey lines represent the extent of the revisions to the real-time estimates of the smoothed indicators. As with Figure 1, the dates beyond the collapse of Lehman Brothers near the beginning of the global financial crisis are shaded.

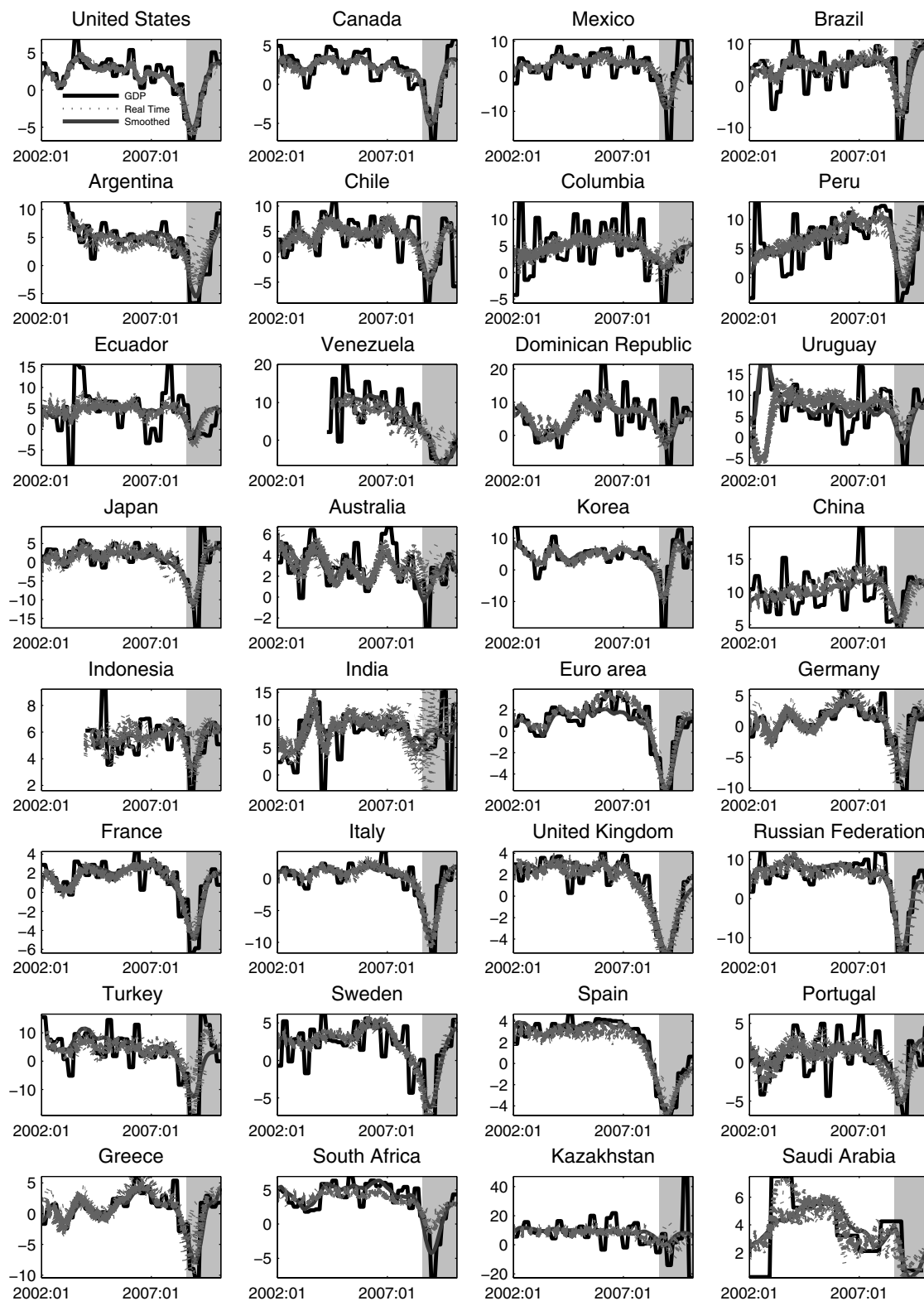
For most countries, the revisions to the indicators are small relative to the variance of GDP growth, and they do not show a marked deterioration since the global financial crisis. However, the revisions to the indicators for some countries – particularly Australia, India and Saudi Arabia – are large and warrant further discussion.

Recall that the relative forecasting performance of the DFM for Australia and India was not as good as for other countries. This accounts from some of the large revisions associated with these countries. But short sample periods and instabilities in the specifications of the DFM over time are perhaps more important.

In each month of the out-of-sample period, all DFM parameters are re-selected and re-estimated on the basis of the criteria described in Section 2. In fact, we find that the specifications tend to change more for India and Saudi Arabia than for most other countries in the recursive experiment, introducing volatility into the estimates of the indicators. While the recursively estimated indicators for Australia are not subject to this small sample problem, a relatively low R-squared in explaining quarterly real GDP growth

Figure 3. **Interpolated GDP growth and smoothed indicators in real time**

Percentage, at an annual rate



Notes: Monthly growth indicators estimated recursively are shown in several curves of increasing length (one curve per adjustment). Shaded areas correspond to the period after the collapse of Lehman Brothers. yyyy:01 = January of the year yyyy.

Source: Haver Analytics, www.haver.com/our_data.html, author's calculations.

for an advanced economy of 59%, even after including six factors, suggests a relatively weak factor structure in the Australian data.

Fortunately, the revision properties of the indicators seem to have improved over the past couple of years, suggesting that any inefficiencies in estimation seen over history may well become less important with time.

7. Concluding remarks

We developed monthly indicators for tracking short-run trends in real GDP growth for 32 advanced and emerging-market economies. For each country, the indicators were estimated using a dynamic factor model and a large number of economic time series. We find that our growth indicators did a good job at describing the business cycle and they produced reliable short-term forecasts relative to a range of time series models in a quasi-real-time forecasting experiment. The revision properties of the indicators were shown to be good for most of the countries and for the countries where revisions were historically large, there was some evidence that the properties of the indicators may well improve with time.

The indicators are currently used to evaluate the state of the business cycle in the World Economic Outlook published by the IMF, and it is hoped that applying the indicator in real time will prompt further refinements to the framework over time. The list of countries for which indicators are estimated is also likely to be expanded in the future.

Notes

1. See www.oecd.org/std/cli and www.conference-board.org/data/bci.cfm.
2. Altissimo et al. (2007).
3. For a review of methods that deal with the “jagged edge” in forecasting see Banbura et al. (2011).
4. Prior to estimation, the monthly rate of GDP growth has been achieved by using a linear interpolation of published GDP, usually on a quarterly basis. For more details, see Section 1.2.
5. As explained in more detail in Section 3 on data, series released at non-monthly intervals were made monthly before estimation. In particular, those published at a longer frequency (quarterly or annual series) were made monthly by linear interpolation.
6. Effectively, the Kalman filter computes the factors by weighting the innovation content of each variable by its signal to noise ratio. The restrictions state that this will go to zero when the data are unobserved.
7. Preliminary analysis with data from a small subset of the countries revealed that estimating the model with the EM algorithm described in Doz et al. (2007) produced little or no forecasting gains relative to the two-step procedure employed in this paper.
8. Real GDP for Saudi Arabia is the only series that is initially measured at the annual frequency.
9. These eight series are: the industrial production index (Federal Reserve Bank), three retail sales series (total retail sales, excluding automobiles and total including food services) published by the Bureau of Economic Analysis, a manufacturing index derived from the Institute for Supply Management (ISM) manufacturing survey, the employment and unemployment rates (US Bureau of Labor Statistics), and the consumer confidence indicator by the Conference Board.
10. Following Bai and Ng (2007) and using their notations (page 56), we use $\delta = 0.1$, and $m = 1$ for $\hat{q}_3$ and $\hat{q}_4$ (the covariance matrix of the VAR residuals is used, rather than the correlation matrix). We take $\hat{q}$ as being the (rounded) average of $\hat{q}_3$ and $\hat{q}_4$.
11. A more complicated way of producing a smoothed indicator would be to model the idiosyncratic term in Equation 1 as an AR process, as in Banbura and Modugno (2010), and Jungbacker et al. (2011).
12. See, for example, Kitchen and Monaco (2003) and Baffigi et al. (2004).

13. Note that a more general specification would allow for lags of $y_{i,t}^o$ on the right hand side of this equation. In our application, however, we found that allowing for such lags generally led to a deterioration in forecast accuracy.
14. Recall, we linearly interpolate low-frequency data to estimate the DFM. We instead employ the Chow-Lin procedure here in the hope that it will better capture monthly relationships between real GDP growth the indicators.
15. All indicators initially measured at frequencies higher than monthly are interpolated in the same way as those used in the bivariate VARs described above.
16. For some countries, due to a lack of available data, we replaced one or more of these series with series that have a similar economic interpretation.
17. The results for the one-step-ahead forecasts are qualitatively very similar to the nowcasting results, and are available from the author on request.

References

- Altissimo, F., et al. (2007), "New eurocoin: Tracking economic growth in real time", *Bank of Italy Working Papers* (631).
- Angelini, E., J. Henry and M. Marcellino (2006), "Interpolation and backdating with a large information set", *Journal of Economic Dynamics and Control* 30 (12), 2693-2724.
- Aruoba, B.S. and F.X. Diebold (2010), "Real-time macroeconomic monitoring: Real activity, inflation, and interactions", *American Economic Review* 100 (2), 20-24.
- Aruoba, B.S., F.X. Diebold and C. Scotti (2009), "Real-time measurement of business conditions", *Journal of Business & Economic Statistics* 27 (4), 417-427.
- Baffigi, A., R. Golinelli and G. Parigi (2004), "Bridge models to forecast the euro area GDP", *International Journal of Forecasting*, 20, 447-460.
- Bai, J. and S. Ng (2002), "Determining the number of factors in approximate factor models", *Econometrica* 70 (1), 135-172.
- Bai, J. and S. Ng (2007), "Determining the number of primitive shocks in factor models", *Journal of Business and Economic Statistics* 25 (1), 52-60.
- Banbura, M., D. Giannone and L. Reichlin (2010), "Large bayesian vector auto regressions", *Journal of Applied Econometrics* 25 (1), 71-92, <http://ideas.repec.org/a/jae/japmet/v25y2010i1p71-92.html>.
- Banbura, M., D. Giannone and L. Reichlin (2011), "Nowcasting", In: Clements, M. P., Hendry, D.F. (eds.), *The Oxford Handbook of Economic Forecasting*, Oxford University Press, Ch. 7.
- Banbura, M. and M. Modugno (2010), "Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data", *European Central Bank Working Paper Series* 1189.
- Barhoumi, K., et al. (2008), "Short-term forecasting of gdp using large monthly datasets: A pseudo real-time forecast evaluation exercise", *European Central Bank Occasional Paper* 84.
- Camacho, M. and G. Perez-Quiros (2010), "Introducing the euro-sting: Short-term indicator of euro area growth", *Journal of Applied Econometrics* 25 (4), 663-694, <http://ideas.repec.org/a/jae/japmet/v25y2010i4p663-694.html>.
- Chow, G.C. and A. Ioh Lin (1971), "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series", *Review of Economics and Statistics* 53 (4), 372-375.
- Doan, T., R. Litterman and C. Sims (1984), "Forecasting and conditional projections using realistic prior distributions", *Econometric Reviews* 3, 1-100.
- Doz, C., D. Giannone and L. Reichlin (2007), "A two-step estimator for large approximate dynamic factor models based on Kalman filtering", *Discussion Paper* 6043, Centre for Economic Policy Research (CEPR).
- Evans, M.D.D. (2005), "Where are we now? real-time estimates of the macroeconomy", *International Journal of Central Banking* 1 (2), <http://ideas.repec.org/a/ijc/ijcjou/y2005q3a4.html>.
- Giannone, D., L. Reichlin and L. Sala (2005), "Monetary policy in real time", In: Gertler, M., Rogoff, K. (eds.), *NBER Macroeconomics Annual 2004*, MIT Press, Cambridge, Mass.

- Giannone, D., L. Reichlin and D. Small (2008), "Nowcasting: The real-time informational content of macroeconomic data", *Journal of Monetary Economics* 55 (4), 665-676.
- Jungbacker, B., S. Koopman and M. van der Wel (2011), "Maximum likelihood estimation for dynamic factor models with missing data", *Journal of Economic Dynamics and Control* 8 (35), 1358.
- Kitchen, J. and R. Monaco (2003), "Real-time forecasting in practice: The U.S. Treasury Staff's real-time GDP forecast system", *Business Economics* 38, 10-28.
- Litterman, R. (1986), "Forecasting with bayesian vector autoregressions – five years of experience", *Journal of Business and Economic Statistics* 4, 25-38.
- Mariano, R.S. and Y. Murasawa (2003), "A new coincident index of business cycles based on monthly and quarterly series", *Journal of Applied Econometrics* 18 (4), 427-443, <http://ideas.repec.org/a/jae/japmet/v18y2003i4p427-443.html>.
- Matheson, T.D. (2010), "An analysis of the informational content of New Zealand data releases: The importance of business opinion surveys", *Economic Modelling* 27, 304-314.
- Stock, J.H. and M.W. Watson (1989), "New indexes of coincident and leading economic indicators", In: *NBER Macroeconomics Annual 1989*, Vol. 4, NBER Chapters, National Bureau of Economic Research, pp. 351-409, <http://ideas.repec.org/h/nbr/nberch/10968.html>.
- Stock, J.H. and M.W. Watson (2002), "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics* 20 (2), 147-62.

ANNEX A

Data transformation

We apply the following to each country's data set prior to estimation:

1. Missing values within the sample are linearly interpolated.
2. The seasonal series are adjusted using X12.
3. Quarterly and annual series are interpolated to the monthly frequency using linear interpolation; the daily and weekly series are converted into monthly averages.
4. Log quarterly differences are taken of the non-stationary series, $x_{i,t} = \ln(z_{i,t}) - \ln(z_{i,t-3})$, except those that are measured in percentages or can take negative values, in which case quarterly differences are taken, $x_{i,t} = z_{i,t} - z_{i,t-3}$. The remaining series are left as levels, $x_{i,t} = z_{i,t}$.
5. The series that only change less than 10% of the time are discarded.
6. The series with less than three years worth of data are discarded.
7. The series not released in the past year are discarded (to avoid discontinued data).
8. Outliers are removed, where observations greater/less than six times the interquintile range are replaced with the upper/lower bound of six times the interquintile range.
9. Missing observations at the beginning of the sample are backdated using the DFM, with the number factors set to explain 60% of the variation in the data.

ANNEX B

Bayesian VAR

The basic principle of the Litterman (1986) prior (often referred to as the Minnesota prior) is that all of the VAR model (11) are “centered” around a random walk with drift. This amounts to shrinking the diagonal elements of β_1 towards one and all other coefficients in $\beta_1, \dots, \beta_p$ towards zero:

$$Z_t = \gamma + Z_{t-1} + \varepsilon_t \quad (\text{B.1})$$

This embodies the belief that the more recent lags provide more useful information than the more distant ones. More formally, these priors can be imposed by setting the following moments for the prior distribution of the coefficients:

$$E[(\beta_s)_{l,j}] = \begin{cases} \delta_l & \text{if } j=l \text{ and } s=1 \\ 0, & \text{otherwise} \end{cases} \text{ and } V[(\beta_s)_{l,j}] = \left(\frac{1}{\mu_1 \cdot s^\lambda} \frac{\sigma_l}{\sigma_j} \right)^2 \quad (\text{B.2})$$

where $\delta_l = 1, \forall l$ reflects the random walk prior. The researcher can also incorporate priors where some variables are characterised by a degree of mean-reversion, $0 \leq \delta_l < 1$. In our application, we estimate BVARs on stationary data, so we set $\delta_l = 0, \forall l$. The hyper-parameter μ_1 controls the overall tightness of the prior distribution around δ_l , and the factor $1/s^\lambda$ is the rate at which the prior standard deviation decreases with the lag length of the VAR. See Banbura et al. (2010) for more details.

Dynamic factor models: A review of the literature

by

Karim Barhoumi*, Olivier Darné** and Laurent Ferrara***

In the last few years, the growth in the amount of economic and financial data available has prompted econometricians to develop or adapt new methods enabling them to summarise efficiently the information contained in large databases. Of these methods, dynamic factor models have seen rapid growth and become very popular among macroeconomists. In this paper, we carry out a survey of recent literature on dynamic factor models. We start by presenting the models used before looking at parameter estimation methods and statistical tests available for choosing the number of factors. We then focus on recent empirical applications dealing with the construction of economic outlook indicators, macroeconomic forecasts, and both macroeconomic and monetary policy analyses.

Keywords: Dynamic factor models, estimation, tests for the number of factors, macroeconomic applications.

JEL classification: C13, C51, C32, E66, F44

*Corresponding author: International Monetary Fund (IMF), 700 19th Street, N.W., Washington, DC 20431, United States, kbarhoumi@imf.org.

**University of Nantes (LEMNA) and Bank of France.

***Bank of France, International Macroeconomics Division.

The authors would like to thank Éric Dubois and Hélène Erkel-Rousse for their comments and suggestions, as well as two anonymous reviewers. This paper does not necessarily reflect the opinions of the Bank of France or the IMF.

1. Introduction

In recent years, the volume of available financial and economic data has led econometricians to develop or adapt methods that efficiently summarise the information contained in these large databases. In applied macroeconomics, frequently the practitioner has the tricky task of identifying, from among the large number N of variables available, the few variables of interest that will enable the best solution to their problem.

For example, economic growth and inflation forecasts are made in national and international institutions that have access to large volumes of data from surveys of households and businesses, and various series on prices and real activity such as the industrial production index (IPI), household consumption, the unemployment rate, etc. Similarly, central banks conduct monetary policy in a data-rich environment, looking at macroeconomic activity and the different financial markets on a regular basis, and tracking numerous monetary aggregates.

A number of econometric methods have been proposed in the literature for working in such data-rich environments. For example, to explain the changes in a particular variable using a vast set of exogenous variables N in a linear regression model, the so-called “general-to-specific” method (Krolzig and Hendry, 2001) proposes an algorithm that makes it possible to select just a few variables from among those variables N . Similarly, vector autoregressive (VAR) models are recognised as allowing for simultaneous modelling of variables in a multivariate context. Traditionally, VAR models use a small number of variables to avoid inflating the number of parameters to be estimated. To remedy this problem, Bayesian approaches have been proposed to estimate VAR models with a high number of variables N by imposing restrictions (see for example, De Mol, Giannone and Reichlin, 2008). Finally, if we consider the problem of predicting a particular variable when we have a large number of variables N that are potentially very relevant, we can imagine estimating N linear regressions, which then provide N forecasts that we will seek to combine (for forecast combination methods, see Newbold and Harvey, 2002). Eklund and Kapetanios (2008) also provide a review of the literature on the various forecasting techniques using large data sets.

Among the different methodologies proposed in the literature, dynamic factor models have grown significantly in popularity since the early 2000s and have been shown to be very useful in macroeconomic analysis and forecasting in a data-rich environment. These models can be used to summarise the information contained in a large number of economic variables into a small number of factors common to the set of variables. In this type of model, the N variables (x_{it}) , for $i = 1, \dots, N$ and $t = 1, \dots, T$, where t refers to the time index, are each assumed to be the sum of two unobservable orthogonal components: one component resulting from the factors that are common to the set of variables, (χ_{it}) , and an idiosyncratic component (ξ_{it}) . The component (χ_{it}) is obtained by extracting a small number $r \geq 1$ of common factors (F_{jt}) , $j = 1, \dots, r$ from all of the variables present in the data set. Often, by extension, this component (χ_{it}) is identified by the term “common

component,” which we will also use in this paper. The idiosyncratic component (ξ_{it}) covers the shocks specific to each of the variables. Thus, in a factor model of dimension ($N \times 1$), each element of the vector $X_t = (x_{1t}, \dots, x_{Nt})'$, assumed to be zero mean, can be written as follows:

$$x_{it} = \chi_{it} + \xi_{it},$$

or:

$$x_{it} = \lambda_{i1}F_{1t} + \dots + \lambda_{ir}F_{rt} + \xi_{it},$$

For $i = 1, \dots, N$ and $t = 1, \dots, T$. The loadings (λ_{ij}) for $i = 1, \dots, N$ and $j = 1, \dots, r$, represent the contributions of the variable i to the common factor (F_j) of dimension ($r \times 1$) such as $F_t = (F_{1t}, \dots, F_{rt})'$. The vector $(\xi_t) = (\xi_{1t}, \dots, \xi_{Nt})'$ of dimension ($N \times 1$) is a vector consisting of N idiosyncratic components. The vectorial form of the model is presented as follows, for all $t = 1, \dots, T$:

$$X_t = \Lambda F_t + \xi_t, \quad (1)$$

where Λ is the weighting matrix of dimension ($N \times r$). The matrix version is given as:

$$X = F\Lambda' + \xi, \quad (2)$$

where X is of dimension ($T \times N$), F is of dimension ($T \times r$), Λ is of dimension ($N \times r$), and ξ is of dimension ($T \times N$).

Given the rapid development of dynamic factor models in applied macroeconomics, we felt that the time was right to propose a review of the literature on these models to recap the current situation for practitioners. In Section 2, we begin by presenting the so-called traditional or classical factor models, which were developed initially for a small number of variables with common movements. We distinguish between static and dynamic approaches for these models. In Section 3, we describe approximate factor models, which can take a large number of variables into account, again in a static or dynamic context. Section 4 presents some estimation models proposed in the literature. A crucial aspect of these models is the selection of the number of common factors r to use in the analysis, so Section 5 provides a review of the various information criteria developed to select the optimal number of factors. There are many applications of factor models in the empirical economic literature including, for example, asset-pricing models (Ross, 1976), consumer theory (Gorman, 1981; Lewbel, 1991), performance assessment and risk measurement in finance (Campbell et al., 1997). In Section 6, we focus on some recent applications that underscore the interest of this approach for macroeconomists, specifically: (6.1) for the construction of short-term economic indicators; (6.2) for macroeconomic forecasting; and, (6.3) for international macroeconomics and monetary policy analysis.

2. Factor models for a small number of variables (small N)

In this section, we present factor models used to model a small number of variables N , where in practice, N is generally lower than six or seven variables. We begin with the simplest non-dynamic models (static factors), then look at dynamic models, and end with a few recent extensions of this type of model.

2.1. Static factor models (SFM)

In this type of model, a small number of unobservable variables r provides a linear explanation of a small number of observed variables N so that $r < N$. In the applications presented in Section 6, the number of variables is such that $N \leq 7$ and a single factor can generally explain most of the variance, that is, $r = 1$. The series are assumed to be stationary, to have finite variance and to be standardised. We put forward the following hypotheses, which could subsequently be abandoned:

- (SH1) The factors (F_t) are centred, $E(F_t) = 0$, and are mutually orthogonal for all t , that is: $\forall t, E(F_{jt}F_{jt'}) = 0$ for $j \neq j'$. Consequently, the variance-covariance matrix of (F_t) , $\Sigma_F = E(F_t F_t')$, where F_t' is the transpose of F_t , is a diagonal matrix.
- (SH2) The idiosyncratic processes (ξ_{it}) and $(\xi_{it'})$ are mutually orthogonal for all $i \neq i'$, with $E(\xi_{it}) = 0$. Consequently the variance-covariance matrix (ξ_t) is a diagonal matrix: $\Sigma_\xi = E(\xi_t \xi_t') = \text{diag}(\sigma_1^2, \dots, \sigma_N^2)$.
- (SH3) The factors (F_t) and idiosyncratic noise $(\xi_{it})_{i=1, \dots, N}$ are not correlated, i.e.: $\forall i, j, t, t'$, we have: $E(F_{jt} \xi_{it'}) = 0$.
- (SH4) The variables are assumed to be independent and identically distributed over time (the so-called IID hypothesis), so that $t \neq t', E(F_{jt} F_{jt'}) = 0$ and $E(\xi_{it} \xi_{it'}) = 0$.

The model given by Equation (1) represents the static factor model (SFM) in which the factors (F_t) do not possess their own dynamic and the relationship between the factors and variables is linear with constant weights over time. This model can be estimated either by assuming that the variables are IID (Hypothesis SH4), or by assuming that there is a time dynamic within the variables (SH4 is abandoned).

Assuming that (F_t) and (ξ_t) are not correlated and are zero mean, then the variance-covariance matrix for the static factor model, denoted $\Sigma_X = E(X_t X_t')$, is given by:

$$\Sigma_X = \Lambda \Sigma_F \Lambda' + \Sigma_\xi \quad (3)$$

By normalising the variance-covariance matrices of (F_t) , $\Sigma_F = I_r$, and by assuming that the diagonal elements of the variance-covariance matrix Σ_ξ of (ξ_t) are bounded, we obtain:

$$\Sigma_X = \Lambda \Lambda' + \Sigma_\xi \quad (4)$$

For additional details, we refer to Lawley and Maxwell (1971) and Anderson (1984). The static factor model can, thus be identified and estimated. The factorial analysis method is used for the static estimation of the factors. The weighting matrix Λ can be estimated by minimising the sum of the squared residuals as follows:

$$\sum_{t=1}^T (X_t - \Lambda F_t)' (X_t - \Lambda F_t) \quad (5)$$

subject to the constraint $\Lambda \Lambda' = I_r$.

In this context, Doz and Lenglart (1999) establish the asymptotic properties of the estimator. Specifically, they show that this method produces convergent estimators even when the data used are autocorrelated, as is the case with time series. Moreover, they show empirically that this method provides a very good approximation of the dynamic method,

while being easier to establish, which is an essential quality for forecasters who regularly estimate the model.

2.2. Exact or strict dynamic factor models (DFM)

Static factor models (SFM) are different from exact or strict dynamic factor models (DFM) in the sense that the latter incorporate a time dynamic. Thus, the common component in the DFM is seen as a sum of common shocks, whether contemporaneous or lagged. The model is defined as follows:

$$x_{it} = \chi_{it} + \xi_{it}, \quad (6)$$

where:

$$\begin{aligned} \chi_{it} = & b_{i1}^0 u_{1t} + \dots + b_{i1}^s u_{1,t-s} + b_{i2}^0 u_{2t} + \dots + b_{i2}^s u_{2,t-s} \\ & + \dots + b_{iq}^0 u_{qt} + \dots + b_{iq}^s u_{q,t-s}, \end{aligned} \quad (7)$$

where (u_t) , of dimension $(q \times 1)$, is the vector of common shocks such as $u_t = (u_{1t}, u_{2t}, \dots, u_{qt})'$, with $q \leq N$, and where s is the number of lags included in the model. The parameters (b_{il}^s) , for $\tau = 0, \dots, s, i = 1, \dots, N$ and $l = 1, \dots, q$, represent the weights of the dynamic factors with finite order s . We speak of both a “restricted” DFM when s is finite and a “generalised” DFM when s is infinite.¹

Equation (7) can be rewritten as follows:

$$\chi_{it} = \sum_{l=1}^q b_{il}(L) u_{lt}, \quad (8)$$

where $b_{il}(z) = b_{il}^0 + b_{il}^1 z + \dots + b_{il}^s z^s$ are polynomials of degree s and where L is the lag operator so that, for all s , $L^s u_t = u_{t-s}$. In a matrix form, Equation (7) can be rewritten as:

$$\chi_{it} = B_i(L) u_t, \quad (9)$$

where $B_i(L) = (b_{i1}(L), \dots, b_{iq}(L))$ is a q -vector of polynomials with degree s .

Moreover, we assume (Bai and Ng, 2007) that the vector $B_i(L)$ can be decomposed as follows:

$$B_i(L) = \lambda_i^*(L) C(L) \quad (10)$$

where $\lambda_i^*(L)$ is a r -vector of polynomials with degree s and where $C(L)$ is a matrix of dimension $(r \times q)$. Using Equations (9) and (10), we can then write the common component in the following manner:

$$\chi_{it} = \lambda_i^*(L) F_t^*, \quad (11)$$

where $F_t^* = C(L) u_t$ is a vector of dimension r that refers to the *static factors*, and the common shocks u_t of dimension $(q \times 1)$ to the *dynamic factors*. A model with q dynamic factors can, thus, be considered as a model with $r = q(s+1)$ static factors.

In the context of small-dimension dynamic factor models, the estimation is generally done in the time domain by likelihood maximisation, as proposed by Dempster et al. (1977), Shumway and Stoffer (1982), Watson and Engle (1983), and Stock and Watson (1989).²

To estimate the DFM when N is small, the following hypotheses are generally put forward:

- (DH1) The factors (F_{jt}) and $(F_{j't})$ are mutually orthogonal, but the factors (F_{jt}) can be autocorrelated and are variance-covariance stationary, that is: $\forall j \neq j', \tau \neq 0, E(F_{j,t}) = 0$, $\text{cov}(F_{j,t}, F_{j',t-\tau}) = 0$, and $\text{cov}(F_{j,t}, F_{j,t-\tau})$ depends only on τ .
- (DH2) The idiosyncratic processes (ξ_{it}) and $(\xi_{i't})$ are mutually orthogonal, but the processes (ξ_{it}) can be autocorrelated and covariance-stationary, that is: $\forall i \neq i', \tau \neq 0, E(\xi_{i,t}) = 0$, $\text{cov}(\xi_{i,t}, \xi_{i',t-\tau}) = 0$, and $\text{cov}(\xi_{i,t}, \xi_{i,t-\tau})$ depends only on τ .
- (DH3) The factors (F_{jt}) and the idiosyncratic processes (ξ_{it}) are orthogonal for all i, j .

Based on these hypotheses, we can attempt to estimate a dynamic factor model by likelihood maximisation in the time domain, with the additional hypothesis of normality for the model residuals. The maximum likelihood estimator is calculated by placing the model in a space-state form, and then using a Kalman-type recursive filter.

The DFM can be written in a space-state form, assuming that the common factors follow a VAR process of order p such as:

$$\Phi(L)F_t = \varepsilon_t \quad \Leftrightarrow \quad F_t = \sum_{\tau=1}^p \Phi_{\tau} F_{t-\tau} + \varepsilon_t, \quad (12)$$

and, for a given index i , the idiosyncratic process (ξ_{it}) follows an AR process of order p' in the following form:

$$\varphi_i(L)\xi_{it} = \eta_{it} \Leftrightarrow \xi_{it} = \sum_{\tau=1}^{p'} \varphi_{i\tau} \xi_{i,t-\tau} + \eta_{it}, \quad (13)$$

where (ε_t) and (η_{it}) are innovations of (F_t) and (ξ_{it}) , respectively, so that (ε_t) and (η_{it}) are independent. $\Phi(\cdot)$ and $\varphi_i(\cdot)$ are polynomials of order p and p' , respectively, so that:

$$\Phi(L) = I - \Phi_1 L - \dots - \Phi_p L^p \quad \text{and} \quad \psi_i(L) = I - \psi_{i1} L - \dots - \psi_{ip} L^{p'}.$$

The hypothesis of normality is put forward for (ε_t) and (η_{it}) . In practice, the orders p and p' of the lag polynomials must be selected prior to the estimation stage. This selection is generally done by minimising an Akaike information criterion (AIC) or a Bayesian information criterion (BIC), or by using the Doz and Lenglart (1999) test. In empirical studies, $p = 2$ and $p' = 1$ are often shown to be sufficient to whiten the residuals.

This type of model as shown by Equations (1), (12) and (13) allows a space-state representation as follows:

$$X_t = c_t \beta_t + m_t Z_t + w_t \quad (14)$$

where (Z_t) is a vector of n explanatory variables, for example, the lagged values of the observed variables (X_t) , and where:

$$\beta_t = a_t \beta_{t-1} + v_t \quad (15)$$

Equation (14) is the measure equation, which describes the relations between the unobservable states, of dimension r , and the observable variables, of dimension n , where β_t represents the state vector:

$$\beta_t = \begin{bmatrix} F_t \\ \vdots \\ F_{t-p+1} \\ \xi_t \\ \vdots \\ \xi_{t-q+1} \end{bmatrix}$$

Equation (15) represents the state or transition equation, which describes the development of unobservable states. We see that a_t , c_t and m_t are matrices that can depend on time, of the dimensions $((pr + qN) \times (pr + qN))$, $(N \times (pr + qN))$, and $(N \times n)$, respectively, and where v_t is a Gaussian white noise vector of dimension $(pr + qN)$, w_t is a Gaussian white noise vector of dimension N , of the variance-covariance matrices Q_t and R_t , respectively. In practice, the system is generally assumed to be invariant over time, that is, a_t , c_t and m_t are constant. It is also assumed that for all t , $t' \neq t$, $E(v_t w_{t'}) = 0$.

The model in its space-state form can then be estimated by maximum likelihood using a filtering method such as the Kalman filter. We refer, for example, to Hamilton (1994) for a description of the filtering algorithm. The maximum likelihood estimation algorithm can take a great deal of time, as it requires inversion of a large dimensional matrix, even when N is small. In general, in the case of numerical optimisation, the expectation-maximisation (EM) algorithm is used, as proposed by Dempster et al. (1977) or Shumway and Stoffer (1982).³

2.3. Recent extensions of factor models with small N

Several extensions of factor models with a small number of variables have recently been proposed, to take certain data characteristics into account. We present two of these extensions below.

2.3.1. Markov regime-switching models

These models are directly linked to the Markov regime-switching processes introduced by Hamilton (1989) and assume that the common unobservable factors have their own dynamics governed by a two-regime Markov chain, denoted (S_t) , with for all t , $S_t \in \{1, 2\}$. The idea of these models is to assume that the factors are related to the state of the economy, which itself evolves cyclically but non-periodically, based on two economic phases that follow one another. We assume that, for example, the first regime ($S_t = 1$) corresponds to the low phase of the business cycle and the second regime ($S_t = 2$) to the high phase of the business cycle. The model can easily be extended to a larger number of regimes, but the estimation problems become complicated as the model contains two latency levels, that is, the common factors and the Markov chain.

An initial model of this type was proposed by Diebold and Rudebusch (1996), but the theoretical and empirical aspects were more broadly considered by Kim and Yoo (1995), and Kim and Nelson (1998). At the same time, Chauvet (1998) independently proposed a

similar model. For example, in the case of a single factor (that is, $r = 1$) for N centred stationary variables, the Markov regime-switching model can be defined as follows for $i = 1, \dots, N$ and $t = 1, \dots, T$:

$$x_{it} = \lambda_i' F_t + \xi_{it} \quad (16)$$

with:

$$\phi(L)F_t = \mu(S_t) + \varepsilon_t, \quad (17)$$

where the $\lambda_i' = (\lambda_{i1}, \dots, \lambda_{ir})$ are the loadings, where, for each i , (ξ_{it}) follows a Gaussian autoregressive process of order one (AR(1)) of finite variance σ_i^2 , where (ε_t) is a Gaussian white noise of unit variance and where $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$. If we assume that (S_t) is a first-order two-regime Markov chain, this means that the probability of S_t belonging to a regime at date t depends only on the probability of being in a certain regime at date $(t - 1)$, or:

$$P(S_t | S_{t-1}, S_{t-2}, S_{t-3}, \dots) = P(S_t | S_{t-1}) \quad (18)$$

The transition probabilities $p_{12} = P(S_t = 2 | S_{t-1} = 1)$ and $p_{21} = P(S_t = 1 | S_{t-1} = 2)$ measure the probability of switching from one regime to the other. Similarly, the probabilities $p_{11} = 1 - p_{12}$ and $p_{22} = 1 - p_{21}$ measure the probability of remaining in the same regime, thus reflecting the degree of persistence of each regime. The estimation stage allows for the estimation for each date t of the expected, filtered and smoothed probabilities of being in a particular regime, respectively given by $P(S_t | \hat{\theta}, X_{t-1}, \dots, X_1)$, $P(S_t | \hat{\theta}, X_t, \dots, X_1)$, and $P(S_t | \hat{\theta}, X_T, \dots, X_1)$, where $\hat{\theta}$ represents the set of estimated parameters of the model, which includes the N autoregressive parameters of the AR(1) models, the N idiosyncratic variances and the p parameters of the polynomial $\phi(\cdot)$.

The parameters of this model can be estimated simultaneously by maximum likelihood as proposed by Kim and Nelson (1998), or in two steps, by first estimating the common factor (F_t) in the time or spectral domain (see above), and then adjusting an autoregressive regime-switching process on the estimated factor (see Diebold and Rudebusch, 1996). From a theoretical standpoint, simultaneous estimation is preferable but empirically, the maximisation algorithm is often difficult to converge, especially if the variables are very volatile. The two-stage estimation is more practical (for an application, see Darné and Ferrara, 2011), but Equation (17) includes an error in the measure of the estimated factors that is not explicitly integrated in the model, which can thus create statistical inference problems.

2.3.2. Mixed frequency models

Numerous macroeconomic series are available to forecasters but do not necessarily have the same sampling frequency (or periodicity). The national accounts in particular, which most economists seek to predict, are available only on a quarterly basis while many economic indicators, such as the industrial production index (IPI), household consumption or opinion surveys are monthly. To be able to handle simultaneously these two periodicities in a single model, Mariano and Murasawa (2003) proposed a dynamic factor

model in a space-state form that considers quarterly series as monthly series with missing values.

The aim of this type of model is to estimate a common factor with N variables, some of which are quarterly and some of which are monthly. Thus, we write $(Y_{1,t})$ a vector of N_1 quarterly variables that are observable only in the third month t of the quarter and $(Y_{2,t})$ a vector of N_2 monthly variables, so that $N_1 + N_2 = N$. We assume here that these series (in logarithms) are integrated of order one. We also assume that there is a vector $(Y_{1,t}^*)$ of N_1 unobservable monthly variables so that for all t , $Y_{1,t}$ is the geometric mean of $Y_{1,t}^*$ over the three months in a given quarter, that is:

$$\log(Y_{1,t}) = \frac{1}{3} [\log(Y_{1,t}^*) + \log(Y_{1,t-1}^*) + \log(Y_{1,t-2}^*)] \quad (19)$$

We note that this identity (19) differs from the arithmetical mean generally used in the quarterly accounts but makes it possible to implement a linear space-state form, in contrast to the arithmetical identity, that requires a non-linear form.

Mariano and Murasawa (2003) show that:

$$y_{1,t} = \frac{1}{3} y_{1,t}^* + \frac{2}{3} y_{1,t-1}^* + y_{1,t-2}^* + \frac{2}{3} y_{1,t-3}^* + \frac{1}{3} y_{1,t-4}^* \quad (20)$$

where $y_{1,t} = Y_{1,t} - Y_{1,t-3}$ and $y_{1,t}^* = Y_{1,t}^* - Y_{1,t-3}^*$, $(y_{1,t})$ being observable in all three periods and $(y_{1,t}^*)$ being unobservable.

Under standard hypotheses of factor dynamics and idiosyncratic errors, and the normality of residuals (see Mariano and Murasawa, 2003, p. 430), one can show that the model with a factor (F_t) is written in the following form for all t :

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} a_1(\frac{1}{3}F_t + \frac{2}{3}F_{t-1} + F_{t-2} + \frac{2}{3}F_{t-3} + \frac{1}{3}F_{t-4}) \\ a_2F_t \end{pmatrix} \quad (21)$$

$$+ \begin{pmatrix} \frac{1}{3}u_{1,t} + \frac{2}{3}u_{1,t-1} + u_{1,t-2} + \frac{2}{3}u_{1,t-3} + \frac{1}{3}u_{1,t-4} \\ u_{2,t} \end{pmatrix}$$

where $a = (a_1', a_2')'$ is the weight vector of dimension N , (F_t) is the scalar common factor, $u_t = (u_{1,t}', u_{2,t}')'$ is the idiosyncratic component of dimension N , and where $y_{2,t} = Y_{2,t} - Y_{2,t-1}$. The model is then put in a space-state form (see Mariano and Murasawa, 2003, p. 431) and estimated by maximum likelihood with the help of a Kalman filter.

Other approaches have been proposed to manage data with different periodicities simultaneously in factor models. For example, Aruoba, Diebold and Scotti (2009) also propose a factor model containing four variables of different periodicities (daily, weekly, monthly and quarterly) to estimate US GDP with a high periodicity. This indicator is actually updated weekly by the Philadelphia Federal Reserve on its internet site. Similarly, Camacho and Perez-Quiros (2010, 2011) propose a factor model that simultaneously treats variables with different periodicities and regime switching, in factors to estimate GDP growth in the euro area and Spain. For French data, Cornec and Desperaz (2006) construct a composite activity indicator for services based on monthly and quarterly survey data. Cornec (2006) also develops an indicator based on a mixed frequency factor model (see Section 6 for applications). More generally, we should point out that this type of approach makes it possible to treat the problem of missing data in series in an econometric model.

3. Approximate factor models (large N)

Although the concept of factor models is attractive, the traditional approach presented in Section 2 has a number of limitations that are both theoretical and practical in nature:

1. The number of variables (N) is often larger than the number of observations (T) in economic data series. Consequently, potentially important information is lost when a small number of variables must be selected to respect the constraint that N be small.
2. Asymptotic convergence of estimators is assured when T tends to infinity and N is fixed, but not when N also tends to infinity.
3. IID hypotheses and hypotheses on the diagonality of the variance-covariance matrix of the idiosyncratic component Σ_ε , which prohibit the cross correlation, are often too strong for economic data. This can result in a risk of misspecification.
4. The maximum likelihood estimation (MLE) is generally considered unachievable for factor models of large dimensions because the number of parameters to be estimated is too large (Bai, 2003; Bai and Ng, 2002).
5. The traditional approach makes it possible to consistently estimate the coefficients of the weighting factors (λ_i) by MLE when T is large, but not the common factors (F_t), for which only the expected value can be obtained (Steiger, 1979). Meanwhile, in most economic problems, these common factors are of greatest interest since for example, they represent the common shocks, the diffusion indices, etc.

To respond to a number of these limitations, the idea of factor models was generalised to allow for the manipulation of less strict hypotheses on the variance-covariance matrix of the idiosyncratic components by proposing an approximate factor structure. Non-parametric estimators of common factors based on the principal components have been suggested (Forni et al., 2000; Stock and Watson, 2002), their asymptotic properties being known when N is large. New methodologies have consequently, been proposed.

3.1. Approximate static factor models (SFM)

Chamberlain and Rothschild (1983) are the first to introduce the so-called “approximate” factor structure concept by abandoning the hypothesis that idiosyncratic disturbances are not mutually correlated within the so-called “strict” factor structure, that is, allowing for idiosyncratic errors to be weakly correlated. This concept makes it possible to obtain a non-diagonal variance-covariance matrix $\Sigma_\varepsilon = E(\xi_t \xi_t')$. Moreover, Chamberlain and Rothschild (1983) show that a principal components analysis (PCA) is equivalent to a factor analysis (or to a maximum likelihood analysis under the hypothesis of normality of the idiosyncratic component (ξ_{it})) when N increases to infinity. However, they assume that the variance-covariance matrix of the population, Σ_X , of dimension $(N \times N)$, is known. Connor and Korajczyk (1986, 1988, 1993) study the case of an unknown variance-covariance matrix Σ_X and suggest that, when N is larger than T , the factor model can be estimated by applying a PCA to the variance-covariance matrix Σ_X , of dimension $(T \times T)$.

Connor and Korajczyk (1986) establish the coherence of factors estimated by PCA when T is fixed and N tends to infinity in the context of approximate factor models but they provide no formal argument when N and T simultaneously tend to infinity.⁴ Stock and Watson (1999) study the uniform coherence of estimated factors and derive convergence rates for large T and N . Bai and Ng (2002) also study the convergence rate. Finally, Bai (2003) shows that the PCA estimator of the common component is asymptotically Gaussian,

converging to a rate equal to $\min(N^{1/2}, T^{1/2})$, even when the idiosyncratic component is serially correlated and/or heteroskedastic when N and T are large.⁵

3.2. Approximate dynamic factor models (DFM)

Forni and Lippi (1997), Forni and Reichlin (1998) and Forni et al. (2000, 2004) extend approximate factor models by considering dynamic factor models of large dimensions and introduce different methods for the estimation of this type of model. These models are referred to as “generalised” because they combine both dynamic and approximate structures, that is, they generalise exact dynamic factor models by assuming that the number of variables N tends to infinity and by allowing idiosyncratic processes to be mutually correlated.

Forni et al. (2000, 2004) expand the dynamic principal components analysis introduced by Brillinger (1981) when N is large. The estimation proposed by Brillinger (1981) generalises the static PCA by placing the analysis in the frequency domain. First, the spectral density of the vector (X_t) is estimated using a consistent spectral density estimator, denoted $\hat{S}(\omega)$, for a frequency $\omega \in]0, 2\pi]$. Then, the eigenvectors corresponding to the largest q eigenvalues of this spectral matrix are calculated. Finally, one returns to the time domain by applying the inverse Fourier transform to these eigenvectors, to recover the estimators of the time series in dynamic principal components (see next section).

Brillinger (1981) obtains distributional results when N is fixed and T tends to infinity. Forni et al. (2000) show that the dynamic PCA provides a consistent estimate of the common component when both N and T increase. Forni et al. (2004) discuss the coherence conditions and convergence rates. It has been shown that the principal components are convergent estimators of factors, both in the static context (Bai and Ng, 2002; Stock and Watson, 2002; Bai, 2003) and in the dynamic context (Forni et al., 2000, 2004).

Approximate factor models have several advantages over strict models. They are flexible and appropriate under general hypotheses on measurement errors, and usually on the cross-correlation of idiosyncratic components. The misspecification error resulting from the approximate structure of the idiosyncratic component disappears when N and T are large, as long as the cross-correlation of the idiosyncratic processes is relatively small and that of the common components increases across the transverse dimension when N increases. These conditions are introduced in Chamberlain and Rothschild (1983) and are used, reinterpreted and expanded in Connor and Korajczyk (1986, 1988, 1993), Forni and Lippi (1997), Forni et al. (2000) and Stock and Watson (2002). In short, approximate factor models have two important advantages over traditional factor models:

1. The idiosyncratic components can both be weakly mutually correlated and show little heteroskedasticity. This can reflect the condition in which all the eigenvalues of the idiosyncratic variance-covariance matrix $\Sigma_\xi = E(\xi_t \xi_t')$ are bounded. Thus, the absolute mean of the covariances is bounded, that is, $\lim_{N \rightarrow \infty} N^{-1} \sum_{i=1}^N \sum_{j=1}^N |E(\xi_{it} \xi_{jt})| < \infty$ (Stock and Watson, 2002).⁶
2. In this type of model, it is possible to have a weak correlation between the factors (F_t) and the idiosyncratic components (ξ_t) .

3.3. Recent extensions of approximate factor models

Among the recent extensions of dynamic factor models when N is large, we would like to mention FAVAR models, models whose parameters vary over time and mixed frequency models. The applications of these different types of models are presented in Section 6. We will see that these models assume that the variables are stationary. For the development of dynamic factor models on data that are non-stationary, we refer to, for example, Peña and Poncela (2006a, 2006b). Moreover, Banerjee and Marcellino (2009) have extended the FAVAR models to the factor-augmented error correction model, which makes it possible to integrate variables that are non-stationary.

3.3.1. FAVAR models

To remedy the problem of missing variables, generally encountered in traditional VAR and structural VAR (SVAR) modelling, Bernanke, Boivin and Elias (2005) propose using factor-augmented VAR (FAVAR) models, particularly in the context of monetary policy analysis.

The FAVAR model can be described by the following equation:

$$X_t = \Lambda F_t + B X_{t-1} + \xi_t \quad (22)$$

where (X_t) represents the endogenous variables of a traditional VAR model, such as in Bernanke et al. (2005) and Boivin et al. (2009), (F_t) the common factor, Λ the weighting matrix, and (ξ_t) the idiosyncratic component. In Stock and Watson (2005a), B is a diagonal matrix $D(L)$ with a lag polynomial $\delta_i(L)$ on the i^{th} diagonal. It is also conceivable to specify a short-term dynamics, of the first-order autoregressive type, for example, on the common factor (F_t) and on the idiosyncratic component (ξ_t) .

Stock and Watson (2005a) propose using an iterative procedure to estimate the FAVAR model given by Equation (22). This procedure begins with an initial estimation of the static factor $\hat{F}_t$ using a PCA. Then, the weighting matrix $\hat{\Lambda}$ and the coefficients $\hat{B}$ are estimated by ordinary least squares. Finally, the $\hat{F}_t$ factors are re-estimated by the principal components of $X_t - \hat{B}X_{t-1}$ and this procedure is iterated until convergence. Boivin et al. (2009) also use this iterative procedure with $B = 0$ as the initial estimation. Bernanke et al. (2005) propose estimating the unobservable factors in two steps:

1. the principal components of the informational variables are calculated by ignoring the presence of the observable variables;
2. Equation (22) is estimated by integrating the factors estimated in the previous step.

Other works on monetary policy analysis involve the estimation of factors using a dynamic factor model, these estimated factors then being introduced into a VAR model as additional regressors (see, for example, Giannone et al., 2004, and Favero et al., 2005).

3.3.2. Time-varying parameter models

Some authors are also interested in factor models for which, for example, the weights grouped in the matrix Λ of Equation (1) vary over time (see, for example, Motta, Hafner and von Sachs, 2011). This type of approach is promising since it makes it possible to integrate structural changes with respect to the source and amplitude of the shocks into the modelling, as well as their channels of transmission to the economy. This type of

modelling also incorporates non-linearity into relations, which makes it possible to assess whether behaviours change over the course of a business cycle.

In the area of FAVAR models, some recent papers incorporate a dynamic structure by allowing a change over time of the: i) weights; ii) autoregressive dynamics of the factors; or, iii) of the variance of innovations. Thus, a possible specification of a time-varying FAVAR (TV-FAVAR) model is as follows:

$$X_t = \Lambda_t F_t + B_t X_{t-1} + \xi_t \quad (23)$$

with $E(\xi_t) = 0$ and $E(\xi_t \xi_t') = \Sigma_{\xi t}$.

Generally, the FAVAR model given by Equation (23) is estimated in a Bayesian context, as is done in the papers by Del Negro and Otrok (2008), Mumtaz and Surico (2009) or Baumeister, Liu and Mumtaz (2010). Similarly, Kose, Otrok and Whiteman (2003) propose an estimation of dynamic factor models based on a Monte Carlo Markov-Chain (MCMC)⁷ approach in a Bayesian context.

Eickmeier, Lemke and Marcellino (2011) propose an alternative approach by developing a two-step standard estimation model by maximum likelihood. The first step consists in estimating the factors in a static context (see Section 4). Then, in a second step, the model with non-constant parameters is estimated equation by equation. Thus, each univariate regression equation is put in a space-state form, then estimated in the traditional fashion using a Kalman filter. This approach requires assuming that FAVAR model equations are conditionally independent.

3.3.3. Mixed frequency models

Finally, in the case of a large number N of available variables in the data set, variables with different periodicities can be managed using a mixed data sampling (MIDAS) regression proposed by Ghysels et al. (for a presentation, see Ghysels, Sinko and Valkanov, 2007). The MIDAS approach can be used to explain a sampled variable with a certain periodicity (such as annual or quarterly) by variables with a higher periodicity (such as monthly or daily) without having to first aggregate the higher periodicity data. This approach is based on a standard linear regression equation but involves the estimation of a weight function depending on a hyper-parameter with a smaller dimension compared with the initial dimension of the problem.

Let us assume for example, that we are seeking to estimate the quarterly GDP growth rate of an economy, denoted (y_t) , assumed to be stationary, for a number of quarters T , the index t here designating the quarter. Let us also assume that, with the help of one of the methods shown in the next section, we have estimated a single monthly stationary factor $(\hat{F}_t^{(m)})$, (that is, $r = 1$) based on a large monthly data set. We observe $m = 3$ times $(\hat{F}_t^{(m)})$ over the period $[t - 1, t]$. The standard MIDAS equation makes it possible to link the quarterly variable to be explained to the monthly estimated factor as follows:

$$y_t = c_0 + c_1 G(\theta) \hat{F}_t^{(m)} + \varepsilon_t, \quad (24)$$

where c_0 and c_1 are parameters to be estimated, and where (ε_t) is assumed to be a Gaussian white noise of finite variance that will also need to be estimated. The term $G(\theta)$ controls the polynomial weights, which allows the mixing of frequencies. In fact, the

MIDAS specification smoothes the past values of the $(\hat{F}_t^{(m)})$ using the polynomial $G(\theta)$ of the following form:

$$G(\theta) = \sum_{k=1}^K gk(\theta) L^{(k-1)/m} \quad (25)$$

where K is the number of points on which the smoothing operates, L is the lag operator, so that, for any monthly variable $x_t^{(m)}$, $L^{1/m} x_t^{(m)} = x_{t-1}^{(m)}$ and $g_k(\cdot)$ is the weight function, which can take various forms. As in Ghysels et al. (2007), we generally use a two-parameter Almon polynomial $\theta = (\theta_1, \theta_2)$ so that:

$$gk(\theta) \equiv gk(\theta_1, \theta_2) = \frac{\exp(\theta_1 k + \theta_2 k^2)}{\sum_{k=1}^K \exp(\theta_1 k + \theta_2 k^2)} \quad (26)$$

The θ parameter is part of the estimation problem. It is influenced by the information contained in the last K values of $(\hat{F}_t^{(m)})$, the size of the window K being an exogenous parameter. Other specifications can be considered in the literature for Equation (24), particularly by adding monthly explanatory variables or autoregressive terms for the target variable (y_t). Similarly, other weight functions can be considered.

In terms of application, Marcellino and Schumacher (2010) propose an approach in which they first estimate the monthly factors based on a data set of 111 representative variables of the German economy. Then use these factors to forecast the quarterly German GDP using a MIDAS regression (so-called factor-MIDAS approach). The authors show the usefulness of such an approach in making better use of the most recent data for short-term macroeconomic forecasting purposes. As for the MIDAS approach itself, some examples of applications in forecasting are found in the papers by Clements and Galvao (2008), and Ferrara and Marsilli (2013).

4. Estimation of factor models for large N

In this section, we present the main estimation methods of factor models, whether static or dynamic, when the number of variables is high (large N). In this case, the usual methods based on maximising likelihood run into the problem of the dimension of the parameter to be estimated.

4.1. Static factor models: The Stock and Watson (2002) approach

One of the first approximate factor models is the one proposed by Stock and Watson (SW) (2002), which is based on a static PCA. The PCA is used since it allows for the estimation of both the parameters and the factors of the model given by Equation (1) by maximising the variance explained by the initial variables, for a small number r of static factors (F_t). The main aim of the SW approach is to approximate the factors by a linear combination of the data $\hat{F}_{j,t} = \hat{W}_j' X_t$, for $j = 1, \dots, r$, that maximises the variance of the estimated factors $\hat{W}_j' \hat{\Sigma}_x \hat{W}_j$, where $\hat{\Sigma}_x = (1/T) \sum_{t=1}^T X_t X_t'$ is the empirical variance-covariance matrix of the vector of the initial standardised data X_t .

Under the following normalisation assumption: $\hat{W}_j' \hat{W}_{j'} = 1$ for $j = j'$, and $\hat{W}_j' \hat{W}_{j'} = 0$ for $j \neq j'$, the maximisation problem can be transformed into the solution of an eigenvalues' problem:

$$\hat{\Sigma}_x \hat{W}_j = \hat{\mu}_j \hat{W}_j \quad (27)$$

where $\hat{\mu}_j$ is the j^{th} eigenvalue and $\hat{W}_j$ is the associated eigenvector of dimension $(N \times 1)$. Once they have been calculated, the highest N eigenvalues are classified in decreasing order. The eigenvectors are, in turn, classified in decreasing order with respect to the highest r eigenvalues. The factors proposed by SW are written as follows:

$$F_t^{\text{SW}} = \hat{W}'X_t, \quad (28)$$

where $\hat{W}$ is the matrix of dimension $(N \times r)$ of the stacked eigenvectors $\hat{W} = (\hat{W}_1, \dots, \hat{W}_r)$.⁸

However, the Stock and Watson approach does not allow for use of the different dynamics that may exist between the variables used. To take account of this dynamic structure in factor models, several alternatives to the static factor model have been proposed in the literature. Specifically, there are two main types of dynamic factor models or approaches. Developed by Doz, Giannone and Reichlin (2011, 2012), the first approach is based on a space-state representation of models in the time domain. The second, proposed by Forni, Hallin, Lippi and Reichlin (2004, 2005), is based on the spectral domain.

4.2. Dynamic factor models

4.2.1. Time domain approach

Doz, Giannone and Reichlin (DGR) (2011, 2012) propose a dynamic factor model that can be represented in a space-state form. Specifically, they estimate their dynamic factor model using two different approaches. The first one is the so-called two-step approach (DGR, 2011). The second one is based on the quasi-maximum likelihood (DGR, 2012).

According to DGR (2011), for a number r of factors and q of dynamic shocks, the estimation is carried out in two steps:

1. $\hat{F}_t$ is estimated using a PCA, as an initial estimate, then Equations (6) and (11) are estimated using the estimated factor $\hat{F}_t$, to obtain both $\hat{\lambda}_i^*(L)$ and the variance-covariance matrix of the residuals $\hat{\xi}^*$, denoted $\hat{\Sigma}_{\xi^*}$. To obtain an estimate of $C(L)$, appearing in Equation (10), DGR (2011) apply a decomposition of eigenvalues to the matrix $\hat{\Sigma}_{\xi^*}$ by taking into account the number of dynamic shocks q . Let us introduce the matrix M of dimension $(r \times q)$ corresponding to the largest q eigenvalues and the matrix P of dimension $(q \times q)$ containing the largest q eigenvalues in its diagonal and zeros elsewhere. The estimate of $C(L)$ is, then, obtained by $\hat{C}(L) = M \times P^{-\frac{1}{2}}$.
2. The coefficients and parameters of the system described by Equations (6) and (11) are considered to be known and provided by the first step. The model is written in a space-state form and the Kalman filter is applied to obtain new estimates of the factors.

In their alternative approach, DGR (2012) estimate an approximate dynamic factor model using the quasi maximum likelihood method.⁹ The main aim of this approach is to consider the strict factor model as a misspecification of the approximate factor model and to analyse the properties of the maximum likelihood indicator of the factors under this misspecification. This estimator is called the *quasi* maximum likelihood in the sense of White (1982). By analysing the properties of the maximum likelihood estimator under several sources of misspecifications, such as an omitted serial correlation of the observations or a cross-sectional correlation of the idiosyncratic components, DGR (2012) show that these misspecifications do not affect the robustness of the common factors, particularly for fairly large N and T . More specifically, this estimator is a valid parametric alternative for the estimator resulting from a PCA. The model defined by means of

Equations (6) and (11) can be put in a space-state form, with a number of states equal to the number of common factors r . It is noteworthy that the estimation of the parameters of the model, particularly the common factors, by the quasi maximum likelihood can be approximated by their anticipated values, using the Kalman filter.¹⁰

These dynamic factor models have also been called restricted dynamic factor models, since the r static factors are caused by a number q of dynamic factors, with $q \leq r$ (Forni et al., 2005; Hallin and Liska, 2007).

Kapetanios and Marcellino (2004) also propose an approach based on a space-state representation. Their approach is based on the use of specific subspaces in which the factors are estimated. This subspace algorithm can be used to estimate factors without having to specify or identify the model entirely in its space-state form.

4.2.2. Frequency domain approach

In a series of papers, Forni et al. (FHLR) (2000, 2003, 2004, 2005) propose a dynamic PCA in the frequency domain, also called a generalised dynamic factor model, to estimate dynamic factors.¹¹ The purpose of their model is to identify the dynamic structure of a factor model. The dynamic factor model is given by Equations (6) and (7). The method proposed by FHLR makes it possible to estimate dynamic factors and then obtain the static factors from the estimated dynamic factors. The approach proposed by FHLR aims to estimate both the dynamic factors and their covariances. This estimation is performed to maximise the variance of the common component under certain orthogonality restrictions. The optimisation program is likened to a problem to determine the dynamic eigenvalues of the spectral density matrix of the variables observed. The spectral matrix $I_x(\omega)$ of X_t is estimated using a representation of time series in the frequency domain for each frequency ω in the interval $[0, 2\pi]$. The estimated spectral matrix contains information on both the cross-correlation between variables and their dynamic relations. Thus, we write $\hat{\Sigma}_x(\tau)$ the estimated autocovariance matrix between X_t and $X_{t-\tau}$ for a particular lag τ . The estimated spectral density of the vector of observed variables is given by:

$$\hat{I}_x(\omega_h) = \sum_{\tau=-H}^H \hat{\Sigma}_x(\tau) \left(1 - \frac{|\tau|}{H+1}\right) e^{-i\tau\omega_h} \quad (29)$$

for each Fourier frequency $\omega_h = \frac{2\pi h}{2H+1}$ and for each $h = 0, \dots, 2H$, with i representing the imaginary number such as $i^2 = -1$. For each frequency ω_h , the dynamic eigenvalues and eigenvectors resulting from $\hat{I}_x(\omega_h)$ are calculated. The eigenvectors are classified in decreasing order. More specifically, the eigenvectors $\hat{p}_l(\omega_h)$, of dimension $(N \times 1)$, are collected for $l = 1, \dots, q$ (the highest q eigenvalues). To return to the time domain, the eigenvectors are obtained based on the inverse Fourier transform:

$$\hat{P}(L) = \sum_{\tau=-H}^H \hat{P}_{\tau} L^{\tau}, \text{ with } \hat{P}_{\tau} = \frac{1}{2H+1} \sum_{h=0}^{2H} \hat{p}_l(\omega_h) e^{i\tau\omega_h}, \quad (30)$$

for $\tau = -H, \dots, H$ and $j = 1, \dots, q$. The j^{th} dynamic principal component $\hat{F}_{j,t}$, is, then, given by the j^{th} component of $\sum_{l=1}^q \hat{P}'_l(L) \hat{P}_l(L) X_t$.

Thus, the dynamic principal components are obtained from a decomposition of the spectral density matrix into dynamic eigenvalues and eigenvectors. This breakdown also makes it possible to divide the spectral density matrix into a spectral density matrix of common components $I_x(\omega)$ and a spectral density matrix of idiosyncratic components $I_{\xi}(\omega)$.

Moreover, the estimator of the frequency domain is reduced to a symmetric filter. This presents problems at the end of samples, particularly when future observations are useful to estimate the principal components. To remedy this problem, FHLR (2005) suggest a refinement of their procedure that maintains the advantages of the dynamic approach, while basing the estimation of the common components on an asymmetric filter.¹²

With this procedure, the space of the factors is approximated by adding the r static factors rather than the q dynamic principal components. However, the mean resulting from the r contemporary static factors is based on information from the dynamic approach. The estimation of the model, then, consists in maximising the variance of the common components or minimising the variance of the idiosyncratic components. Thus, a convergent estimator of the spectral density matrix of the common component is given by:

$$\hat{I}_x(\omega) = \hat{P}(\omega)\Omega(\omega)\hat{P}'(\omega), \quad (31)$$

where $\Omega(\omega)$ is a diagonal matrix of dimension $(q \times q)$ containing the largest q dynamic eigenvalues on the diagonal and $\hat{P}(\omega) = (\hat{P}_1(\omega), \dots, \hat{P}_q(\omega))$ is a matrix of dimension $(N \times q)$ containing the eigenvectors corresponding to the frequency ω . We, then, deduce the spectral density matrix of the idiosyncratic component:

$$\hat{I}_\xi(\omega) = \hat{I}_x(\omega) - \hat{I}_x(\omega) \quad (32)$$

This estimation in the frequency domain is carried out in two steps. The first one is based on the dynamic approach, by which we obtain the variance-covariance matrices of the common components $\hat{I}_x(\omega)$ and idiosyncratic components $\hat{I}_\xi(\omega)$, respectively, estimated by an inverse Fourier transform. Thus, the variance-covariance matrix of the common components is estimated as follows:

$$\hat{\Sigma}_x(\tau) = \frac{1}{2H+1} \sum_{h=0}^{2H} \hat{I}_x(\omega_h) e^{i\tau\omega_h} \quad (33)$$

for $\tau = -H, \dots, H$. The variance-covariance matrix of the idiosyncratic components is estimated in the same way:

$$\hat{\Sigma}_\xi(\tau) = \frac{1}{2H+1} \sum_{h=0}^{2H} \hat{I}_\xi(\omega_h) e^{i\tau\omega_h} \quad (34)$$

In a second step, this information is used to construct the factor space using the r aggregated means. More specifically, the variables are weighted in terms of the common-to-idiosyncratic variance ratio, obtained by means of the variance-covariance matrices estimated in the first step. These r aggregated means are defined as the solutions to a generalised principal components problem and they have the advantage of minimising the idiosyncratic quadratic errors of the common factors by selecting only those variables with the highest common-to-idiosyncratic variance ratio. The number of these aggregated means is equal to $r = s(p+1)$, which represents the rank of the spectral density matrix of the common factors, where s indicates the number of lags for $\lambda_i^*(L)$ in Equation (10). FHLR (2005) show that to determine the number of aggregated means r , the problem of maximisation can be converted into a problem of generalised eigenvalues:

$$\hat{\Sigma}_x(0)\hat{Z}_j = \hat{\mu}_j \hat{\Sigma}_\xi(0)\hat{Z}_j \quad (35)$$

where $\hat{\mu}_j$ is the j^{th} generalised eigenvalue, $\hat{Z}_j$ its associated eigenvector of dimension $(N \times 1)$, and $\hat{\Sigma}_\kappa(0)$ and $\hat{\Sigma}_\xi(0)$ are the contemporaneous variance-covariance matrices ($\tau = 0$) of the common and idiosyncratic components, respectively. Moreover, FHLR (2005) impose the following normalisation $\hat{Z}_j' \hat{\Sigma}_\xi(0) \hat{Z}_{j'} = 1$ for $j = j'$ and $\hat{Z}_j' \hat{\Sigma}_\xi(0) \hat{Z}_{j'} = 0$ for $j \neq j'$. Then, the eigenvalues are classified in decreasing order and the factors obtained correspond to the product of the r eigenvectors corresponding to the highest eigenvalues and the vector X_t . The estimator proposed by FHLR (2005) is written as follows:

$$F_t^{\text{FHLR}} = \hat{Z}' X_t \quad (36)$$

where $\hat{Z} = (\hat{Z}_1, \dots, \hat{Z}_r)$ is a matrix of dimension $(N \times r)$ of the stacked eigenvectors.

To conclude this section, we point out that the estimation methods proposed are relatively recent and the literature does not yet have sufficient perspective to systematically choose one method over another. Because it is easy to use, the Stock and Watson (2002) approach is naturally attractive and the empirical results, particularly in a forecasting context, show that the results yielded by this approach are not significantly poorer than the other approaches in terms of forecasting error (on this point, see D'Agostino and Giannone, 2012; Barhoumi, Darné and Ferrara, 2013).

The recent literature has also looked at the estimation of factor models in a Bayesian context. This approach makes it possible to reduce uncertainty regarding the parameters by first applying hypotheses on the distributions of these parameters. In this regard, we refer interested readers to Kose, Otrok and Whiteman (2003, 2008) or Lopes and West (2004).

Finally, the asymptotic properties of the estimators presented above are proved under the simple hypothesis “ N and T tend to infinity”, the interpretation of which is sometimes rather vague. Bai (2003) and Forni et al. (2004) emphasise that asymptotic properties, such as convergence, hold along specific trajectories $\{(N, T(N)); N \in \mathbb{N}\}$. For example, a property that holds for $\min(N, T)$ holds along the entire trajectory $(N, T(N))$, while a property that holds for $N = O(T^k)$ requires that the number of observations T be at least on the order of $N^{\frac{1}{k}}$. In fact, three concepts of limits exist: i) sequential, ii) pairwise, and iii) simultaneous. Let $g(N, T)$ be a function that one wishes to study. A sequential limit stretches N and T to infinity, one after the other. A pairwise limit stretches (N, T) to infinity only along a particular trajectory, which can be denoted $\lim_{N, T \rightarrow \infty} g(N, T(N))$. A simultaneous limit authorises (N, T) to increase along all possible trajectories: $\lim_{N, T \rightarrow \infty} g(N, T)$. It is noteworthy that the existence of a simultaneous limit implies the existence of a pairwise limit and a sequential limit, but the reverse is not true. Another approach using the theory of random matrices postulates that N and T tend to infinity with $N/T \rightarrow c \in (0, \infty)$, where c is a constant. For a more detailed discussion, see Bai and Ng (2008b), and Harding (2009).

5. Selection of the number of factors

An important step in the statistical analysis of static and dynamic factor models is the preliminary identification of the number of factors. A number of papers focus on the problem of determining the number of factors. For example, Forni and Reichlin (1998) suggest a graphic approach to identify the number of factors when $N \rightarrow \infty$ and T is fixed

but no theory is proposed. Stock and Watson (1998) modify the BIC criterion to select the optimal number of factors in forecasting when $N, T \rightarrow \infty$ with

$$\frac{\sqrt{N}}{T} \rightarrow \infty.$$

However, their criterion is restrictive since it requires that $N \gg T$ and it is appropriate only in a forecasting context. Forni et al. (2000) consider a multivariate version of the AIC criterion but no theoretical or empirical property is known for their criterion.

In this section, we present the criteria most used in the empirical literature, including: Bai and Ng (2002) and Alessi et al. (2010) for static factor models; and, those of Stock and Watson (2005a), Amengual and Watson (2007), Bai and Ng (2007), Hallin and Liska (2007), and Breitung and Pigorsch (2013) for dynamic factor models. Note that Barhoumi, Darné and Ferrara (2013) have compared those criteria using a forecasting framework.

5.1. Selection of the number of factors for static factor models

To specify the number of factors, Bai and Ng (2002) suggest using information criteria to select the optimal number of static factors r when N and T tend to infinity. Bai and Ng (2002) propose information criteria based on the quality of adjustment of the model to the data measured by the variance $V(j, F)$ such that:

$$V(j, F) = (NT)^{-1} \sum_{t=1}^T (X_t - \hat{F}_t)' (X_t - \hat{F}_t), \quad (37)$$

where j is a given number of factors such as $\hat{F}_t = (\hat{F}_{1t}, \dots, \hat{F}_{jt})'$. Thus, if the number of factors j increases, the variance of the factors increases mechanically and the sum of the squares of the residuals decreases in turn. Bai and Ng (2002) suggest introducing a penalty function in the criterion to be optimised and propose the following three criteria, corresponding to different penalty functions:

$$IC_1(j) = \ln(V(j, F)) + j \left(\frac{N+T}{NT} \right) \ln \left(\frac{NT}{N+T} \right), \quad (38)$$

$$IC_2(j) = \ln(V(j, F)) + j \left(\frac{N+T}{NT} \right), \quad (39)$$

$$IC_3(j) = \ln(V(j, F)) + j \left(\frac{\ln C_{NT}^2}{C_{NT}^2} \right), \quad (40)$$

where $C_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$ and $\ln$ denotes the natural logarithm. The estimation of the number of factors r is obtained by minimisation of the information criteria for $j = 0, \dots, r_{max}$, where r_{max} is the maximum number of static factors. These criteria reflect the trade-off between the quality of the adjustment and the risk of overadjustment.¹³ Bai and Ng (2002) show that their criteria are robust to the presence of a heteroskedastic component in the time and cross-section dimensions between variables, but also in the presence of weak serial and cross-section dependence.

Subsequently, Alessi et al. (2010) extend this criterion by modifying the strength of the penalty function that appears in the preceding three criteria given by Equations (38), (39), and (40). Alessi et al. (2010) propose an alternative to the criteria proposed by Bai and Ng (2002) by multiplying the penalty function by a positive constant c , suggested originally by

Hallin and Liska (2007), representing the strength of the penalty function. The authors, thus, propose the following two criteria:

$$IC_1^*(j) = \ln(V(j, F)) + c \cdot j \cdot \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right) \quad (41)$$

$$IC_2^*(j) = \ln(V(j, F)) + c \cdot j \cdot \left(\frac{N+T}{NT}\right), \quad (42)$$

where $V(j, F)$ is given by Equation (37). The estimation of the number of factors r is obtained by minimisation of the information criteria IC_1^* and IC_2^* for $j = 0, \dots, r_{\max}$, where $r_{\max}$ is the maximum number of static factors. The procedure for the selection of the number of static factors depends both on the variance of the number of estimated factors $V_c(r)$ (for N and T tending to infinity) and on the constant $c \in [0, c_{\max}]$. Alessi et al. (2010) suggest estimating this variance $V_c(r)$ by reiterating the procedure for estimating r for a finite number of subsets of the initial N variables, also making the number of observations T vary.

Kapetanios (2010) proposes a concurrent method to the information criterion to estimate the number of static factors, based on the random matrix theory. His approach is based on a series of tests on the largest eigenvalues of the variance-covariance matrix of the initial data, which we have denoted Σ_x . Alternative procedures have been suggested by Yao and Pan (2008) and Onatski (2010).

5.2. Selection of the number of factors for dynamic factor models

5.2.1. The Bai and Ng (2007) criterion

In the context of dynamic factor models, the number of dynamic shocks q (for the estimation of factors in dynamic principal components and their space-state form) can be determined using the Bai and Ng (2007) information criterion. This criterion is obtained by considering the r estimated static factors as given and then estimating a VAR model of order p on these factors, where the order p is selected using the BIC criterion. Next, a spectral decomposition of the variance-covariance matrix of the estimated residuals of the VAR model, denoted $\hat{\Sigma}_e$, of dimension $(r \times r)$, is calculated. Then, the j^{th} ordered eigenvalue $\hat{c}_j$, where $\hat{c}_1 \geq \hat{c}_2 \geq \dots \geq \hat{c}_r \geq 0$, is recovered. Finally, for $l = 1, \dots, r-1$, Bai and Ng (2007) propose the following two quantities:

$$\hat{D}_{1,l} = \left(\frac{\hat{c}_{l+1}}{\sum_{j=1}^r \hat{c}_j} \right)^{1/2}$$

$$\hat{D}_{2,l} = \left(\frac{\sum_{j=l+1}^r \hat{c}_j}{\sum_{j=1}^r \hat{c}_j} \right)^{1/2}$$

where $\hat{D}_{1,l}$ represents a measure of the marginal contribution of the $l+1^{\text{th}}$ eigenvalue and $\hat{D}_{2,l}$ represents a measure of the cumulative contribution of the eigenvalues, under the hypotheses that $\hat{\Sigma}_e$ is the unit matrix of dimension $r \times r$ and that $c_l = 0$ for $l > q$.¹⁴

Thus, according to the selected marginal contribution measure, the number of dynamic factors q is obtained by minimising:

$$\left\{ l \text{ such that } : \hat{D}_{1,l} \leq \frac{c}{\min(N^{\frac{2}{5}}, T^{\frac{2}{5}})} \right\}$$

or:

$$\left\{ l \text{ such that } : \hat{D}_{2,l} \leq \frac{c}{\min(N^{\frac{2}{5}}, T^{\frac{2}{5}})} \right\}$$

Bai and Ng (2007) suggest using $c = 1$ based on Monte Carlo simulations.

In practice, these different criteria are used at three stages:

1. One of the Bai and Ng (2002) criteria is used to determine the optimal number of factors $r \in \{1, \dots, r_{max}\}$ in a static context.¹⁵
2. A VAR(p) is estimated on these r estimated factors and the order p of the VAR is selected to minimise the BIC criterion.
3. The Bai and Ng (2007) criteria are applied to the variance-covariance matrix or correlation matrix of the residuals (ε_t) of the VAR(p) to obtain the optimal number of dynamic factors q .

5.2.2. The Stock and Watson (2005a) and Amengual and Watson (2007) criteria

Stock and Watson (2005a), and Amengual and Watson (2007) show that the Bai and Ng (2002) estimator can be used to estimate the number of dynamic factors. To do this, they propose applying this estimator to the errors resulting from the projection of observed data on the lagged values of static factors, that is, to

$$v_t = X_t - \sum_{\tau=1}^p \Lambda \Phi(L) F_{t-\tau}.$$

They propose two ways of estimating the errors:

$$\hat{v}_t^A = X_t - \sum_{\tau=1}^p \hat{\Lambda} \hat{\Phi}_\tau \hat{F}_{t-\tau}$$

$$\hat{v}_t^B = X_t - \sum_{\tau=1}^p \hat{\Pi}_\tau \hat{F}_{t-\tau}$$

where $(\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p)$ are the ordinary least squares estimators of the regression of $\hat{F}_t$ on $(\hat{F}_{t-1}, \dots, \hat{F}_{t-p})$ and $(\hat{\Pi}_1, \hat{\Pi}_2, \dots, \hat{\Pi}_p)$ are the ordinary least squares estimators of the regression of X_t on $(\hat{F}_{t-1}, \dots, \hat{F}_{t-p})$

5.2.3. The Breitung and Pigorsch (2013) criteria

Breitung and Pigorsch (2013) also propose two information criteria to select the number of dynamic factors. Their criteria are based on an analysis of the canonical correlations of static factors (obtained by a principal components analysis) and depend on the estimation of a VAR(p) model on these factors, where the order p is selected by the BIC criterion.

The first criterion is based on the following statistic:

$$\zeta(q^*) = \tilde{C}_{NT}^{2-\delta} \sum_{j=1}^{r-q^*} (1 - \tilde{\lambda}_j)$$

where $\tilde{C}_{NT}^{2-\delta} = (2-\delta)N^{-1} + (2-\delta)T^{-1}$, with $0 < \delta < 2$, and $\tilde{\lambda}_j$ are values resulting from the solution to the following problem $|\tilde{\lambda}_j \tilde{S}_{00} - \tilde{S}_{01} \tilde{S}_{11}^{-1} \tilde{S}_{01}'| = 0$, with

$$\tilde{S}_{00} = \sum_{t=\tau+1}^T \hat{F}_t \hat{F}_t', \tilde{S}_{01} = \sum_{t=\tau+1}^T \hat{F}_t \hat{G}_{t-1}', \tilde{S}_{11} = \sum_{t=\tau+1}^T \hat{G}_{t-1} \hat{G}_{t-1}', \text{ and } \hat{G}_{t-1} = [\hat{F}_{t-1}', \dots, \hat{F}_{t-\tau}'].$$

The number of dynamic factors can be estimated using a large number of q^* resulting from this sequence $q^* = r-1, r-2, \dots, 0$, where the statistic $\zeta(q^*)$ is larger than the level of the threshold κ , or:

$$q = \max\{q^* \text{ such that: } \zeta(q^*) > \kappa\}$$

Breitung and Pigorsch (2013) suggest using the following values of the parameters: $\tau = 1, \delta = 0.5$, and $\kappa = 1$.

The second criterion is based on the following statistic:

$$LR(q^*) = T \sum_{j=r-q^*+1}^r \tilde{\lambda}_j$$

where the null hypothesis is $H_0: q^* = q$ as against the alternative $H_1: q^* < q$.

5.2.4. The Hallin and Liska (2007) criterion

Hallin and Liska (2007) develop an information criterion for generalised dynamic factor (GDF) models. This criterion is based on the spectral density matrix of the observations. It is written as follows:

$$IC_{2,N}^T(j) = \ln \left[\frac{1}{N} \sum_{s=j+1}^N \frac{1}{2M_T+1} \sum_{h=-M_T}^{M_T} \lambda_{N_s}^T(\omega_h) \right] + c \cdot j \cdot p(N, T), \quad (43)$$

with $0 \leq j \leq q_{\max}$, $\omega_h = \frac{\pi h}{M_T + \frac{1}{2}}$, for $h = -M_T, \dots, M_T$ with the truncation parameter $M_T > 0$, c is a positive constant such that $c = [0.01, 0.02, \dots, 3.00]$, $N_s < N$, is the number of variables contained in a given subset and $p(N, T)$ is a penalty function¹⁶ such that:

$$p(N, T) = (-M_T^{-2} + M_T^{1/2} T^{-1/2} + N^{-1}) \times \ln[\min(N, M_T^2, M_T^{-1/2} T^{1/2})]$$

The eigenvalues $\lambda_{N_s}^T$ result from $\hat{\Gamma}_x(\omega)$, which represents the spectral density matrix estimator of X_t with $\omega \in [-\pi, \pi]$. The number of estimated factors is, then, given by:

$$q = \operatorname{argmin} IC_{2,N}^T(j)$$

$$0 \leq j \leq q_{\max}$$

The procedure for selecting the number of dynamic factors is similar to that used by Alessi et al. (2010), that is, by examining the variance of the number of estimated factors, $V_c(r)$ for N and T tending to infinity and for an interval of values for the constant c . In their numerical illustration, Hallin and Liska (2007) propose retaining $M_T = \lceil 0.75\sqrt{T} \rceil$ and $q_{\max} = 13$.

6. Recent results in the empirical literature

Applications of dynamic factor models abound in the empirical economic literature. A few examples include asset pricing models (Ross, 1976), consumer theory (Gorman, 1981;

Lewbel, 1991), and the assessment of performance and risk measurement in finance (Campbell et al., 1997). In this section, we present some recent applications of these models that underscore the interest of this approach for: i) the construction of short-term economic indicators; ii) macroeconomic forecasting; and, iii) international macroeconomics and monetary policy analysis.

6.1. Tools for short-term economic monitoring

Dynamic factor models are useful for developing economic activity indicators based on the mass of data available to short-term forecasters. These models can be used to synthesise large datasets into a composite indicator that reflects the most relevant data available on a given date.

One of the most common indicators to which forecasters working on the US economy refer is the Chicago Fed National Activity Indicator (CF-NAI) developed using the Stock and Watson (1999) approach. This indicator is based on 85 monthly series representative of the US economy, covering production, income, employment, personal consumption, housing, sales, inventories and orders. The CF-NAI corresponds to the first factor estimated in a principal components analysis. A value close to zero means that activity is close to its long-term trend. As well, the CF-NAI can be used to detect recessions in the US by using historical estimated thresholds. The Philadelphia Federal Reserve publishes each week, a daily economic activity indicator that is based on the mixed frequency model in Aruoba, Diebold and Scotti (2009), and includes daily, weekly, monthly and quarterly data. Given its high frequency, this indicator is interesting for forecasting since it makes it possible to provide an early signal very rapidly.

In the euro area, the Centre for Economic Policy Research (CEPR) has for some years now been disseminating the EuroCoin indicator developed at the Bank of Italy by Altissimo et al. (2001, 2010). The purpose of this indicator is to estimate a monthly GDP growth in the euro area for the coincident quarter using smoothing to remove very short-term effects (frequency of less than one year). The model used to synthesise the information is a generalised dynamic factor model proposed by Forni, Hallin, Lippi and Reichlin (2000) applied to a very large number of variables. From an economic standpoint, this measure corresponds to a kind of medium-term quarterly growth, but does not aim to estimate precisely the quarterly accounts figures provided by Eurostat. The first version of EuroCoin has a smaller variance than quarterly GDP growth. The economic outlook that it provides deviates considerably from the national accounts. CEPR has recently tried to construct a more effective new version of EuroCoin (Altissimo et al., 2010). The aim of the new indicator is the same but the number of input variables used has been reduced from 951 to 145 (IPI, monetary aggregates, interest rates, financial variables, demand indicators, surveys, trade variables and labour market variables). The input variables have not been smoothed by statistical means, which eliminates some side effects that might have introduced a bias into the old EuroCoin.

For France, Doz and Lengart (1999) have developed a summary indicator for the industry tendency survey conducted by the French National Institute of Statistics and Economic Studies (INSEE) by means of six monthly balances of opinion. This approach is now commonly used by INSEE to calculate composite indicators based on data from short-term tendency surveys in the various sectors. Clavel and Minodier (2009) extend this business activity indicator relating to industry by proposing a business climate indicator for the economy as a whole, which incorporates balances of opinion derived from the

INSEE tendency surveys conducted in the various sectors such as services, construction, and wholesale and retail sales.

Many economic activity indicators have been developed based on regime-switching factor models. Kim and Nelson (1998) propose an application for the four main US economic series (growth rates of IPI, employment, income and retail sales), also considered by the National Bureau of Economic Research (NBER) for the dating of US recessions. These same series are considered by Diebold and Rudebusch (1996), Chauvet (1998), and Chauvet and Piger (2008) who simultaneously estimate a similar model to construct a real-time business cycle indicator. In France, Nguiffo-Boyom (2006) estimates the Kim and Nelson (1998) model simultaneously on four series derived from the INSEE tendency surveys to reproduce the growth cycle, which measures the long-term deviation from trend. For the euro area, Darné and Ferrara (2011) propose a two-step dynamic factor model. They estimate a factor which follows a regime-switching model, based on a set of six business confidence index series for the six main euro area countries. The authors, thus, develop an indicator to detect the acceleration cycle in the monetary zone in real time.

As for the mixed frequency approach, Mariano and Murasawa (2003) have applied their model to estimate an economic activity indicator in the US. Mariano and Murasawa (2010) also use a version of their model to calculate a monthly US GDP series by estimating a mixed frequency VAR that includes the quarterly GDP growth series and the four monthly series traditionally used by NBER to assess the US business cycle (growth rates of employment, income, industrial production and retail sales). Cornec (2006) also uses this approach, first to provide a monthly dating of the French business cycle using two quarterly series (GDP growth rate and employment) and two monthly series (IPI and household consumption expenditure). Cornec estimates a composite activity indicator similar to that of Doz and Lenglart (1999), but includes the quarterly GDP growth rate series as supplementary information. The empirical results of this application emphasise that the contribution of GDP to the first factor is negligible in comparison with the composite business indicator relating to industry. Cornec and Deperraz (2006) use a mixed frequency model to develop an activity indicator in the French service sector based on three monthly balances, and three quarterly balances from the INSEE tendency survey relating to services. This indicator can usefully supplement the indicator on the business climate in manufacturing for short-term forecasting. Clavel and Minodier (2009) also develop a mixed frequency approach to incorporate the various short-term tendency surveys conducted by INSEE, which are sampled monthly, bimonthly and quarterly, into their business climate indicator. More recently, Camacho and Perez-Quiros (2010, 2011) develop two short-term growth indicators for Spain and the euro area using a mixed frequency factor model that also includes Markov regime-switching to take account of the business cycle. Frale et al. (2010) develop a small-dimension mixed frequency factor model to provide a measure of monthly GDP in the euro area. Finally, Frale et al. (2011) propose a mixed frequency model to estimate a monthly GDP indicator in the euro area called EuroMInd, based on a disaggregation between supply and demand. This indicator is based on the official database for the euro area developed by Eurostat (Euro-IND).

6.2. Macroeconomic forecasting

Dynamic factor models are widely used, particularly by central banks, as a tool for forecasting various macroeconomic variables, such as the GDP growth rate or inflation (for example, see the survey in Stock and Watson, 2006, or Eickmeier and Ziegler, 2008). When

the forecasting horizon covers the current period, such forecasting is termed “nowcasting” (on this point see Giannone, Reichlin and Small, 2008). The factors are estimated from monthly data used to track countries’ economic situations, such as household and business survey data (*soft data*), variables for the real economy (*hard data*), including indices of industrial production, household consumption, retail sales or new vehicle registrations, and financial variables (stock prices, oil prices, interest rates, etc.). For a given country, such a database can include several hundred variables. It is useful, therefore, to be able to synthesise this large data set into a small dimension vector to be included in standard models.

Based on asymptotic theoretical results on the convergence of estimators in this type of model, early work used the largest possible number of variables available. More recent work addresses the question whether including the largest number of variables is appropriate or not to improve the accuracy of forecasts. For example, Barhoumi, Darné and Ferrara (2010) empirically show that in the case of France, increasing the data set by disaggregation does not result in significant improvements in the accuracy of short-term GDP forecasts. Boivin and Ng (2006) identify the conditions under which expanding the database could result in less accurate factor estimates and provide empirical rules for eliminating redundant variables. It is shown that expanding the data set is not preferable if the new series add too much idiosyncratic noise and/or increase the cross correlation between idiosyncratic errors too much. Bai and Ng (2008a) use least angle regressions (LARS) type statistical methods, which are weighted regressions, to identify optimal subsets of (targeted) predictors from a large dataset. Schumacher (2010) underscores the effectiveness of this approach in using an international database to predict German growth, which is very sensitive to fluctuations in the international environment. Charpin (2009) applies this approach to French data, which also seems to provide encouraging results.

Once the factors have been estimated, the forecasting of the variable of interest Y_t over a horizon h is derived from either an autoregressive distributed lags (ARDL) type univariate regression equation (Equation 44) or a VAR-type multivariate process. When the aim is to forecast over a horizon h exceeding one step, two approaches co-exist: the recursive approach, which for a given step, uses the forecasts made for the previous steps; and, the direct approach, which seeks to predict the value over the horizon h directly, that is without trying to forecast the variable of interest in the previous steps. In a general context, direct forecasting of the variable over the horizon h makes it possible to reduce the forecast bias resulting from the estimation of the parameters that may appear in the case of a multistep recursive forecast (see, for example, Chevillon, 2007). In the particular context of factor models, according to simulations done by Boivin and Ng (2005), there does not seem to be a significant difference between a direct forecast and a recursive forecast if estimated factors are used. However, the direct approach is preferred in many applications.

Thus, the univariate equation of direct forecasting to a horizon h is written:

$$\hat{Y}_{t+h|t} = \hat{\alpha}_h + \sum_{j=1}^m \hat{\beta}_{hj}' \hat{F}_{t-j+1} + \sum_{j=1}^p \hat{\phi}_{hj} Y_{t-j+1}, \quad (44)$$

where $\hat{F}_t$ is the vector of dimension r of the estimated factors, m and p are the autoregressive orders, and $\hat{\beta}_{hj}$ is a vector of estimated coefficients of dimension r . The parameters α_h , β_{hj} and ϕ_{hj} depend on the horizon h since, in the context of a direct forecast, they vary on the basis of the horizon considered. The $mr + p + 1$ parameters of the model are estimated

using ordinary least squares. In Equation (44), the number of factors r can be specified by one of the tests presented previously. However, $r = 3$ is often used in practice, since three factors are often sufficient to explain a significant portion of the data variance. Three variants of the model given in Equation (44) are generally used (Stock and Watson, 2002; Boivin and Ng, 2005). The first, denoted *DI* (diffusion index) is obtained with $m = 1$ and the terms depending on Y_{t-j+1} , $j=1, \dots, p$, being suppressed in Equation (44) and includes only contemporaneous information $\hat{F}_t$. The second, *DI-AR* authorises a dynamics on the series Y_t and corresponds to $m = 1$ and $1 \leq p \leq 6$ in (44). The optimal autoregressive order p is obtained by minimisation of an AIC or BIC-type information criterion. Finally, the specification *DI-AR,Lag* of (44) corresponds to $1 \leq m \leq 3$ and $1 \leq p \leq 3$, thus allowing lags on the factors and on the variable Y_t . Once again, the optimal parameters m and p are obtained by minimisation of an information criterion. We should point out that the specification *DI-AR,Lag* is not used for a dynamic factor since such a factor is assumed to already include a time dynamics.

Boivin and Ng (2005) use a simulation to show that the differences between using static or dynamic factors are negligible in forecasting. Barhoumi, Darné and Ferrara (2010) empirically find the same result with French data and show that the specification of the model used to make the forecasts has only a marginal impact on the quality of the forecast, particularly when the number of observations is high.

One of the major problems that appears when these models are used for real-time forecasting results from the fact that data arrive in a staggered fashion, leading to missing values at the end of samples (known by forecasters as the *ragged-edge data* problem). Several solutions have been proposed in the empirical literature, such as the projection of missing data using either an autoregressive-type parametric model or moving averages, or the realignment of the database on the last points available if the number of variables is high. In the context of factor models, two-step estimation using a Kalman filter (Doz, Giannone and Reichlin, 2011) solves this problem elegantly (Giannone, Reichlin and Small, 2008; Angelini et al., 2011).

Among the many applications in forecasting the GDP growth rate, see: Stock and Watson (2002), and Banerjee and Marcellino (2006) for the US; Barhoumi, Darné and Ferrara (2010, 2013), and Bessec and Doz (2013) for France; Forni et al. (2000, 2003), Camba-Mendez and Kapetanios (2005), Marcellino, Stock and Watson (2003), Banerjee, Marcellino and Masten (2005), Ruenstler et al. (2009), and Angelini et al. (2011) for the euro area; Schumacher (2007, 2010), Schumacher and Breitung (2008), Eickmeier and Ziegler (2008), and Marcellino and Schumacher (2010) for Germany; Artis, Banerjee and Marcellino (2005) for the United Kingdom; and, Van Nieuwenhuyze (2006) for Belgium. Matheson (2011) also develops GDP forecasts for a large number of advanced and emerging countries.

It should be noted that applications to inflation forecasting are much rarer, see Forni et al. (2003) or Camba-Mendez and Kapetanios (2005) for the euro area, or de Bandt et al. (2007) for France. Boivin and Ng (2005) also consider series of US prices from a forecasting point of view. It appears to be difficult to improve the accuracy of inflation forecasts by using a large number of variables in comparison with an approach based on a precise selection of variables of interest. In contrast, measurements of underlying inflation have been conducted using this type of approach, see Cristadoro et al. (2005) for the euro area, or Kapetanios (2004) for the United Kingdom.

Various works have endeavoured to identify the contribution of financial variables to macroeconomic forecasting using factor models applied to a database on financial markets activity. For example, Forni et al. (2003) show that in the euro area, financial variables help to forecast inflation but cannot be used to accurately forecast industrial output. Bellégo and Ferrara (2009, 2012) use a factor model to assess the likelihood of a recession in the euro area based on a large set of monthly variables (factor-probit-type model).¹⁷ Specifically, Bellégo and Ferrara (2009) show that by using only financial variables, this approach would have made it possible to anticipate the 2008-09 recession in the euro area in real time as early as 2007Q4.

A review of the literature on the results of factor models in forecasting is found in Eickmeier and Ziegler (2008), who conduct a meta-analysis of the performance of models to forecast GDP and inflation. They conclude that factor models generally improve smaller-scale econometric models, but that methods of combining forecasts constitute a competing alternative approach.¹⁸

6.3. Applications in monetary policy and the international economy

An abundant literature, particularly for the US, analyses the impact of monetary policy shocks on the macroeconomy and the evolution of the transmission mechanism of these shocks. Traditionally the impact of monetary policy shocks is measured using small dimension VAR models (fewer than six variables). Typically, trivariate SVAR models containing the interest rate, output and inflation are used. Beyond this small number of variables, it is difficult to estimate this type of model using standard methods (in a Bayesian context, see de Mol, Giannone and Reichlin, 2008). As noted in Bernanke and Boivin (2003), now monetary policy is conducted by central banks in a data-rich environment. In this context, Bernanke, Boivin and Elias (2005) initiated a category of papers analysing monetary policy using FAVAR models. This type of modelling is chosen to remedy the problem of missing variables generally encountered in traditional VAR modelling. Bernanke, Boivin and Elias (2005), Stock and Watson (2005a), and Favero, Marcellino and Neglia (2005) use FAVAR models to analyse monetary policy in the US and some euro area countries. They all conclude that adding estimated factors from factor models to VAR models allows for a finer-grained analysis of the phenomena in question, particularly in terms of structural shocks. For example, Del Negro and Otrok (2007) estimate a factor common to changes in residential housing prices in various US states, and then introduce it into a FAVAR to assess the extent to which monetary easing helped create a housing bubble (using data up to 2005). They show that the impact of monetary policy shocks is weak in comparison with the scope of the price fluctuations observed through to the end of their sample.

Another portion of the literature addresses the question of whether and how the mechanism for the transmission of shocks has changed over time. Time-varying FAVAR (TV-FAVAR) models provide an extremely flexible modelling that can shed light on this question. The literature appears to agree that this mechanism has changed, although there is no consensus as to how. For example, Eickmeier et al. (2011) show that based on a set of 803 quarterly variables from 1972-2007, the volatility of monetary shocks in the US declined significantly from the early 1980s to the eve of the subprime crisis, and that the negative impact of a shock on US activity and prices declined over this period. The authors also emphasise that the negative impact of a monetary policy shock on inflation expectations and long-term interest rates weakened over time. They reason that these evolutions are

due to changes in monetary policy, and the globalisation of trade and finance. Finally, Eickmeier et al. indicate that the transmission mechanism appears to be the same in both periods of expansion and contraction. Baumeister, Liu and Mumtaz (2010) also show that for the US economy, the reaction of GDP, consumption and investment to a monetary shock declined over the period 1960-2008. We should point out however, that most of these studies do not include the period 2008-09, during which industrialised countries suffered the worst recession since the 1920s.

FAVAR models have also been used to analyse changes in the synchronisation of global business cycles, making it possible to discriminate between different types of shocks. For example, Stock and Watson (2005b) estimate a FAVAR on the GDP of the G7 countries, enabling them to identify common international shocks and domestic effects due to an international shock, and those due to an idiosyncratic shock. They conclude that with the exception of Japan, the reduction in the volatility of cycles in the G7 observed between the mid-1980s and the mid-2000s is primarily due to a reduction in the amplitude of common international shocks (the so-called “Great Moderation”). Kose, Otrok and Whiteman (2003) consider a similar model to show the existence of a global cycle based on a set of 60 countries. They also show that factors specific to the region play only a minor role in the explanation of macroeconomic fluctuations. Bordo and Helbling (2010) explore the historical angle by using annual GDP data from 1880-2008 for 16 industrialised countries, and show a trend toward increasing synchronisation among those countries. The authors show the role of common shocks in this change using a restricted FAVAR model estimated on this database.

There are also several applications of factor models to the measurement of international cycles and their transmission between countries. Mansour (2003), and Helbling and Bayoumi (2003) estimate a global business cycle for the world and the G7 countries, and analyse the contribution of this common cycle to economic changes in each country. Kose, Otrok and Whiteman (2008) use a factor model in a Bayesian context to estimate the common and idiosyncratic components of the G7 countries for a set of economic aggregates. They show that the factor common to these countries explains a larger portion of the variation of these aggregates over the period 1986-2003 than in previous periods, thus demonstrating an increased synchronisation of the business cycles in the G7. Eickmeier (2007) analyses the transmission of US structural shocks to Germany using the approach put forward by Forni et al. (2004). After analysing common economic movements in the euro area, Marcellino, Stock and Watson (2003), and Eickmeier (2005) try to give economic interpretations to the common factors by linking them to the various countries in the zone and/or certain variables, based on correlation measures.

7. Conclusion

In this paper, we reviewed the recent literature on dynamic factor models. There has been an increasing interest in these models on the part of researchers since they can adequately respond to certain problems encountered in practice, particularly inflation of the number of available data. We have presented the models and their most interesting extensions, main estimation methods, and tests of the number of factors. In Section 6, we presented a few recent examples of the application of dynamic factor models to macroeconomic forecasting, the construction of short-term economic indicators and the analysis of monetary policy and the international economy. The success of dynamic factor models means that this review of the literature cannot claim to be exhaustive. Extensions

have recently been developed, particularly to add flexibility by means of non-linearities or mixed frequencies, and numerous applications continue to be published. Moreover, these models are now increasingly weighted against other econometric methods that also make it possible to reduce the scale of the problem. It appears to us, therefore, that the research into dynamic factor models will continue to thrive.

Notes

1. For a discussion of the relationship between restricted DFMs and generalised DFMs, see Giannone et al. (2006) or Forni et al. (2009).
2. Another method of estimating this type of model has been proposed by Sargent and Sims (1977) and Geweke (1977) in the frequency domain, based on a spectral analysis. We come back to this type of frequency domain estimation in the Section 4.
3. Alternatively, the Fisher scoring algorithm is used by Watson and Engle (1983).
4. Ding and Hwang (1999) obtain results on the coherence for PCA estimation of traditional exact factor models when N and T tend to infinity.
5. Jones (2001) and Boivin and Ng (2005) propose weighted PCA estimators by considering the problem of non-linear generalised least squares as follows:

$$\min_{F_1, \dots, F_T, \Lambda} \sum_{t=1}^T (X_t - \Lambda F_t)' \Sigma_{\xi}^{-1} (X_t - \Lambda F_t),$$
with (ξ_t) being IID and of the normal distribution $N(0, \Sigma_{\xi})$. Stock and Watson (2005a) extend the weighted PCA approach by assuming an autoregressive structure of weak order for (ξ_t) .
6. The precise technical condition allowing for a weak correlation of idiosyncratic terms varies from one study to another but in general, this condition limits the contribution of the idiosyncratic covariances to the variance-covariance matrix of X_t when N is large.
7. MCMC methods are numerical methods that create long Markov chains, making it thus possible to obtain samples distributed asymptotically according to a certain distribution.
8. Stock and Watson (1998) develop theoretical results for this methodology.
9. Jungbacker and Koopman (2008) propose new results for the estimation of a dynamic factor model using the maximum likelihood method and a Bayesian method based on Markov chains. Jungbacker et al. (2011) adapt this approach in the context of missing data.
10. The likelihood can be maximised by means of the EM algorithm, which requires the use of the Kalman filter for each iteration.
11. FHLR generalise the dynamic factor model of Sargent and Sims (1977) and Geweke (1977) by raising the hypothesis of the orthogonality of the idiosyncratic factors (see also Forni et al., 2009). Hallin and Liska (2011) recently adapted these models to estimate common factors specific to “blocks” of data, that is, large subpanels of variables.
12. See also Forni and Lippi (2011).
13. Bai and Ng (2002) also propose another class of information criteria for which the variance $V(j, F)$ replaces $\ln(V(j, F))$ in Equations (38), (39) and (40). Bai and Ng (2002, Theorem 2) give the results of the convergence of these criteria when N and T tend to infinity.
14. Bai and Ng (2007) show that $\hat{D}_{1,l}$ and $\hat{D}_{2,l}$ converge toward zero when $l > q$.
15. The criterion IC_2 is used more often in practice.
16. Hallin and Liska (2007) also propose two other penalty functions. Onatski (2009) suggests alternative tests in the context of approximate dynamic factor models. Jacobs and Otter (2008) propose a test based on a canonical correlation procedure to determine simultaneously the number of dynamic factors q and the order of the lag p in a dynamic factor model, but for fixed N and T .
17. A factor-probit model is obtained by, first, estimating the factors from a data set, and then, incorporating them in a standard probit-type model.
18. Forecast combinations generally use a weighted average of forecasts of a single target variable, based on a large number of different models. See Timmerman (2006).

References

- Alessi, L., M. Barigozzi and M. Capasso (2010), "Improved penalization for determining the number of factors in approximate factor models", *Statistics and Probability Letters*, 80, pp. 1806-1813.
- Altissimo, F. et al. (2001), "EuroCOIN: A real time coincident indicator of the euro area business cycle", *Discussion Paper*, No. 3108, CEPR.
- Altissimo, F. et al. (2010), "New EuroCOIN: Tracking economic growth in real time", *The Review of Economics and Statistics*, 92, 4, pp. 1024-1034.
- Amengual, D. and M.W. Watson (2007), "Consistent estimation of the number of dynamic factors in a large N and T panel", *Journal of Business and Economic Statistics*, 25, pp. 91-96.
- Anderson, T.W. (1984), *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.
- Angelini, E. et al. (2011), "Short-term forecasts of Euro area GDP growth", *Econometrics Journal*, 14, 1, pp. C25-C44.
- Artis, M., A. Banerjee and M. Marcellino (2005), "Factor forecasts for the UK", *Journal of Forecasting*, 24, pp. 279-298.
- Aruoba, S.B., F.X. Diebold and C. Scotti (2009), "Real-time measurement of business conditions", *Journal of Business and Economic Statistics*, 27, 4, pp. 417-27.
- Bai, J. (2003), "Inferential theory for factor models of large dimensions", *Econometrica*, 71, 1, pp. 135-171.
- Bai, J. and S. Ng (2002), "Determining the number of factors in approximate factor models", *Econometrica*, 70, pp. 191-221.
- Bai, J. and S. Ng (2007), "Determining the number of primitive shocks in factor models", *Journal of Business and Economic Statistics*, 25, pp. 52-60.
- Bai, J. and S. Ng (2008a), "Forecasting economic time series using targeted predictors", *Journal of Econometrics*, 146, pp. 304-317.
- Bai, J. and S. Ng (2008b), "Large dimensional factor analysis", *Foundations and Trends in Econometrics*, 3, pp. 89-168.
- de Bandt, O. et al. (2007), "Forecasting inflation using economic indicators: The case of France", *Journal of Forecasting*, 26, 1, pp. 1-22.
- Banerjee, A. and M. Marcellino (2006), "Are there any reliable leading indicators for US inflation and GDP growth?", *International Journal of Forecasting*, 22, pp. 137-151.
- Banerjee, A. and M. Marcellino (2009), "Factor-augmented error correction models", in Castle, J. and N. Shepard (eds.), *The Methodology and Practice of Econometrics – A Festschrift for David Hendry*, Oxford: Oxford University Press.
- Banerjee, A., M. Marcellino and I. Masten (2005), "Leading indicators for Euro area inflation and GDP growth", *Oxford Bulletin of Economics and Statistics*, 67, pp. 785-813.
- Barhoumi, K., O. Darné and L. Ferrara (2010), "Are disaggregate data useful for forecasting French GDP with dynamic factor models?", *Journal of Forecasting*, 29, 1-2, pp. 132-144.
- Barhoumi, K., O. Darné and L. Ferrara (2013), "Testing the number of factors: An empirical assessment for forecasting purpose", *Oxford Bulletin of Economics and Statistics*, 75, 1, pp. 64-79.
- Baumeister, C., P. Liu and H. Mumtaz (2010), "Changes in the transmission of monetary policy: Evidence from a time-varying factor-augmented VAR", *Working Paper*, No. 401, Bank of England.
- Bellégo, C. and L. Ferrara (2009), "Forecasting euro area recessions using time-varying binary response models for financial variables", *Working Paper*, No. 259, Banque de France.
- Bellégo, C. and L. Ferrara (2012), "Macro-financial linkages and business cycles: A factor-augmented Probit approach", *Economic Modelling*, 29, pp. 1793-1797.
- Bernanke, B.S. and J. Boivin (2003), "Monetary policy in a data-rich environment", *Journal of Monetary Economics*, 50, pp. 525.
- Bernanke, B.S., J. Boivin and P. Elias (2005), "Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach", *Quarterly Journal of Economics*, 120, pp. 387-422.
- Bessec, M. and C. Doz (2013), "Short-term forecasting of French GDP growth using dynamic factor models", *Journal of Business Cycle Measurement and Analysis*, 2013(2), published in French as

- “Prévision de court terme de la croissance du PIB français à l'aide de modèles à facteurs dynamiques”, *Économie et Prévision*.
- Boivin, J. and S. Ng (2005), “Understanding and comparing factor-based forecasts”, *International Journal of Central Banking*, 1, pp. 117-151.
- Boivin, J. and S. Ng (2006), “Are more data always better for factor analysis?”, *Journal of Econometrics*, 132, pp. 169-194.
- Boivin, J., M.P. Giannoni and I. Mihov (2009), “Sticky prices and monetary policy: Evidence from disaggregated US data”, *American Economic Review*, 99, pp. 350-384.
- Bordo, M. and T. Helbling (2010), “International business cycle synchronization in historical perspective”, *Working Paper 16103*, NBER.
- Breitung, J. and U. Pigorsch (2013), “A canonical correlation approach for selecting the number of dynamic factors”, *Oxford Bulletin of Economics and Statistics*, 75, 1, pp. 23-36.
- Brillinger, D.R. (1981), *Time Series: Data Analysis and Theory*. Holden-Day, San Francisco.
- Camacho, M. and G. Perez-Quiros (2010), “Introducing the EURO-STING: Short term indicator of euro area growth”, *Journal of Applied Econometrics*, 25, pp. 663-694.
- Camacho, M. and G. Perez-Quiros (2011), “Spain-STING: Spain short-term indicator of growth”, *The Manchester School*, 79, pp. 594-616.
- Camba-Méndez, G. and G. Kapetanios (2005), “Forecasting euro area inflation using dynamic factor measures of underlying inflation”, *Journal of Forecasting*, 25, pp. 491-503.
- Campbell, J.Y., A.W. Lo and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton.
- Chamberlain, G. and M. Rothschild (1983), “Arbitrage, factor structure and mean-variance analysis in large asset markets”, *Econometrica*, 51, pp. 1305-1324.
- Charpin, F. (2009), “Estimation précoce de la croissance. De la régression LARS au modèle à facteurs”, *Revue de l'OFCE*, 2009/1, 108, pp. 31-48.
- Chauvet, M. (1998), “An econometric characterization of business cycle dynamics with factor structure and regime switching”, *International Economic Review*, 39, pp. 969-996.
- Chauvet, M. and J. Piger (2008), “A comparison of the real-time performance of business cycle dating methods”, *Journal of Business Economics and Statistics*, 26, 1, pp. 42-49.
- Chevillon, G. (2007), “Direct multi-step estimation and forecasting”, *Journal of Economic Surveys*, 21, pp. 746-785.
- Clavel, L. and C. Minodier (2009), “A monthly indicator for the French business climate”, *Document de Travail No. G2009/02*, INSEE.
- Clements, M. and A.B. Galvao (2008), “Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States”, *Journal of Business and Economic Statistics*, 26, pp. 546-554.
- Connor, G. and R.A. Korajczyk (1986), “Performance measurement with the arbitrage pricing theory: A new framework for analysis”, *Journal of Financial Economics*, 15, pp. 373-394.
- Connor, G. and R.A. Korajczyk (1988), “Risk and return in an equilibrium APT: Application of a new test methodology”, *Journal of Financial Economics*, 21, pp. 255-289.
- Connor, G. and R.A. Korajczyk (1993), “A test for the number of factors in an approximate factor model”, *Journal of Finance*, 48, pp. 1263-91.
- Cornec, M. and T. Deperraz (2006), “Un nouvel indicateur synthétique mensuel résumant le climat des affaires dans les services en France”, *Économie et Statistique*, 395-396, pp. 13-38.
- Cornec, M. (2006), “Analyse factorielle dynamique multifréquence appliquée à la datation de la conjoncture française”, *Économie et Prévision*, 2006/1, 172, pp. 29-43.
- Cristadoro, R. et al. (2005), “A core inflation index for the Euro area”, *Journal of Money, Credit and Banking*, 37, 3, pp. 539-560.
- D'Agostino, A. and D. Giannone (2012), “Comparing alternative predictors based on large-panel factor models”, *Oxford Bulletin of Economics and Statistics*, 74, 2, pp. 306-326.

- Darné, O. and L. Ferrara (2011), "Identification of slowdowns and accelerations for the euro area economy", *Oxford Bulletin of Economics and Statistics*, 73, 3, pp. 335-364.
- Del Negro, M. and C. Otrok (2007), "99 Luftballons: Monetary policy and the house price boom across US states", *Journal of Monetary Economics*, 54, pp. 1962-1985.
- Del Negro, M. and C. Otrok (2008), "Dynamic factor models with time-varying parameters: Measuring changes in international business cycles", *Technical report*, Federal Reserve Bank of New York.
- De Mol, C., D. Giannone and L. Reichlin (2008), "Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?", *Journal of Econometrics*, 146, pp. 318-328.
- Dempster, A., N. Laird and D. Rubin (1977), "Maximum likelihood estimation from incomplete data", *Journal of the Royal Statistical Society*, 14, pp. 1-38.
- Diebold, F.X. and G.D. Rudebusch (1996), Measuring business cycles: A modern perspective, *Review of Economics and Statistics*, 78, 67-77.
- Ding, A.A. and J.T.G. Hwang (1999), "Prediction intervals, factor analysis models, and high-dimensional empirical linear prediction", *Journal of the American Statistical Association*, 94, pp. 446-455.
- Doz, C. and F. Lenglart (1999), "Analyse factorielle dynamique : test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie", *Annales d'Économie et Statistique*, 54, pp. 91-127.
- Doz, C., D. Giannone and L. Reichlin (2011), "A two-step estimator for large approximate dynamic factor models based on Kalman filtering", *Journal of Econometrics*, 164, pp. 188-205.
- Doz, C., D. Giannone and L. Reichlin (2012), "A quasi maximum likelihood approach for large approximate dynamic factor models", *Review of Economics and Statistics*, 94, 4, pp. 1014-1024.
- Eickmeier, S. (2005), "Common stationary and non-stationary factors in the euro area analysed in a large-scale factor model", *Discussion Paper*, No. 2/2005, Bundesbank.
- Eickmeier, S. (2007), "Business cycle transmission from the US to Germany – A structural factor approach", *European Economic Review*, 51, pp. 521-551.
- Eickmeier, S., W. Lemke and M. Marcellino (2011), "Classical time-varying FAVAR models – Estimation, forecasting and structural analysis", *Discussion Paper*, No. 8321, CEPR.
- Eickmeier, S. and C. Ziegler (2008), "How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach", *Journal of Forecasting*, 27, 3, pp. 237-265.
- Eklund, J. and G. Kapetanios (2008), "A review of forecasting techniques for large data sets", *Working Paper*, No. 625, Department of Economics, Queen Mary, University of London.
- Favero, C., M. Marcellino and F. Neglia (2005), "Principal components at work: The empirical analysis of monetary policy with large datasets", *Journal of Applied Econometrics*, 20, pp. 603-620.
- Ferrara, L. and C. Marsilli (2013), "Financial variables as leading indicators of GDP growth: Evidence from a MIDAS approach during the Great Recession", *Applied Economics Letters*, 20, pp. 233-237.
- Forni, M. and M. Lippi (1997), *Aggregation and the Microfoundations of Dynamic Macroeconomics*, Oxford University Press.
- Forni, M. and M. Lippi (2011), "The general dynamic factor model: One-sided representation results", *Journal of Econometrics*, 163, pp. 23-28.
- Forni, M. and L. Reichlin (1998), "Let's get real: A factor-analytic approach to disaggregated business cycle dynamics", *Review of Economic Studies*, 65, pp. 453-473.
- Forni, M., M. Hallin, M. Lippi and L. Reichlin (2000), "The generalized dynamic-factor model: Identification and estimation", *Review of Economics and Statistics*, 82, 540-554.
- Forni, M. et al. (2003), "Do financial variables help forecasting inflation and real activity in the euro area?", *Journal of Monetary Economics*, 50, pp. 1243-1255.
- Forni, M. et al. (2004), "The generalized factor model: Consistency and rates", *Journal of Econometrics*, 119, pp. 231-255.
- Forni, M. et al. (2005), "The generalized dynamic factor model: One-sided estimation and forecasting", *Journal of the American Statistical Association*, 100, pp. 830-840.
- Forni, M. et al. (2009), "Opening the black box: Structural factor models with large cross-sections", *Econometric Theory*, 25, pp. 1319-1347.

- Frale, C. et al. (2010), "Survey data as coincident or leading indicators", *Journal of Forecasting*, 29, pp. 109-131.
- Frale, C. et al. (2011), "EUROMIND: A monthly indicator of the euro area economic conditions", *Journal of the Royal Statistical Society: Series A*, 174, 2, pp. 439-470.
- Geweke, J. (1977), "The dynamic factor analysis of economic time series", in Aigner, D.J. and A.S. Goldberger (eds.), *Latent Variables in Socio Economic Models*, North Holland, Amsterdam.
- Ghysels, E., A. Sinko and R.I. Valkanov (2007), "MIDAS regressions: Further results and new directions", *Econometric Reviews*, 26, pp. 53-90.
- Giannone D., L. Reichlin and L. Sala (2004), "Monetary policy in real time", in Gertler M. and K. Rogoff (eds.), *NBER Macroeconomics Annual*, MIT Press.
- Giannone, D., L. Reichlin and L. Sala (2006), "VARs factor models and the empirical validation of equilibrium of business cycle models", *Journal of Econometrics*, 132, pp. 257-279.
- Giannone, D., L. Reichlin and D. Small (2008), "Nowcasting: The real-time informational content of macroeconomic data", *Journal of Monetary Economics*, 55, pp. 665-676.
- Gorman, W.M. (1981), "Some Engel curves", in A. Deaton (ed.) *Essays in the Theory and Measurement of Consumer Behaviour in Honor of Sir Richard Stone*, Cambridge University Press, New York.
- Hallin, M. and R. Liska (2007), "Determining the number of factors in the general dynamic factor model", *Journal of the American Statistical Association*, 102, pp. 603-617.
- Hallin, M. and R. Liska (2011), "Dynamic factors in the presence of blocks", *Journal of Econometrics*, 163, pp. 29-41.
- Hamilton, J.D. (1989), "A new approach to the economic analysis of nonstationary time series and the business cycle", *Econometrica*, 57, 2, pp. 357-384.
- Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press, Princeton.
- Harding, M.C. (2009), *Structural estimation of high-dimensional factor models*, Mimeo, Department of Economics, Stanford University.
- Helbling, T. and T. Bayoumi (2003), "Are they all in the same boat? The 2000-2001 growth slowdown and the G7-business cycle linkages", *Working Paper*, WP/03/46, IMF.
- Jacobs, J.P.A.M. and P.W. Otter (2008), "Determining the number of factors and lag order in dynamic factor models: A minimum entropy approach", *Econometric Reviews*, 27, pp. 385-397.
- Jones, C.S. (2001), "Extracting factors from heteroskedastic asset returns", *Journal of Financial Economics*, 62, pp. 293-325.
- Jungbacker, B. and S.J. Koopman (2008), "Likelihood-based analysis for dynamic factor models", *Discussion Paper*, No. 2008-0007/4, Tinbergen Institute, Amsterdam University.
- Jungbacker, B., S.J. Koopman and M. van der Wel (2011), "Maximum likelihood estimation for dynamic factor models with missing data", *Journal of Economic Dynamics and Control*, 35, pp. 1358-1368.
- Kapetanios, G. (2004), "A note on modelling core inflation for the UK using a new dynamic factor estimation method and a large disaggregated price index dataset", *Economics Letters*, 85, pp. 63-69.
- Kapetanios, G. (2010), "An alternative method for determining the number of factors in factor models with large datasets", *Journal of Business and Economic Statistics*, 28, pp. 397-409.
- Kapetanios, G. and M. Marcellino (2004), "A comparison of estimation methods for dynamic factor models of large dimension", *Queen Mary Working Paper*, No. 489, London.
- Kim, C.-J. and C.R. Nelson (1998), "Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching", *Review of Economics and Statistics*, 80, pp. 188-201.
- Kim, C.-J. and J.S. Yoo (1995), "New index of coincident indicators: A multivariate Markov Switching factor model approach", *Journal of Monetary Economics*, 36, pp. 607-630.
- Kose, M.A., C. Otrok and C. Whiteman (2003), "International business cycles: World, region and country-specific factors", *American Economic Review*, 93, 4, pp. 1216-1239.
- Kose, M.A., C. Otrok and C. Whiteman (2008), "Understanding the evolution of world business cycles", *Journal of International Economics*, 75, pp. 110-130.
- Krolzig, H.M. and D. Hendry (2001), "Computer automation of general-to-specific model selection procedures", *Journal of Economic Dynamics and Control*, 25, pp. 831-866.

- Lawley, D.N. and A.E. Maxwell (1971), *Factor Analysis in a Statistical Method*, Butterworth, London.
- Lewbel, A. (1991), "The rank of demand systems: Theory and nonparametric estimation", *Econometrica*, 59, pp. 711-730.
- Lopes, H.F. and M. West (2004), "Bayesian model assessment in factor analysis", *Statistica Sinica*, 14, pp. 41-67.
- Mansour, M.J. (2003), "Do national business cycles have an international origin?", *Empirical Economics*, 28, pp. 223-247.
- Marcellino, M., J. Stock and M. Watson (2003), "Macroeconomic forecasting in the euro area: Country specific versus euro wide information", *European Economic Review*, 47, pp. 1-18.
- Marcellino, M. and C. Schumacher (2010), "Factor-MIDAS for nowcasting and forecasting with ragged-edge data: A model comparison for German GDP", *Oxford Bulletin of Economics and Statistics*, 72, pp. 518-550.
- Mariano, R. and Y. Murasawa (2003), "A new coincident index of business cycles based on monthly and quarterly series", *Journal of Applied Econometrics*, 18, pp. 427-443.
- Mariano, R. and Y. Murasawa (2010), "A coincident index, common factors and monthly real GDP", *Oxford Bulletin of Economics and Statistics*, 72, 1, pp. 27-46.
- Matheson, T. (2011), "New indicators for tracking growth in real time", *Working Paper*, No. WP/11/43, International Monetary Fund.
- Motta, G., C. Hafner and R. von Sachs (2011), "Locally stationary factor models: Identification and nonparametric estimation", *Econometric Theory*, 27, 6, pp. 1279-1319.
- Mumtaz, H. and P. Surico (2009), "The transmission of international shocks: A factor-augmented VAR approach", *Journal of Money, Credit and Banking*, 41, pp. 71-100.
- Newbold, P. and A. Harvey (2002), "Forecast combination and encompassing", in Clements M. and D. Hendry (eds.), *A Companion to Economic Forecasting*, Basil Blackwell, Oxford.
- Nguiffo-Boyom, M. (2006), "Un indicateur de retournement conjoncturel pour la France: une application du modèle à facteur avec changements de régimes", *Économie et Prévision*, 172, pp. 101-114.
- Onatski, A. (2009), "Testing hypotheses about the number of factors in large factor models", *Econometrica*, 77, pp. 1447-1479.
- Onatski, A. (2010), "Determining the number of factors from empirical distribution of eigenvalues", *The Review of Economics and Statistics*, 92, 4, pp. 1004-1016.
- Peña, D. and P. Poncela (2006a), "Forecasting with non-stationary dynamic factor models", *Journal of Econometrics*, 119, pp. 291-321.
- Peña, D. and P. Poncela (2006b), "Non-stationary dynamic factor analysis", *Journal of Statistical Planning and Inference*, 136, pp. 1237-1257.
- Ross, S. (1976), "The arbitrage theory of asset pricing", *Journal of Economic Theory*, 13, pp. 341-360.
- Ruenstler, G. et al. (2009), "Short-term forecasting of GDP using large monthly data sets: A pseudo real-time forecast evaluation exercise", *Journal of Forecasting*, 28, 7, pp. 595-611.
- Sargent, T. and C. Sims (1977), "Business cycle modelling without pretending to have too much a priori economic theory", in C. Sims (ed.), *New Methods in Business Cycle Research*, Federal Reserve Bank of Minneapolis, Minneapolis.
- Schumacher, C. (2007), "Forecasting German GDP using alternative factor models based on large datasets", *Journal of Forecasting*, 26, pp. 271-302.
- Schumacher, C. (2010), "Factor forecasting using international targeted predictors: The case of German GDP", *Economics Letters*, 107, 2, pp. 95-98.
- Schumacher, C. and J. Breitung (2008), "Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data", *International Journal of Forecasting*, 24, pp. 368-398.
- Shumway, R.H. and D.S. Stoffer (1982), "An approach to time series smoothing and forecasting using the EM algorithm", *Journal of Time Series Analysis*, 3, pp. 253-264.
- Steiger, J.H. (1979), "Factor indeterminacy in the 1930s and the 1970s: Some interesting parallels", *Psychometrika*, 40, pp. 157-167.

- Stock, J. and M. Watson (1989), "New Indexes of Coincident and Leading Economic Indicators", NBER *Macroeconomics Annual*, NBER.
- Stock, J. and M. Watson (1998), "Diffusion indexes", *Working Paper*, No. 6702, NBER.
- Stock, J. and M. Watson (1999), "Forecasting inflation", *Journal of Monetary Economics*, 44, pp. 293-335.
- Stock, J. and M. Watson (2002), "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics*, 20, pp. 147-162.
- Stock, J. and M. Watson (2005a), "Implications of dynamic factor models for VAR analysis", *Working Paper*, No. 11467, NBER.
- Stock, J. and M. Watson (2005b), "Understanding changes in international business cycle dynamics", *Journal of the European Economic Association*, 3, 5, pp. 968-1006.
- Stock, J. and M. Watson (2006), "Forecasting with many predictors", in G. Elliott, C.W.J. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Vol. 1, North-Holland, Amsterdam.
- Timmermann, A. (2006), "Forecast combinations", in Elliott, G., C.W.J. Granger and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Vol. 1, North-Holland, Amsterdam, pp. 135-196.
- Van Nieuwenhuyze, C. (2006), "A generalized dynamic model for the Belgian Economy – useful business cycle indicators and GDP growth forecasts", *Working Paper*, No. 80, National Bank of Belgium.
- Watson, M.W. and R.F. Engle (1983), "Alternative algorithms for estimation of dynamic factor, MIMIC, and time varying coefficient regression models", *Journal of Econometrics*, 23, pp. 385-400.
- White, H. (1982), Maximum likelihood estimation of misspecified models. *Econometrica*, 50, pp. 1-25.
- Yao, Q. and J. Pan (2008), "Modelling multiple time series via common factors", *Biometrika*, 95, pp. 365-379.

Constructing a conditional GDP fan chart with an application to French business survey data

by
Matthieu Cornec*

Interval confidence and density forecasts, notably in the form of “fan charts”, are useful tools to describe the uncertainty inherent to any point forecast. However, the existing techniques suffer from several drawbacks. We propose a new method to represent uncertainty in real-time that is conditional upon the economic outlook, non-parametric and reproducible. Moreover, we build a Forecasting Risk Index associated with our fan chart to measure the intrinsic difficulty of the forecasting exercise. Using balances of opinion of different business surveys carried out by the French statistical institute INSEE, our GDP fan chart efficiently captures the growth stall during the crisis on a real-time basis. Our Forecasting Risk Index has increased substantially in this period of turbulence, showing signs of growing uncertainty.

Keywords: Density forecast, quantile regressions, business tendency surveys, fan charts

JEL classification: E32, E37, E66, C22

* Insee, Business Tendency Surveys Division, Economic Studies and National Accounts Directorate, Department of Short-Term Analysis, Timbre G120, 15 Boulevard Gabriel Péri, 92245 Malakoff Cedex, France.

The author thanks the anonymous referees for the careful reading and the valuable comments and suggestions to improve this paper. Any remaining errors or omissions are the responsibility of the author.

1. Introduction

Usually, financial institutions publish economic point forecasts. However, the forecasts are not free of uncertainty; assuming that forecasts are not biased, they are considered as the expected figure given the available information. Indeed, many shocks can affect the forecast for oil prices, exchange rates, interest rates and other variables. Another reason is that the behaviours of economic agents can only be estimated imprecisely over the past (even when they do not change). The potential scenarios are therefore numerous: forecasters thus condense their forecast into a single, “baseline” scenario.

Ultimately, readers may lose sight of the uncertainty inherent to this type of exercise. Among economic forecasters, it has therefore become common practice to provide point projection with a density forecast. The longest running series of macroeconomic density forecasts dates back to 1968, when the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER) initiated a survey of forecasters. For a detailed historical review, see Tay and Wallis (2003). In particular, to dispel their risk, the Central Bank of England (CBE) (see Britton et al., 1998) and the French National Institute for Statistics and Economic Studies (INSEE) display a “fan chart” to provide a concise illustration of the uncertainty affecting point forecasts. This realistic view recognises that the future evolution of the economic outlook cannot be predicted with absolute certainty. Confidence intervals and density forecasts have emerged as useful tools to describe, in probability terms, the uncertainty inherent to any point forecast.

Both methodologies suffer from some drawbacks. First, the INSEE fan chart is unconditional; which means that if the methodology is strictly applied, the magnitude of the uncertainty displayed is the same, whatever the economic outlook is. However, it is a common belief among practitioners that the forecasting exercise greatly depends on the state of the economy, especially during crisis. Secondly, the CBE fan chart is conditional but this conditionality comes from subjective assessments by the members of the Monetary Policy Committee, and is therefore not reproducible. Lastly, another limitation is the parametric shape of the distribution, which is usually assumed to be exponential, that is, without fat tails. In this paper, we tackle these issues by providing a reproducible methodology to build non-parametric conditional density forecasts.

This paper is organised as follows: in Section 2 we review the main tools used to describe uncertainty, namely confidence intervals and density; we describe in detail the methodologies used by both INSEE and the CBE to derive density forecasts. In Section 3, we highlight the drawbacks of those approaches. We give a brief description of French business surveys in Section 4. In Section 5, we introduce our new statistical methodology to derive conditional density forecasts and to construct our Forecasting Risk Index. We then apply the proposed methodology to the French economy in Section 6 using business surveys with a focus on the recent financial crisis.

2. Uncertainty description

2.1. Confidence intervals

Confidence intervals may be the first simple method to describe point forecast uncertainty (for a review, see Tay and Wallis, 2002). The principle is the following: the economic forecaster provides an interval together with his point forecast; the future observed value is then supposed to lie within this interval with a specified probability. For example, if the forecaster provides 95% confidence intervals, the observed value is supposed to be within these intervals around 95% of the time. It is also common to provide confidence intervals at different probability levels, that is 20%, 50% and 99% confidence intervals.

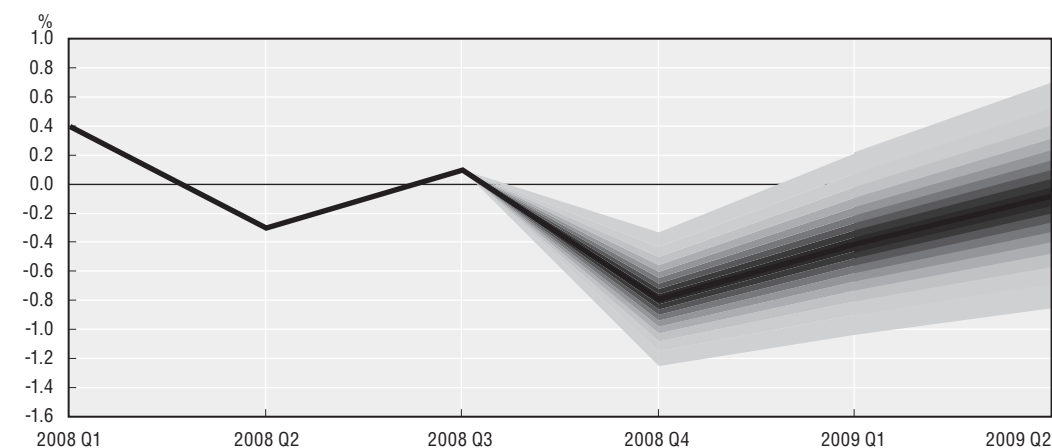
In the case of confidence intervals, the width of the error band measures the level of uncertainty. For example, we usually expect uncertainty to grow as the forecasting horizon increases. This is displayed by an increasing width of error bands for a specified level of probability (Figure 1). However, if the confidence intervals simplicity makes them very attractive, they do not completely describe uncertainty.

2.2. Density forecasts

A common way to describe uncertainty is to consider the future evolution of the economy (that is, GDP growth rate) y_{t+1} , as the realisation of a continuous random variable Y_{t+1} , where t denotes the current quarter. The uncertainty is then fully characterised by the random variable density $f_{Y_{t+1}}(\cdot)$. At this stage, note that the density is not supposed to be normal. An appealing feature of density forecasts is that they can also be used to easily derive confidence intervals based on the appropriate distribution quantiles. For example, to get a 90% confidence interval I_{90} for the next unknown outcome y_{t+1} , it is sufficient to find real numbers a and b such that $\int_a^b f_{Y_{t+1}}(y)dy = 0.9$ and to define I_{90} as $I_{90} \equiv [a, b]$.

In the case of density, the level of uncertainty is related to the sharpness of the distribution. A peaked density means a small uncertainty. On the contrary, a large uncertainty is translated into a loose distribution. Many indicators have been proposed to

Figure 1. INSEE GDP fan chart



Source: Insee (2008a), *Note de Conjoncture*.

quantify the level of uncertainty, among them variance and entropy. For examples of applications of these statistical elements, we refer to Osterholm (2009) and Robertson, Tallman and Whiteman (2005), who employ relative entropy for constructing forecasting bounds for conditional forecasts. Waggoner and Zha (1999) developed Bayesian methods to compute conditional error distributions in vector autoregressive models.

A central question is then: how to construct such density forecasts? Since 1996, the CBE has published a density forecast of inflation in its quarterly *Inflation Report*, the so called “fan chart”. More recently, INSEE has also published a fan chart of its GDP prediction in the *Note de Conjoncture* (see Figure 1). We will first review the two methodologies used respectively by INSEE and the CBE, and then highlight their drawbacks.

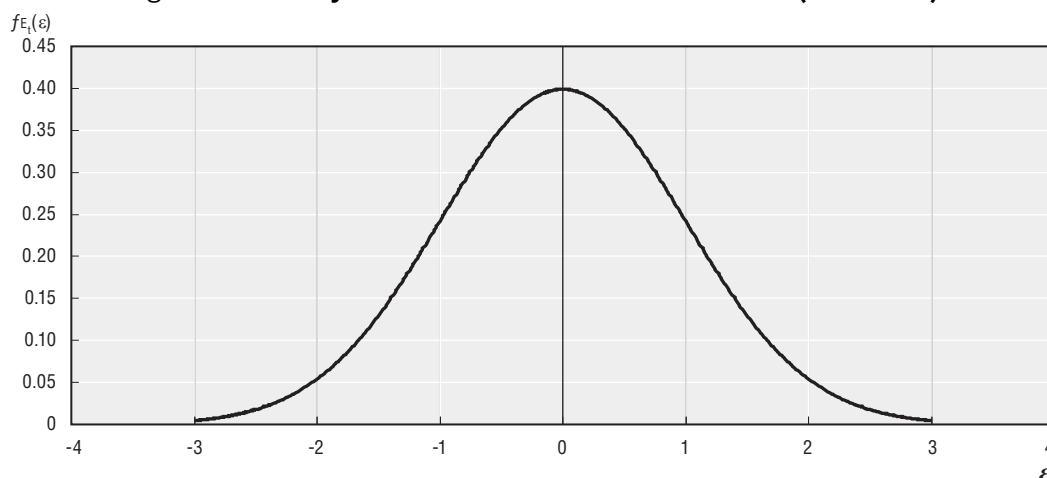
2.3. INSEE methodology

For a complete description, we refer to INSEE (2008b). Forecasting errors are computed by comparing GDP flash estimates published in the national accounts y_t and the GDP growth forecast $\hat{y}_t$: $\varepsilon_t \equiv y_t - \hat{y}_t$. These errors are supposed to be drawn from a zero-mean normal distribution:

$$f_{\varepsilon_t}(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right).$$

Thus, INSEE perceives the possible GDP outcomes symmetrically dispersed around the central most probable value. This is the famous bell-shaped curve (Figure 2 for the density of a standard normal distribution).

Figure 2. **Density of a standard normal distribution (bell curve)**



To construct the fan chart, INSEE (2008b) estimated the standard deviation σ from the time series of forecasting errors in *Note de Conjoncture* reports published by INSEE between 1989 and 2006:

$$\hat{\sigma} \equiv \sqrt{\frac{1}{T} \sum_{t=1}^T \varepsilon_t^2}.$$

which gives a standard deviation of 0.28. It is now possible to derive the INSEE fan chart, which is simply:

$$\begin{aligned}\hat{f}_{Y_t}(y) &= \hat{y}_t + \hat{f}_{E_t}(\varepsilon) \\ &= \hat{y}_t + \frac{1}{\sqrt{2\pi\hat{\sigma}}} \exp\left(-\frac{(y - \hat{y}_t)^2}{2\hat{\sigma}^2}\right)\end{aligned}$$

$$\forall y, \varepsilon = y - \hat{y}_t$$

The rationale for this is as follows: “Assuming no change in the volatility of growth figures and the methodology used by INSEE forecasters during the period, the distribution of forecasting errors calculated from past exercises is a reliable indicator of the distribution of future errors and hence of the uncertainty surrounding a given forecast”, INSEE (2008b).

2.4. CBE methodology

The fan chart constructed by the CBE is based on past forecasting errors. However, in some cases they may diverge from the historical record depending on the risk analysis by members of their monetary policy committees.¹ For instance, the recent financial-market turmoil has arguably made forecasts more uncertain than usual. In some circumstances, we may consider that the risks surrounding the likeliest scenario – the one set out in the report – are asymmetrical. This assessment of risk asymmetry is important for central banks, as events that diverge very significantly from the baseline forecast (such as a decline in consumer prices) can have disastrous consequences. The CBE allows subjectivity to deviate from the full-sample variability of the data/model.

Unlike that of INSEE, the CBE fan chart somehow measures the CBE subjective uncertainty. The aim of the fan chart has been to convey to the reader a more accurate representation of the Bank’s subjective assessment of medium-term inflationary pressures. It is therefore a forward-looking view of the risks to the forecast, not an extrapolation of past uncertainty. For more insights about this debate, we refer to de Finetti (1975) and the references therein.

3. Drawbacks of INSEE and CBE methodologies

3.1. Conditional forecasting versus unconditional forecasting

First, the INSEE fan chart is *unconditional*, which means that the magnitude of the displayed uncertainty is the same whatever the economic outlook is; remembering that the standard deviation is estimated on the past sample. On a long-term basis it is supposed to be correct on average. During a recession, however, the usual unconditional central confidence interval can be significantly wrong. Let us take a simple example by comparing forecasting with flying a plane: when crossing a turbulent area, flying gets more difficult. In the same way, during a crisis period, the intrinsic uncertainty of the forecasting exercise increases.

Instead of unconditional forecasting, conditional forecasting, which can change from one period to another, should be favoured to describe uncertainty. Indeed, conditional forecasting adds more information to the forecast than unconditional forecasting. We illustrate the value of conditional forecasting versus unconditional forecasting through the following simple model.

Let Z_t be a random variable that describes the state of the economy at time t and that can only take two values: A for acceleration with probability p ; D for deceleration with probability

$1 - p$. Unlike the hidden Markov models introduced in the seminal work of Hamilton (1989), Z_t can be observed. In Section 5, the role of Z_t will be played by business surveys.

Let Y_t be a random variable describing the GDP growth rate.

To make it simple, we assume the conditional distributions of Y_t to be such that $E(Y_t|Z_t) = \mu_{Z_t}$, $Var(Y_t|Z_t) = \sigma_{Z_t}^2$ with $\mu_D < \mu_A$ (the growth rate is smaller during recession) and $\sigma_A^2 < \sigma_D^2$ (but the volatility is higher). We consider the variance as an indicator to quantify the level of uncertainty.

Then the description of uncertainty by an unconditional forecaster is simply $E(Var(Y_t|Z_t))$ since he does not take into account information from Z_t . Note that $E(Var(Y_t|Z_t))$ is equal to $p\sigma_D^2 + (1-p)\sigma_A^2$.

Conversely, for a conditional forecaster, uncertainty as measured by $Var(Y_t|Z_t)$ would be equal to σ_D^2 during recession and to σ_A^2 during acceleration.

We can deduce that, on a long-term average, the unconditional description is correct. However, at each quarter t , knowing the information Z_t , it is either too small or too big. Indeed, since $\sigma_A^2 < p\sigma_D^2 + (1-p)\sigma_A^2 < \sigma_D^2$, the unconditional description of uncertainty is either too loose (during an acceleration period) or too sharp (during a crisis). The unconditional forecast error does not take into account all the information available.

Eventually, since the CBE methodology allows the injection of subjectivity from one period to another, their fan chart is thus conditional. However, their method is not reproducible as it is a measure of the CBE subjectivity.

3.2. Evaluation of INSEE density forecasts

Another central question is the evaluation of density forecasts. The literature on statistical evaluation of real-time density forecasts is quite new. For a survey, we refer to Tay and Wallis (2002) or Wallis (2003). However, the key feature, namely the probability integral form, can be found in Rosenblatt (1952).

The principle is the following:

1. Compute the probability integral transform of the realised variables with respect to the forecast densities:

$$h_t \equiv \hat{F}_t(y_t) = \int_{-\infty}^{y_t} \hat{f}_{Y_t}(y) dy$$

where $\hat{f}_{Y_t}(\cdot)$ is the forecast distribution function. If $\hat{f}_{Y_t}(\cdot)$ is correct (that is, $\hat{f}_{Y_t}(\cdot) = f_{Y_t}(\cdot)$), then h_t should be distributed as a sequence of independent identically distributed uniform variables on $[0,1]$.

Indeed, we get:

$$\begin{aligned} P_t(h_t \leq h) &= P_t(\hat{F}_t(y_t) \leq h) \\ &= P_t(y_t \leq \hat{F}_t^{-1}(h)) \\ &= F_t(\hat{F}_t^{-1}(h)) \end{aligned}$$

and thus $P_t(h_t \leq h) = h$ if the forecast is correct (that is, $F_t = \hat{F}_t$). In other words, we obtain the cumulative distribution function of a uniform variable on $[0,1]$. When the forecasts comprise sequences of conditional densities, provided that the forecasts are based on an

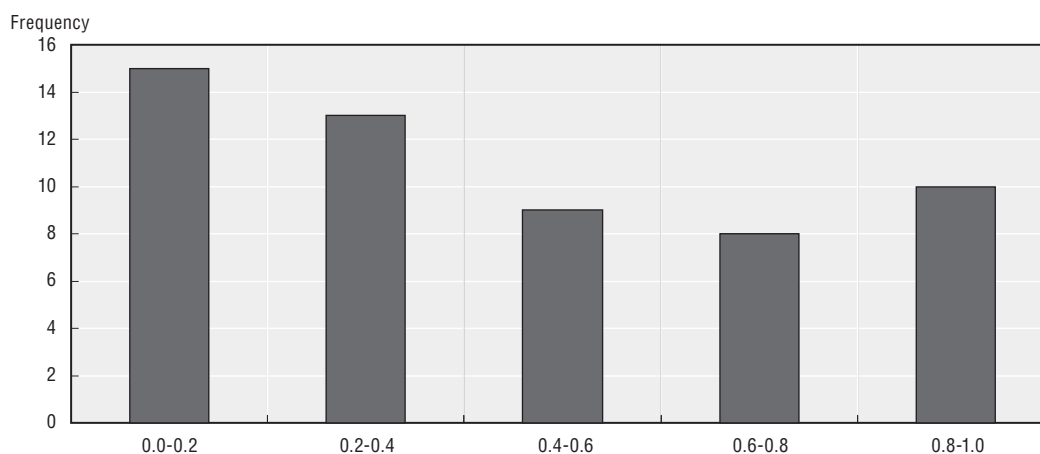
information set that contains the past history of the forecast variable, the independence properties hold.

2. Draw the histogram of h_t , which should be a straight line if our forecast is correct. Deviations from independent and identically distributed (i.i.d.) uniform distributions will indicate that the forecasts have failed to capture some aspects of the underlying data generating process. For example, a concave profile indicates that the probabilities of extreme risks are overestimated. On the contrary, an underestimation of extreme risks leads to a convex profile.

The histogram of h_t is defined as the rank histogram. Other more sophisticated tests might be considered (see Berkowitz, 2001; Wallis, 2003 and references therein).

If we apply this methodology to the INSEE fan chart exercise from 1996Q1 to 2010Q1, we can draw the rank histogram displayed in Figure 3. Instead of a straight line, it exhibits a convex profile indicating that the probabilities of right extreme risks are underestimated; in particular, the left tail risks (stall risk). The reason is that the parametric shape of a normal distribution might not be best suited to modelling the error distribution. Instead of a new parametric distribution, we will propose a new semi-parametric methodology.

Figure 3. Rank histogram of INSEE density forecasts

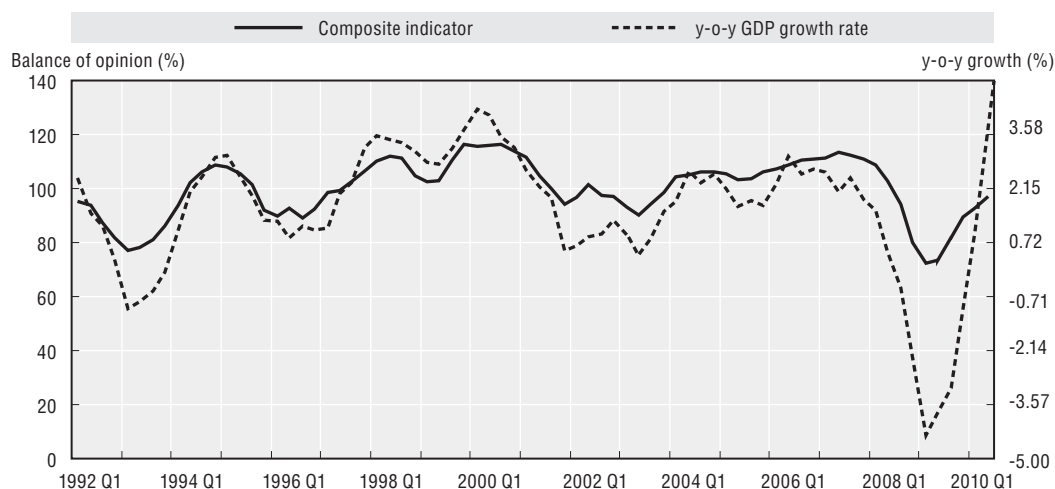


Source: INSEE, author's calculations.

4. Business surveys

Recall that Z_t describes the state of the economy in our toy model. The role of Z_t will be played by business surveys (in practice, this is a multidimensional variable). Indeed, business surveys are a useful source of information when forecasting, for they present three types of advantages: 1) they provide reliable information coming directly from the economic decision makers; 2) they are rapidly available (about one month after the questionnaires are sent), on a monthly, bimonthly or quarterly basis; and, 3) they are subject to small revisions (each publication presents a generally negligible revision, only on the preceding point). The disseminated statistics compiled from these surveys are usually balances of opinions.² Since there are a large number of available survey variables, composite indexes have been developed over the years to provide suitable summaries by extracting the common trend, and suppressing the undesirable “noise” of numerous data. In particular, the French composite indicator (Figure 4) gives an assessment of the global climate of the whole French economy.

Figure 4. French composite indicator and year-on-year GDP growth rate



Sources: INSEE, Quarterly National Accounts and Business Tendency Surveys.

The French composite indicator is denoted by $(i_t)_t$.³ Recall that our goal is to forecast the first release of the French quarterly GDP growth rate y_t . The first release will be published only 45 days after the end of the current quarter. Typically, economists also forecast the next quarters. For the sake of simplicity, we will restrict ourselves to the forecast of the current quarter. Finally, we define the vector of past observations by $\underline{y}_{t-1} = (y_1, \dots, y_{t-1})'$ (the prime denoting the transposed). Our quarterly historical data from French GDP first releases starts from 1988Q1.

5. Methodology

In this section, we first define our goal, which is to forecast the next conditional density of gross domestic product. In a general framework, we then introduce the main tool, namely the quantile regression. We specify our model that will be used to derive point forecasts, confidence intervals and the forecasting risk index.

To model uncertainty, we consider $(z_t)_t$ and $(y_t)_t$ as the outcome of random variables $(Z_t)_t$ and $(Y_t)_t$ such that the process $(Z_t', Y_t')_{t \in \mathbb{Z}}$ is jointly continuous and stationary ergodic (Z_t is a column vector).

Adapting the sequential framework in Biau and Patra (2011), at each quarter $t = 1 \dots T$, we observe business survey data z_t but the first release of the quarter-on-quarter GDP growth rate y_t is unknown. In a sequential version of the density prediction problem, the forecaster is asked to guess the next conditional density of Y_t of a sequence of random variables $Y_1, \dots, Y_{t-1}$ with knowledge of the past observations $\underline{y}_{t-1} = (y_1, \dots, y_{t-1})'$ and $\underline{z}_t = (z_1, \dots, z_t)'$. In other words, the observations $y_1, y_2, \dots$ and $z_1, z_2, \dots$ are revealed one at a time, beginning with $(y_0, z_1'), (y_1, z_2'), \dots$ and the forecaster must guess $f_{Y_t}(\cdot | \underline{y}_{t-1}, \underline{z}_t)$.

Quantile regression techniques provide a suitable semi-parametric tool to achieve a proper density forecast or confidence intervals. It is an applied technique that allows direct estimation of the range of forecast outcomes. The forecast bands are a direct outcome rather than the estimation of an empirical distribution of forecast errors or as an estimate of the standard error. This is why this is a very appealing technique. We introduce here the main outlines of the quantile regression methodology. The reference book on the topic is

Koenker (2005). Other recent statistical literature includes Kocherginsky et al. (2005), Wei and He (2006), and Chernozhukov et al. (2010).

5.1. Introduction to quantile regression

Typically, the quarter-on-quarter GDP growth Y_t is a function of predictor variables Z_t , so that $Y_t = g(Z_t) + \Xi_t$, with Ξ_t a white noise. Most economic forecasting applications focus on estimating rates of change in the mean response variable distribution as a function of a set of predictor variables; in other words, the function is defined for the expected value of Y_t conditional on Z_t : $E(Y_t|Z_t)$. By doing so, most regression analyses miss other parts of the distribution of the response variable. This is especially problematic for regression models with heterogeneous variances, which are common in economics. Quantile regression is a statistical technique intended to estimate, and conduct inference on, conditional quantile functions. It was developed in the 1970s by econometricians (Koenker and Bassett, 1978) as an extension of the linear model. Quantiles seem inseparably related to ordering the sample observations that are usually used to define them. Therefore, it is surprising that quantiles can be defined through an optimisation problem.

To start, note that many probability quantities can be characterised by a minimisation problem. For example, the mean of variables with finite second moment satisfies:

$$E(Y) = \arg \min_{m \in \mathbb{R}} E(Y - m)^2 \quad (1)$$

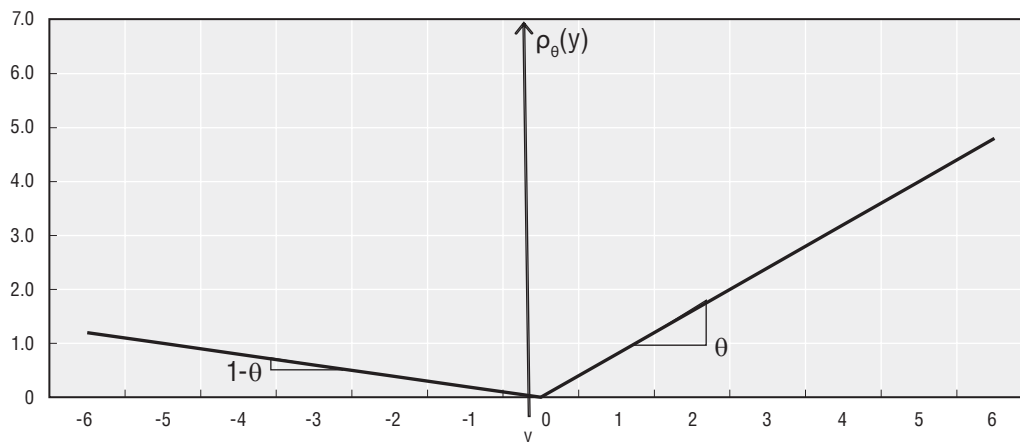
since,

$$\frac{\partial}{\partial m} E(Y - m)^2 = 0 \Rightarrow E(Y - m) = 0$$

In the same way, the median corresponds to $med(Y) = \arg \min_{m \in \mathbb{R}} E|Y - m|$. Thus, the median can be obtained by minimising a symmetrical loss, namely the absolute value. It appears it is also possible to obtain any quantile, as formally defined by $Q_Y(\theta) \equiv \inf \{t \in \mathbb{R} : F_Y(t) \geq \theta\}$, by minimising an asymmetrical loss, called the pinball loss (Figure 5). To do this, it is sufficient to differently weight positive and negative residuals $\rho_\theta(y) \equiv \theta y^+ + (1 - \theta)y^-$ with $y^+ \equiv \max(0, y)$, $y^- \equiv \max(0, -y)$. Note that, in the case where $\theta \equiv 1/2$, we retrieve the absolute value. We thus obtain:

$$Q_Y(\theta) = \arg \min_{m \in \mathbb{R}} E\rho_\theta(Y - m) \quad (2)$$

Figure 5. Pinball loss



Having defined the unconditional quantile, it is not difficult to define the conditional quantile in a similar way. In Equation 1 if we replace the scalar m by a function $m(\cdot)$ and we solve all possible functions (with finite second moment), we get: $E(Y|Z) = \arg \min_{m(\cdot)} E(Y - m(Z))^2$. In the same way, if we replace in Equation 2 the scalar m by a function $m(\cdot)$, we obtain

$$Q_Y(\theta|Z) = \arg \min_{m(\cdot)} E \rho_\theta(Y - m(Z)) \quad (3)$$

For formal proof in the general case, we refer to Biau and Patra (2011).

Lastly, we still need to estimate those quantities based on the knowledge of past observations. To do this, it is sufficient to replace the expected pinball loss by its empirical counterpart when the forecasting exercise is carried out over the period from $t = 1$ to T . The unconditional quantile estimate is thus given by:

$$\hat{Q}_Y(\theta) \equiv \arg \min_{m \in \mathbb{R}} \sum_{t=1}^T \rho_\theta(y_t - m)$$

and the conditional quantile estimate by:

$$\hat{Q}_Y(\theta|z) = \arg \min_{\beta \in \mathbb{R}^p} \sum_{t=1}^T \rho_\theta(y_t - m(z_t, \beta))$$

with $m(z_t, \beta)$ a parametric function dependant on the parameter β of the dimension p . Koenker and Bassett (1978) considered $m(z)$ of a linear form $z'\beta$ and the coefficients are estimated by minimising:

$$\sum_{t=1}^T \rho_\theta(y_t - z_t' \beta) \quad (4)$$

on past sample observations. Let $\hat{\beta}_\theta$ denote the result of this minimisation. This estimation is called linear quantile regression. Their estimated quantile curve is then defined by $\hat{Q}_Y(\theta|z) \equiv z' \hat{\beta}_\theta$. To justify linear quantile regression, assume the data to be generated by the following model with heteroscedasticity:

$$Y_t = \mu_t(\underline{Z}_t) + \sigma_t(\underline{Z}_t) \varepsilon_t \quad (5)$$

with $\mu_t(\underline{Z}_t) \equiv E(Y_t | \underline{Z}_t)$ and $\sigma_t(\underline{Z}_t) \equiv V(Y_t | \underline{Z}_t)$, and ε_t an error term independent of $\underline{Z}_t$. Based on the aforementioned Equation 4, the conditional quantile of Y_t given $\underline{Z}_t$ has the form:

$$Q_Y(\theta | \underline{Z}_t) = \mu_t(\underline{Z}_t) + \sigma_t(\underline{Z}_t) Q_{\varepsilon_t}(\theta)$$

Koenker and Bassett (1982) showed that if $\mu_t(\underline{Z}_t)$ and $\sigma_t(\underline{Z}_t)$ are linear functions of $\underline{Z}_t$, then quantile regression estimates are asymptotically consistent. The appeal of quantile regression is that past observations of the quantiles $Q_Y(\theta | \underline{Z}_{t-i})_{i>0}$ are not required. In the following subsection, we will specify Equation 5 by precisely defining $\underline{Z}_t$. As mentioned in Section 4, the role of $\underline{Z}_t$ will be played by business surveys.

5.2. Model with business surveys data

As mentioned above, there are a large number of balances of opinions. Thus, we face the curse of dimension. The French composite indicator is a way to reduce the dimension by assessing the overall climate of the whole French economy.

In a simplified framework, the French composite indicator would be roughly equal to the GDP growth rate $Y_t \cong a + bI_t$. However, in these surveys, entrepreneurs are asked to give a qualitative appreciation of the occurred or expected changes to some variables of interest (output, order book, foreign order book, inventories...) through the three following categories: increase, no change, decrease. There is thus a difference between qualitative answers from microeconomic surveys and quantitative macroeconomic measures. To fill the gap, we introduce a non-linear variable, namely the signed acceleration of the composite indicator: $\Delta I_t |\Delta I_t|$. At this stage, we do not motivate this introduction by a firm-level model. However, out-of-sample forecasts suggest that this new variable captures GDP growth dynamics more efficiently, especially during slowdowns. We also add the lagged value of quarter-on-quarter growth. Formally, our model is defined by:

$$Y_t = Z_t' \alpha + \sigma(Z_t) u_t \quad (6)$$

with $Z_t \equiv (1, Y_{t-1}, I_t, \Delta I_t |\Delta I_t|)'$, $\alpha \equiv (\alpha_0, \alpha_1, \alpha_2, \alpha_3)'$ and u_t a white noise independent of the available information.

Once the model set, the forecaster first needs a point forecast. It is usual (though not optimal for prediction as we do not need the estimator to be unbiased) to estimate α by Ordinary Least Squares (OLS) and we get:

$$\hat{Y}_t = Z_t' \hat{\alpha}_{t-1}$$

with $\hat{\alpha}_{t-1} := (Z_{t-1}' Z_{t-1})^{-1} Z_{t-1}' Y_{t-1}$, where $Z_{t-1} = (Z_1, \dots, Z_{t-1})'$. The unconditional variance $E(\text{Var}(Y_t | Z_t))$ of the error term can then be estimated by:

$$\hat{E}(\text{Var}(Y_t | Z_t)) \equiv \frac{1}{T} \sum_t \|\hat{Y}_t - Y_t\|^2.$$

A simple way to derive conditional confidence intervals for Y_t is to use quantile regressions. In order to construct a confidence interval with probability at least 90%, it is sufficient to estimate the quantile curve at levels 5% and 95% by quantile regression on the past available observation $(y_0, z_1'), \dots, (y_{t-1}, z_t')$: $\hat{Q}_{Y_t}(0.05|z_t)$ and $\hat{Q}_{Y_t}(0.95|z_t)$. Our confidence interval of 90% at quarter t can be written as:

$$CI^{90\%}_t \equiv [\hat{Q}_{Y_t}(0.05|z_t), \hat{Q}_{Y_t}(0.95|z_t)].$$

We come to our final goal: forecasting the conditional density of GDP growth given the business surveys. Recall that the quantile curve $\theta \rightarrow Q_{Y_t}(\theta|z_t)$ completely describes the distribution. Thus, it would be sufficient to give the forecasted quantile curve $\theta \rightarrow \hat{Q}_{Y_t}(\theta|z_t)$. Instead, it is common practice among forecasters to display a density forecast $y \rightarrow \hat{f}_t(y|z_t)$.

We give here a simple heuristic to derive a density forecast $y \rightarrow \hat{f}_t(y|z_t)$ from the forecasted quantile curve $\theta \rightarrow \hat{Q}_{Y_t}(\theta|z_t)$:

1. Set N a large number (a practical choice is 100). Compute at each $u_i \equiv \frac{i}{N}$; $1 \leq i \leq N$, $y_i \equiv \hat{Q}_{Y_t}(u_i|z_t)$. Recall that $Q_{Y_t}(U|z_t)$ with U a uniform distribution is distributed as $Y_t|z_t$. So our y_i are distributed as a sequence with conditional quantile function $\hat{Q}_{Y_t}(\cdot|z_t)$.
2. Finally, compute $\hat{f}_t(y|z) \equiv \frac{1}{Nh} \sum K(\frac{y-y_i}{h})$ with K an Epanechnikov kernel⁴ and h the bandwidth. Indeed, $\hat{f}_{Y_t}(y|z_t) \equiv \frac{1}{Nh} \sum K(\frac{y-y_i}{h})$ (with K a kernel and $y_i \sim Y_t|z_t$) is a consistent estimator of $f_{Y_t}(y|z_t)$ under suitable conditions.⁵

5.3. Definition of Forecasting Risk Index (FRI)

In this section, we aim to build a Forecasting Risk Index for each quarter t , associated with this fan chart to measure the intrinsic difficulty of the forecasting exercise. At quarter t , the difficulty of forecasting is linked to the sharpness of the distribution: the sharper the distribution, the easier forecasting is. We must thus set a quantitative indicator to measure how much the distribution is spread. In the literature, the classical measures are variance or entropy. As far as the root mean square error is concerned, the root of variance suits our needs. The definition of conditional variance is given by:

$$\sqrt{\text{Var}(Y|z)}$$

We estimate $\sqrt{\text{Var}(Y|z)}$ by $\sqrt{\hat{\text{Var}}(Y|z)} \equiv \sqrt{\text{Var}_{\hat{f}_t}(Y)}$

We are now in a position to define our Forecasting Risk Index for each quarter t as:

$$\text{FRI}_t \equiv \sqrt{\hat{\text{Var}}(Y_t|z_t)}$$

When the point forecast is based on our reference Equation 6, the expected forecasting error is approximately equal to the value of the Forecasting Risk Index.

5.4. Out-of-sample validation

Recall that the forecaster is asked to guess the next conditional density of Y_t of a sequence of random variables $Y_1, \dots, Y_{t-1}$ with knowledge of the past observations $\underline{y}_{t-1} = (y_1, \dots, y_{t-1})'$ and $\underline{z}_t = (z_1, \dots, z_t)'$.

Thus, all previous quantities must be computed on a real time basis: $\hat{Y}_{t+1} = Z_{t+1}'\hat{\alpha}_t$ with $\hat{\alpha}_t \equiv (\underline{Z}_t' \underline{Z}_t)^{-1} \underline{Z}_t' \underline{Y}_t$, $CI^{90\%}_t$, $\hat{f}_{Y_t}(y|z_t)$.

If the confidence intervals for Y_t are correct, y_t should lie in $CI^{90\%}_t$ around 90% of the time.

To assess our density forecasts, we draw the rank histogram of h_t as defined in Section 2.

6. Numerical results with an application to the 2008 recession

All methods are computed thanks to R[®] software package *quantreg* developed by Koenker and *dynlm*.

6.1. Point forecasts

Parameter estimation by OLS (from 1988Q1 to 2010Q1) leads to the following chart (Table 1).

Table 1. **Parameter estimation of Equation 6 by ordinary least squares (OLS)**

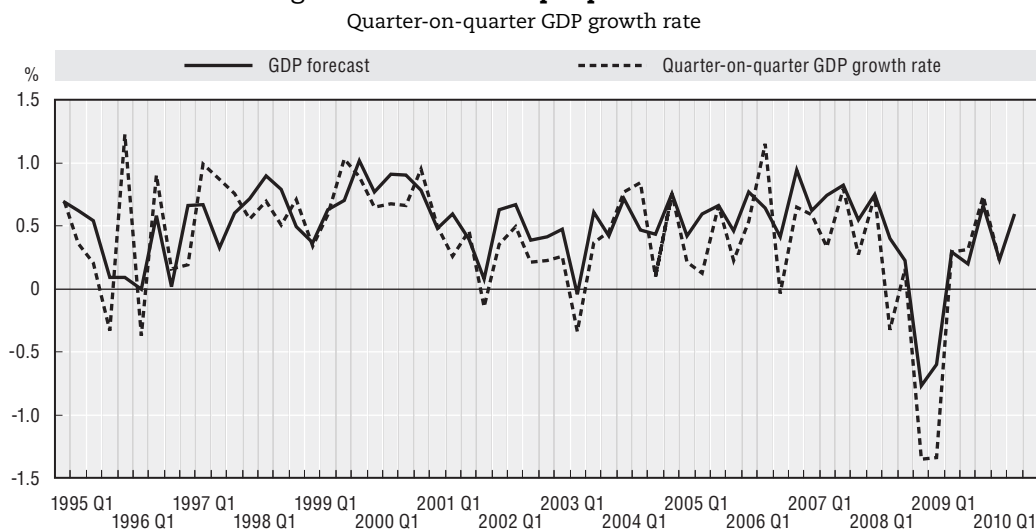
	Estimate	Std. error	T value	p-value
Intercept	0.61	0.05	11.84	***
Y_{t-1}	-0.41	0.09	-4.42	***
I_t	0.09	0.01	6.89	***
$\Delta I_t \Delta I_t $	0.10	0.01	8.28	***

*** stands for significance at the 1% level.

Source: INSEE, Author's calculations.

Figure 6 displays the out-of-sample forecasts together with the first GDP release. The residual standard error is 0.31.

Recall that we have the following decomposition formula $Var(Y_t) = E(Var(Y_t|Z_t)) + Var(E(Y_t|Z_t))$. In other words, the square error of an uninformed forecaster is equal to the error of an informed forecaster plus a residual term (knowledge term). We estimate $Var(Y_t) \cong 0.24$ and $E(Var(Y_t|Z_t)) \cong 0.12$. Thus, our reference Equation 6 gives us a 50% gain in accuracy for the square loss.

Figure 6. **Out-of-sample point forecasts**

Source: INSEE, Author's calculations.

6.2. Conditional confidence intervals

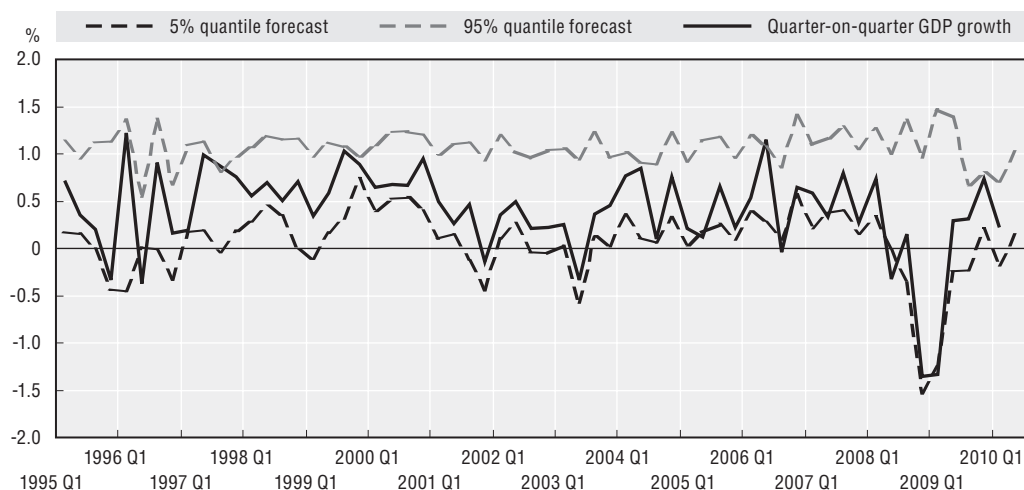
In Table 2, we obtain the estimates (from 1988Q1 to 2010Q1) for the coefficients of the quantile regression 6. Interestingly, the coefficient of the acceleration term is almost zero for the 95% quantile whereas it is significant for the 5% quantile. In other words, the troughs are much more significant than the peaks.

Figure 7 displays GDP first releases together with 90% confidence intervals of its last forecast. Note that the length of the interval depends on the forecasting date. The percentage of GDP first releases outside our 90% confidence intervals is 13% estimated on our historical data.

Table 2. **Parameter estimation of the quantile regression**

	$\theta = 0.05$	[Lower bound, upper bound]	$\theta = 0.95$	[Lower bound, upper bound]
Intercept	0.17	[0.13, 0.23]	1.23	[1.04, 1.57]
Y_{t-1}	-0.35	[-0.71, -0.33]	-0.46	[-0.98, -0.27]
I_t	0.09	[0.08, 0.14]	0.06	[0.03, 0.12]
$\Delta I_t \Delta I_t $	0.10	[0.09, 0.11]	-0.00	[-0.00, 0.18]

Source: INSEE, Author's calculations.

Figure 7. **Out-of-sample 90% confidence intervals**

Source: INSEE, Author's calculations.

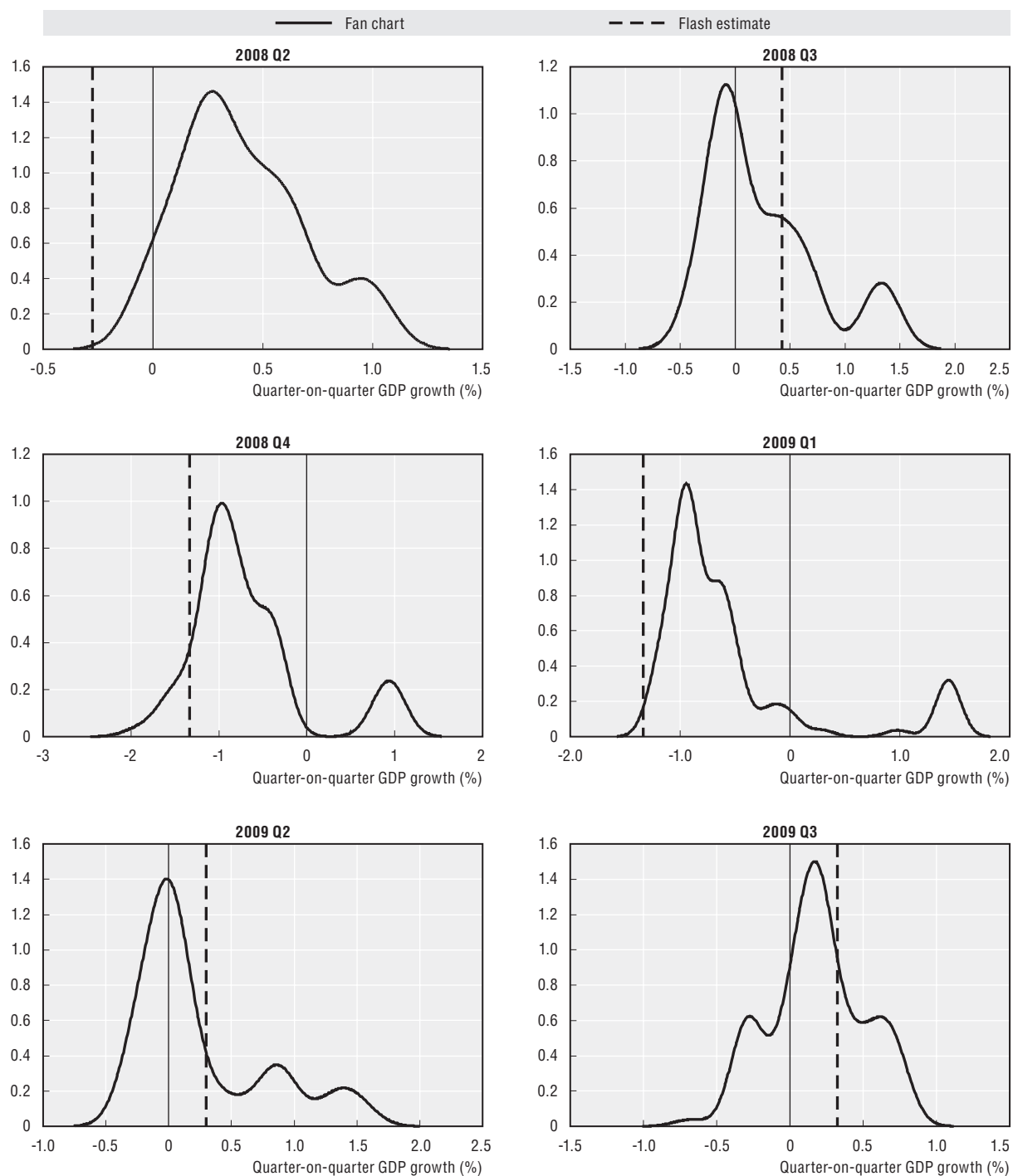
6.3. Conditional fan chart

Figure 8 shows density forecasts (methodology in Section 5) together with GDP first release (vertical line) at different quarters (from 2008Q2 to 2009Q3).

As mentioned above, we draw the rank histogram for our new fan chart (Figure 9). In comparison with the INSEE rank histogram (Figure 3), we find that the tail on the left of the first results is less underestimated by the new methodology than by that of INSEE. The bumpy fan charts come from the semi-parametric methodology as no unimodal distribution has been imposed. This raised the philosophical question of whether the error distribution should be unimodal, parametric or nonparametric. If necessary, it is always possible to display a unimodal density but with a conditional variance. For this, it is sufficient to plug the forecasting risk index, as the conditional standard error of a Gaussian density forecast.

6.4. Forecasting Risk Index

Figure 10 displays the Forecasting Risk Index. On an out-of-sample basis, the FRI exhibits clear signs of turbulence during the 2008 crisis. This gives an early signal of growing uncertainty. For example, when FRI is equal to 0.6 in 2008Q4, it can be interpreted as an expected error of 0.6 for a point forecast based on our reference model. It is worth recalling that our reference model is only based on survey data. Thus, any uncertainty that may not be reflected by business surveys such as uncertainty on oil prices or exchange rates will not be displayed by our forecasting risk index. However, the rank histogram shows that our fan charts enjoy good empirical properties. Moreover, in a Bayesian

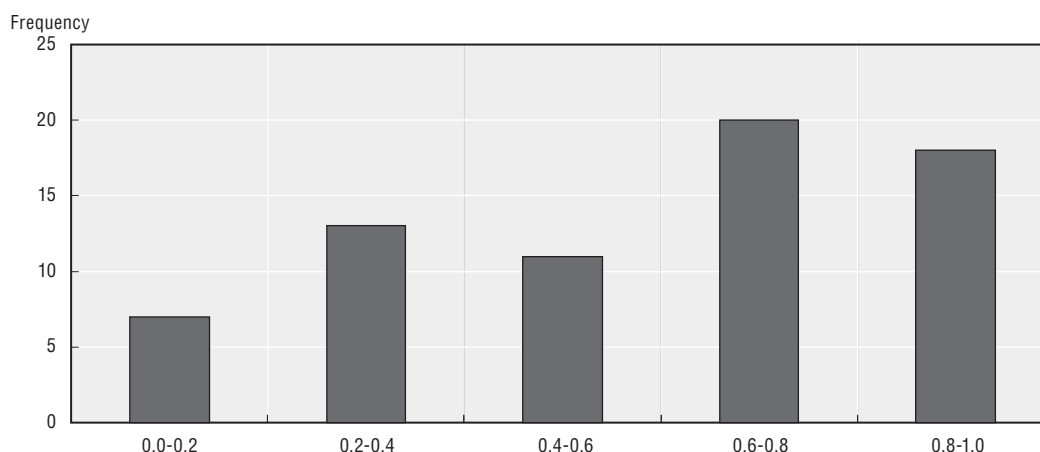
Figure 8. **Density forecasts and GDP first release (vertical line), fan chart**

Note: The density forecast on 2009Q1 shows that an extreme unconditional value for GDP growth rate is not necessarily an extreme value conditional on the available information (meaning that the first result of GDP growth crosses the density where the value of the latter is not a zero measure). For more interpretations, see paragraphs on rank histogram and on forecasting risk index.

Source: INSEE, Author's calculations.

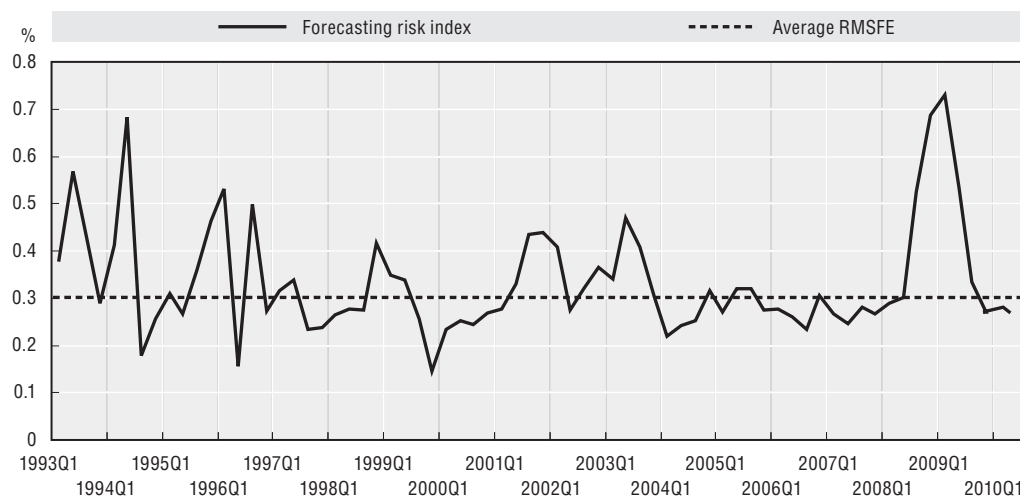
manner, an economic forecaster could always start from the uncertainty as measured by the surveys data using Equation 6 and add his own subjective assessment of uncertainty.

Figure 9. Rank histogram for the new density forecast



Source: INSEE, Author's calculations.

Figure 10. Forecasting Risk Index

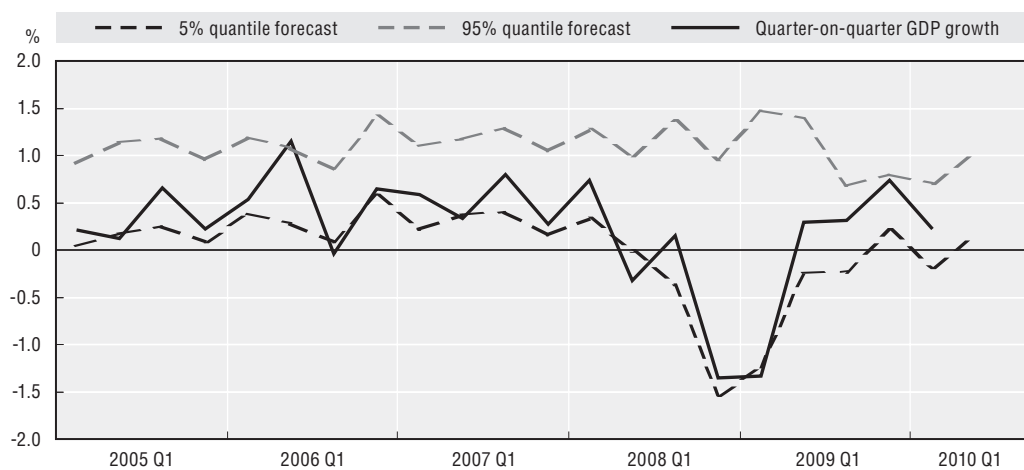


Note: RMSFE = Root mean square forecast error (mean of the sum of squared forecast errors).

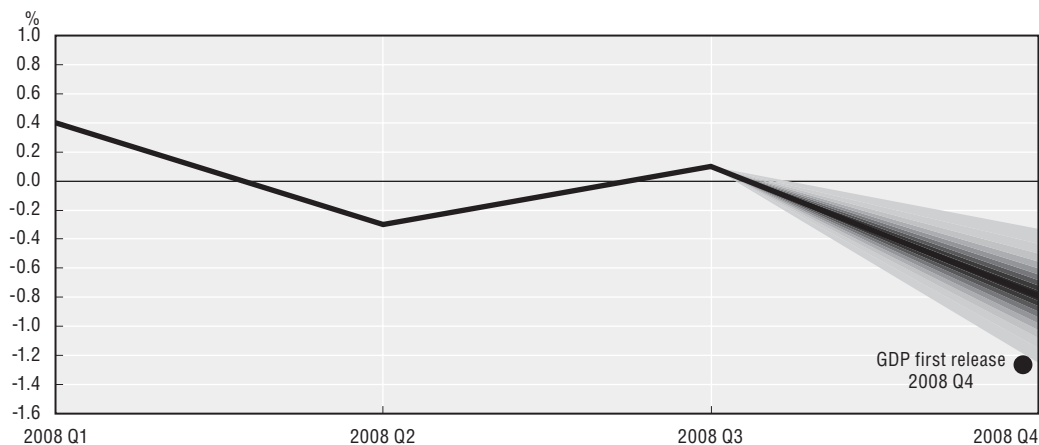
Source: INSEE, Author's calculations.

6.5. Focus on the 2008 crisis

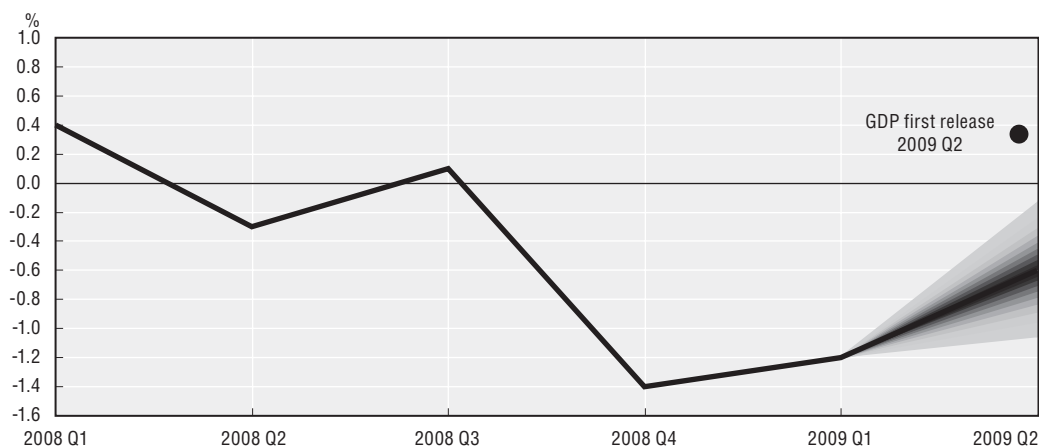
Our new GDP density forecast efficiently captures the growth stall during the crisis on a real time basis. Indeed, the first release of 2008Q4 (-1.3% in volume) lies inside the confidence interval (Figure 11). Conversely, it was almost considered as an outlier by the INSEE fan chart (Figure 12). It is also interesting to compare results during the rebound of 2009Q2. The first release (+0.3%) was in our confidence interval but it was far outside the INSEE fan chart (Figure 13). It could be surprising that our confidence interval efficiently captures the growth stall during the crisis on a real-time basis. Indeed, before 2008Q4, the minimum quarter-on-quarter GDP growth rate of our historical data was -0.6%, far above the -1.3% of 2008 Q4. This feature is made possible both by the linear form of every quantile regression and by new extreme business survey values during the 2008 crisis.

Figure 11. **Out-of-sample confidence intervals during the 2008 crisis**

Source: INSEE, Author's calculations.

Figure 12. **INSEE fan chart on 2008Q3**

Source: INSEE, French quarterly accounts, and INSEE (2008a).

Figure 13. **INSEE fan chart on 2009Q1**

Source: INSEE, French quarterly accounts, and INSEE (2009).

7. Conclusion

In this paper, we proposed a novel way of exploiting the content of survey data. We developed a reproducible methodology to build conditional density forecasts based on business survey data. Even though results are not fully comparable with the INSEE fan chart since the forecasting device is not exactly the same, this methodology seems able to better capture the uncertainty pattern associated with forecasts of French GDP. In particular, we derive a Forecasting Risk Index from the conditional density forecasts, which aims to quantify the intrinsic uncertainty of a point forecast. Interestingly, our index methodology leads to uncertainty that increases in recessions, a characteristic that INSEE forecasts have witnessed in the 2008 crisis.

Notes

1. This is also the case for INSEE (see 2008a). During the 2008 financial crisis, INSEE manually broadened the range of uncertainty of the fan chart (as the underlying method does not allow it) to reflect the increased risk forecasting in a particularly uncertain time. However, this is an exceptional approach in the case of INSEE.
2. The balance of opinion is defined as the difference between the proportion of respondents having expressed a positive opinion and the proportion of respondents having expressed a negative opinion.
3. To be more precise, we define i_t as the mean of the last three known monthly releases when forecasting takes place.
4. $K(u) \equiv \frac{3}{4}(1-u^2)1_{\{|u| \leq 1\}}$
5. For example, $(Y_t)_t$ are independent identically distributed (i.i.d.) conditionally on Z_t , $K(u) \geq 0$, $\int K(u)du = 1$, $\int u^2 K(u)du \neq 0$, and K symmetric, f continuous, $h_t \rightarrow 0, Nh_t \rightarrow \infty$.

References

- Berkowitz, J. (2001), "Testing density forecasts, with applications to risk management", *Journal of Business and Economic Statistics*, Vol. 19, pp. 465-474.
- Biau, G. and B. Patra (2011), "Sequential quantile prediction of time series", *IEEE Transactions on Information Theory*, Vol. 57, pp. 1664-1674.
- Britton, E., P. Fisher and J. Whitley (1998), "The inflation report projections: Understanding the fan chart", *Bank of England Quarterly Bulletin*, Vol. 38/1, pp. 30-37.
- Chernozhukov, V., I. Fernandez-Val and A. Galichon (2010), "Quantile and probability curves without crossing", *Econometrica*, Vol. 78/3, pp. 1093-1125.
- de Finetti, B. (1990), *Theory of probability*, Vol. 1-2, John Wiley & Sons Ltd., Chichester, Reprint of the 1975 translation.
- Hamilton, J.D. (1989), "A new approach to the economic analysis of nonstationary time series and the business cycle", *Econometrica*, Vol. 57/2, pp. 357-384.
- Institut National de la Statistique et des Études Économiques (INSEE) (2008a), *Note de Conjoncture*, December, Paris, p. 15.
- INSEE (2008b), *Note de Conjoncture*, June, Paris, pp. 15-18.
- INSEE (2009), *Note de Conjoncture*, June, Paris.
- Kocherginsky, M., X. He and Y. Mu (2005), "Practical confidence intervals for regression quantiles", *Journal of Computational and Graphical Statistics*, Vol. 14/1, pp. 41-55.
- Koenker, R. (2005), *Quantile Regression*, Vol. 38, Cambridge University Press.
- Koenker, R. and G. Bassett Jr. (1982), "Robust tests for heteroscedasticity based on regression quantiles", *Econometrica*, pp. 43-61.
- Koenker, R. and G. Bassett Jr. (1978), "Regression quantiles", *Econometrica*, pp. 33-50.

- Österholm, P. (2009), "Incorporating judgement in fan charts", *Scandinavian Journal of Economics*, Vol. 111/2, pp. 387-415.
- Robertson, J.C., E.W. Tallman and C.I. Whiteman (2005), "Forecasting using relative entropy", *Journal of Money, Credit, and Banking*, Vol. 37/3, pp. 383-401.
- Rosenblatt, M. (1952), "Remarks on a multivariate transformation", *The Annals of Mathematical Statistics*, Vol. 23/3, pp. 470-472.
- Tay, A.S. and K.F. Wallis (2002), "Density forecasting: A survey", in Clements M.P. et Hendry D.F. (eds.), *Companion to Economic Forecasting*, Oxford, Blackwell, pp. 45-68.
- Waggoner, D.F. and Tao Zha, (1999), "Conditional Forecasts In Dynamic Multivariate Models", *The Review of Economics and Statistics*, MIT Press, Vol. 81(4), pages 639-651, November.
- Wallis, K.F. (2003), "Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts", *International Journal of Forecasting*, Vol. 19/2, pp. 165-175.
- Wei, Y. and X. He (2006), "Conditional growth charts", *The Annals of Statistics*, Vol. 34/5, pp. 2069-2097.

Do business tendency surveys help in forecasting employment? A real-time evidence for Switzerland

by

Boriss Siliverstovs*, **

This study investigates the usefulness of business tendency surveys collected at the KOF Swiss Economic Institute and aggregated in the form of the KOF Employment Indicator for short-term forecasting of employment in Switzerland. We use a real-time dataset in order to simulate the actual predictive process using only information that was available at the time when predictions were made. We evaluate the predictive content of the KOF Employment Indicator both for nowcasts that are published two months before the first official release, and for one-quarter ahead forecasts published five months before the first official release. We find that inclusion of the KOF Employment Indicator leads to a substantial improvement in prediction accuracy of both point and density forecasts compared to the performance of a benchmark autoregressive model.

Keywords: Business tendency surveys, Density forecasting, Real-time data, Bayesian model averaging, Employment

JEL code: C11, C22, C53, E24

* ETH Zurich, KOF Swiss Economic Institute, Weinbergstrasse 35, 8092 Zurich, Switzerland, e-mail: boriss.siliverstovs@kof.ethz.ch.

** The paper has benefited from comments made by Niklas Ahlgren, Richard Etter, Roland Aeppli, two anonymous referees and participants at the KOF Brown Bag seminar (Zurich, Switzerland), the BICEPS seminar (Riga, Latvia), the 23rd International Symposium on Forecasting (Hong Kong, China), the Annual Meeting of the Swiss Economic Society (Geneva, Switzerland), the 10th IWH Workshop on "Recent Advances in Forecasting" (Halle, Germany), the DIW Macroeconometric Workshop (Berlin, Germany) and the research seminar at the National Bank of Estonia (Tallinn, Estonia). All computations were performed in R (www.cran.r-project.org) and in Ox version 6.00 (Doornik, 2007). The GAUSS code of Mico Loretan was used for non-parametric density estimation.

1. Introduction

Decision-making institutions face a great deal of uncertainty regarding not only the future path of the economy but also its current state. The uncertain knowledge about the current state of economic activity stems from the fact that usually relevant economic data are only available with a significant delay. Up to date, a significant body of literature has evolved that attempts to reduce the uncertainty about current and future developments in the economy by relying on coincident/leading indicators constructed on the basis of business (and/or consumer) tendency surveys. Business and consumer tendency surveys reflect an assessment of the current situation as well as recent and expected developments as perceived by business and consumers, respectively. Due to the fact that their publication precedes that of economic data, they are readily available to decision makers and hence can be useful for an early assessment of the state of the economy.

In this article we investigate the usefulness of business tendency surveys (BTS) collected at the KOF Swiss Economic Institute at ETH Zurich for short-term forecasting of employment in Switzerland. To this end, we utilise the KOF Employment Indicator which is constructed using answers to employment-related questions from industry-specific questionnaires. We assess the predictive content of this indicator by evaluating the accuracy of nowcasts for employment growth in the current quarter as well as accuracy of the one-quarter ahead forecasts. Given a nearly two-month publication lag of official data by the Swiss Federal Statistical Office our nowcasts precede the first official publication by two months and our one-quarter ahead forecasts by five months.

Our study contributes to the literature as follows. First, to the best of our knowledge, this is the first attempt in the forecasting literature to assess the usefulness of business tendency surveys (BTS) for short-run forecasting of employment growth in Switzerland. Previous studies have addressed similar questions for Canada (Hartle, 1958) and Germany (Abberger, 2007b). The Hartle study yielded negative conclusions, whereas the Abberger study found that BTS collected at the Ifo Institute in Munich are very useful for predicting employment changes in Germany. Given such contradictory results there is a need for more case studies, laying the motivation for the present study.

We assess the predictive content of the KOF Employment Indicator using real-time data; i.e., for every point in time we utilise historical data vintages reflecting available information at the time of forecasting. The importance of assessing the forecast accuracy of leading indicator models using real-time rather than latest-available data vintages has already been emphasised in the literature (see for example Croushore, 2005; Diebold and Rudebusch, 1991). Taking into account that Abberger (2007b) carries out his exercise with latest-available data, our study further distinguishes itself by utilising the real-time approach.

We employ a Bayesian model averaging (BMA) framework instead of relying on a single-best model approach based either on minimisation of some information criteria or a more sophisticated model selection procedures like PcGets advocated in Hendry and

Krolzig (2001). The latter approach is very popular when forecasting with leading indicator models: see for example a seminal study of Stock and Watson (2002) or more recent studies like Golinelli and Parigi (2008) or Drechsel and Scheufele (2012). The advantages of BMA are well documented in practice (*e.g.*, see Hoeting, Raftery, and Volinsky, 1999), and it is widely used in the forecasting literature. However, to the best of our knowledge, the present study represents the first application of BMA in the context of assessing the usefulness of BTS for forecasting employment.

Finally, we evaluate out-of-sample forecasting performance of competing models both for point and density forecasts, contributing to, the small but growing literature on density forecasting (*e.g.*, Mitchell and Wallis, 2011; Kascha and Ravazzolo, 2010; Hall and Mitchell, 2007; Mitchell and Hall, 2005).

Our main findings suggest that inclusion of the KOF Employment Indicator results in substantial improvement both with in-sample as well as, more importantly, with out-of-sample prediction accuracy compared with the univariate autoregressive benchmark model that excludes the indicator.

The rest of this paper is structured as follows. A brief literature survey is presented in Section 2. Section 3 describes the data used in our predictive exercise. The econometric methodology utilised in our study is described in Section 4. Section 5 discusses in-sample estimation results as well as results of the out-of-sample predictions. The final section concludes.

2. Literature survey

Despite the widespread use of business tendency surveys in forecasting, in most cases the target variables for which forecast are made are either GDP or manufacturing/industrial growth (see *e.g.* Abberger, 2007a; Hansson *et al.*, 2005; Lemmens *et al.*, 2005; Balke and Petersen, 2002; Lindström, 2000; Kauppi *et al.*, 1996; Öller and Tallbom, 1996; Bergström, 1995; Rahiala and Teräsvirta, 1993; Öller, 1990; Hanssens and Vanden Abeele, 1987; Teräsvirta, 1986; Zarnowitz, 1973).

At the same time, the use of BTS for forecasting of employment has received disproportionately less attention. To the best of our knowledge, there are only two academic studies that specifically assess the predictive content of BTS for forecasting employment. As early as in 1958, Hartle (1958) summarised the accuracy of employment predictions derived from the Employment Forecast Survey collected on behalf of the Canadian government for the period from 1946 until 1957. Unfortunately, it proved that these predictions were of questionable value and as noted in Hartle (1958, p. 389), "... there [were] only weak grounds for expecting that the predictions could have been made substantially more reliable in the future". On these grounds further employment predictions with the Employment Forecast Survey indices were discontinued.

More recently, Abberger (2007b) investigated whether qualitative business surveys collected at the Ifo Institute can be used for assessment of employment changes in Germany. To this end, various approaches including non-parametric regression methods, error correction models, and probit models were used. In contrast to the negative conclusion reached by Hartle (1958), Abberger (2007b, p. 258) finds that "All methods indicate that the survey results are very useful for assessing actual employment changes, and they show that the survey-based indicator leads the actual employment by between two and four months". The results reported in Abberger (2007b) are quite strong.

Nevertheless, the dispute on the usefulness of BTS for forecasting changes in employment is not yet fully resolved, and therefore there is a need for more case studies addressing this issue.

3. Data

The time series to be predicted is the quarterly annual growth rate of total employment expressed in full-time equivalent, released at quarterly frequency by the Swiss Federal Statistical Office (SFSO). The publication lag of employment data is about two months after the end of the quarter for which the data are released. We aim to forecast the year-to-year quarterly growth rate of total employment (in full-time equivalent) by means of the KOF Employment Indicator (KOFEI), regularly released by the KOF Swiss Economic Institute at ETH Zurich. This indicator is constructed as a weighted average of 19 industry-specific employment-related business tendency survey indicators. The weights are defined as the employment shares in the respective sectors.

One characteristic feature of the KOFEI is that during the period of investigation the number of industries was gradually increased from three to nine. Information on the changing composition of the KOF Employment Indicator is provided in Table 1, where the column *Coverage* reports the percentage share of employment in the sectors incorporated in the KOFEI in total employment. Observe that this coverage increased from about 35% in 1994 to about 85% at the present. This column also demonstrates the importance of the services sector (labeled as “DLU”) in the Swiss economy. Unfortunately, for this sector employment-related information are only available since 2006Q4.

Table 1. **KOF Employment Indicator: Changes over time**

Period	Sectors covered ¹	No. of indicators	Coverage ² (in %)
1991Q3-1994Q1	IMT, GGT, GHU	6	35.2
1994Q2-1994Q3	IMT, GGT, GHU, DHU	8	44.0
1994Q4-1996Q1	IMT, GGT, GHU, DHU, BAT	10	53.4
1996Q2-2000Q1	IMT, GGT, GHU, DHU, BAT, AIT	11	52.6
2000Q2-2001Q2	IMT, GGT, GHU, DHU, BAT, AIT, BT	14	56.0
2001Q3-2006Q3	IMT, GGT, GHU, DHU, BAT, AIT, BT, VT	17	55.3
2006Q4-2008Q4	IMT, GGT, GHU, DHU, BAT, AIT, BT, VT, DLU	19	84.5

1. “IMT” stands for manufacturing, “GGT” for hotels and restaurants, “GHU” for wholesale trade, “DHU” for retail trade, “BAT” for construction, “AIT” for architects and engineers, “BT” for banking, “VT” for insurance, “DLU” for services. All employment is expressed in full-time equivalent.

2. Reports the share of sectoral employment covered by the KOF Employment Indicator in total employment.

Another important feature of the KOFEI is the presence of both monthly and quarterly survey indicators. However, to the former group belong only surveys collected in industrial “IMT” and retail trade “DHU” sectors. Aggregating from monthly to quarterly frequency monthly observations is done by the simple averaging. The rest of the surveys are collected at a quarterly frequency. The full list of the survey indicators is given in Table 2, which presents the composition of the indicator. For each sector we assigned an equal weight to the questions regarding assessment of current employment and its prospects for the next three months. Observe that due to the fact that for sectors such as construction, banking and architects/engineers the question regarding the current employment assessment is not included in the questionnaires, it was substituted by an assessment of the current

Table 2. **KOF Employment Indicator: Components**

Sector	Label	Indicator	Weight ¹ %
Manufacturing	IMT	Current employment assessment	50
		Employment expectations for next 3 months	50
Hotels/restaurants	GGT	Current employment assessment	50
		Sales expectation for next 3 months	50
Wholesale trade	GHU	Current employment assessment	50
		Employment expectations for next 3 months	50
Retail trade	DHU	Current employment assessment	50
		Employment expectations for next 3 months	50
Construction	BAT	Current business situation assessment	50
		Employment expectations for next 3 months	50
Architects/engineers	AIT	Current business situation assessment	50
		Employment expectations for next 3 months	50
Banking	BT	Frontoffice assessment of current situation	25
		Backoffice assessment of current situation	25
		Employment expectations for next 3 months	50
Insurance	VT	Current employment assessment	50
		Employment expectations for next 3 months	50
Services	DLU	Current employment assessment	50
		Employment expectations for next 3 months	50

Note: The table describes the composition of the KOF Employment Indicator for 2010Q4; i.e. for the last quarter of the prediction sample used in the paper.

1. Denotes weights assigned to each survey question within each sector.

business situation. Similarly, the indicator on future employment prospects in hotels/restaurants is substituted by expectations on future sales.¹

We perform our forecasting exercise in real time, i.e. at each of our forecasting vintages we strictly only use information on employment as it was available at the time of forecast origin. For this purpose we employ the vintages of the real-time data on total employment as well as its sectoral components as they were available, starting with the prediction period from 2004Q4 until 2010Q4.² The parameters of forecasting equations are estimated by means of a rolling regression, fixing the length of the estimation window at the estimation window length employed in order to produce the first nowcast and forecast for 2004Q4 and 2005Q1, respectively. The resulting estimation window is from 1993Q4 until 2004Q3, after taking all necessary transformations including fourth-order differencing of the employment series, and taking up to the fifth-order lags of both variables, yielding the sample size of 44 observations. The next round of nowcast and forecast for 2005Q1 and 2005Q2, respectively, is based on an estimation window from 1994Q1 until 2004Q4, and so on. This implies that the forecast sample is from 2004Q4 until 2010Q4 for nowcasts and from 2005Q1 until 2010Q4 for one-step ahead forecasts, yielding 25 and 24 predictions, respectively.

The relatively short forecast sample is largely justified by the changing composition of the indicators that enter in the KOFEI. Observe that the surveys covering employment outlook in the service sector (“DLU”) only appear in the indicator from 2006Q4 onwards. Hence, by employing the current forecast sample, we attempt to strike a balance between a choice of starting the forecasting exercise as early as in 2006Q4, i.e. evaluating the forecasting properties of the KOFEI in its most recent form, and a choice of extending the forecast sample to several earlier years. In the former case, we significantly shorten the

forecast period, whereas in the latter case in evaluating the forecast accuracy we risk putting too much weight on forecasts made ignoring developments in the service sector.

The timing of our prediction exercise – the end of the third month of a quarter for which a nowcast is made – implies that our nowcasts and forecasts precede the official data release by two and five months, respectively. We compare our predictions with the first published official data. That is, a nowcast made for 2004Q4 is compared with the growth rate in total employment published about two months after the end of the fourth quarter of 2004, i.e. in February 2005. The corresponding nowcast and forecast for 2005Q1 are compared with the growth rate in total employment published in May 2005, and so on.

The quarterly year-on-year growth of total employment and the KOF Employment Indicator time series are shown in Figure 1 using the data vintages that cover the last quarter in our forecast sample, 2010Q4. In the figure both time series were adjusted to have the same mean and range in order to facilitate visual comparison. Both time series display a close comovement. However, the KOFEI has a tendency to lead the reference time series. Below we investigate whether such a qualitatively encouraging conclusion also withstands a quantitative evaluation of the predictive content of the KOFEI.

Figure 1. **Total employment in full-time equivalent (year-on-year growth) and KOF Employment Indicator: Both time series are adjusted to have the same mean and range**



4. Econometric methodology

4.1. Forecasting models

We used autoregressive distributed lag (ARDL) models in the following form for nowcasting:

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \sum_{j=0}^q \beta_j x_{t-j} + \varepsilon_t \quad (1)$$

and for forecasting:

$$y_t = \alpha_0 + \sum_{i=2}^p \alpha_i y_{t-i} + \sum_{j=1}^q \beta_j x_{t-j} + \varepsilon_t, \quad (2)$$

In these equations, y_t is the year-to-year quarterly growth rate of total employment (in full time equivalent) observed in quarter t , and x_t is the KOFEI for quarter t . The structure of these two equations implies that the current growth rate of total employment is projected on its own past values and the contemporaneous and past values of the KOFEI in equation (1) and on its own past values and the past values of the KOFEI in equation (2). The maximum lag length is set to five, i.e. $p = q = 5$, and ε_t is a normally distributed error term with variance σ_ε^2 .

As benchmark models we chose the following univariate autoregressive models which are nested in the ARDL models above used for nowcasts:

$$y_t = \alpha_0 + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t \quad (3)$$

and one-quarter ahead forecasts:

$$y_t = \alpha_0 + \sum_{i=2}^p \alpha_i y_{t-i} + \varepsilon_t, \quad (4)$$

These models retain the same structure as Equations (1) and (2), respectively, but exclude the KOFEI. By comparing the forecasts produced by the model with the leading indicator with those produced by the benchmark model, we can evaluate both the in-sample as well as out-of-sample predictive content of the KOFEI.

In general, an ARDL equation allows 2^k combinations of regressors, where k is the number of regressors except the constant term which is always retained in the estimation. Given such a multitude of model specifications, we chose to conduct our exercise using a Bayesian model averaging (BMA) approach, rather than concentrating on a “single-best”-model approach. The BMA approach allows us to incorporate the following three sources of uncertainty into forecasting process: error term uncertainty, parameter estimation uncertainty and model uncertainty. Observe that predictions based on a single-model approach typically accommodate only the first and at best, the second sources of

uncertainty. Assessment of model uncertainty and henceforth, its incorporation into the prediction process is *per definitionem* ruled out in the latter approach. We believe that accounting for model selection uncertainty is especially important when dealing with real-time data vintages that often undergo (substantial) revisions inducing both changes in the temporal dependence structure of a time series of interest as well as changes in the interdependence structure between the variables in forecasting models.

The BMA approach allows us to consider either all possible combinations of the regressors in our predictive exercise, or to concentrate on a subset of models with highest posterior probability. According to the former approach, for model comparison one has to evaluate model posterior probabilities for all the possible combinations of lags of y and x . This may require significant computational time. To get around this, we follow Madigan and Raftery (1994) and apply an approach of model selection based on Occam's window. According to this approach we exclude

“(a) models that are much less likely than the most likely model – say 20 times less likely, corresponding to a BIC (or BIC') difference of 6; and (optionally) (b) models containing effects for which there is no evidence—that is, models that have more likely submodels nested within them. The models that are left are said to belong to Occam's window, a generalization of the famous Occam's razor, or principle of parsimony in scientific explanation. When both (a) and (b) are used, Occam's window is said to be strict, and when only (a) is used it is said to be symmetric” (Raftery, 1995, p. 146).

One can adjust the severity of model selection either by changing the ratio in (a), and/or by applying a strict rather than symmetric Occam's window.

The parameters of each model specification retained in Occam's window are estimated using the Monte Carlo Markov Chain simulation algorithm. This allows us to produce finite-sample predictive densities, rather than those based on asymptotic approximations. We use the linear opinion pool in order to combine individual densities. This is achieved numerically by simulating from each predictive density with the weights determined by the posterior probability of the respective models.³ Given diffuse priors and equal model prior probabilities, the weights in the BMA are approximated by

$$\omega_j = \frac{\exp\left(-\frac{1}{2} \text{BIC}_j\right)}{\sum_{j=1}^M \exp\left(-\frac{1}{2} \text{BIC}_j\right)}.$$

with non-negative weights w_j that sum to one. The combined predictive density is estimated by the non-parametric method described in Silverman (1986, p. 63-66). The point forecasts are obtained as mean values of combined predictive densities.⁴

4.2. Forecast accuracy evaluation

The formal assessment of the statistical differences in the accuracy of point and density forecasts is carried out within the testing framework suggested in Giacomini and White (2006), which is “... is well suited for comparing methods based on nested models or on different modeling and estimation techniques” (Giacomini and White, 2006, p. 1564). The crucial requirement is that forecasts should be produced using a “rolling-window” rather than “expanding-window” scheme.

We apply unconditional test of equal predictive ability in order to distinguish between the point forecast accuracy of competing models. Denote a pair of h -steps ahead competing forecasts as $(f_{1,t}, f_{2,t})$, such that the corresponding forecast errors are $(y_t - f_{1,t}, y_t - f_{2,t}) = (e_{1,t}, e_{2,t})$. The test is based on

$$E[L(e_{1,t}) - L(e_{2,t})] = 0,$$

where the loss function $L(\cdot)$ provides the metric for evaluation of forecast accuracy. A popular choice for the loss is the quadratic function, *e.g.*, $L(e_t) = e_t^2$. Defining the loss differential, $d_t = L(e_{1,t}) - L(e_{2,t})$, the expectation operator in the above equation is replaced by its empirical counterpart:

$$\bar{d} = \frac{1}{T} \sum_{t=1}^T [L(e_{1,t}) - L(e_{2,t})] = \frac{1}{T} \sum_{t=1}^T d_t, \quad (5)$$

where T specifies the number of forecasts available for a given horizon h . Giacomini and White (2006) prove that under appropriate regularity conditions the unconditional test for equal predictive ability is based on the Diebold-Mariano test statistic:

$$\frac{\bar{d}}{\widehat{\sigma_T} / \sqrt{T}} \xrightarrow{d} N(0, 1), \quad (6)$$

where $\widehat{\sigma_T}^2$ is a heteroskedasticity- and autocorrelation-consistent (HAC) estimator of the asymptotic variance, $\widehat{\sigma_T}^2 = \text{var}(\sqrt{T} \bar{d})$.

Furthermore, we may capitalise on comment 6 in Giacomini and White (2006, p. 1555) that states that the framework is also applicable for testing forecast encompassing using an appropriately defined loss differential. Here we follow Harvey, Leybourne, and Newbold (1998), and define $d_t = e_{1,t} (e_{1,t} - e_{2,t})$. Then for defined as above d_t we can use the test statistic in Equation (6) in order to test whether forecasts of one model encompass those of another model. Naturally, we can test the reverse hypothesis that the forecasts of the second model encompass those of the first model.

In order to compare the accuracy of predictive densities of the competing models we apply the test suggested in Mitchell and Hall (2005).⁵ Denote by $g_1(\cdot)$ and $g_2(\cdot)$ the predictive densities of the competing models. Then for a given outturn y_t the corresponding log scores are $\ln g_1(y_t)$ and $\ln g_2(y_t)$. Higher (log) scores indicate more accurate density forecasts. By defining the loss differential $d_t = \ln g_1(y_t) - \ln g_2(y_t)$, the testing of the null hypothesis of equal accuracy of density forecasts corresponds to testing that the expected loss differential is zero, *i.e.* $E[d_t] = 0$. The resulting test statistic is given in Equation (6).

5. Results

5.1. In-sample estimation results

Tables 3 and 4 present an excerpt of the results of the BMA procedure for the ARDL and AR models, respectively. We report the estimation results of Equations (1) and (3) based on an in-sample window covering the period from 1999Q4 until 2010Q3.⁶ On the basis of these estimation results a nowcast for 2010Q4 is computed using the real-time data vintages. We use the maximum ratio of 20 for excluding models from the symmetric Occam's window.⁷ Together with the model specifications selected in Occam's window, the following

Table 3. **Outcome of BMA applied to the ARDL model in Equation (1)**

	Inclusion frequency	Posterior distribution		Model						
		EV	SD	1	2	3	4	5	..	35
Incpt	100	0.397	0.131	0.368	0.535	0.358	0.447	0.277	..	0.366
y_{t-1}	100	0.720	0.120	0.617	0.738	0.733	0.796	0.714	..	0.743
y_{t-2}	8.3	-0.008	0.053	..	..	..	..	..	..	..
y_{t-3}	10.6	-0.002	0.060	..	..	..	..	..	..	..
y_{t-4}	52.4	-0.129	0.167	..	-0.282	-0.095	-0.256	..	..	-0.320
y_{t-5}	12.6	0.011	0.063	..	..	..	..	..	..	0.201
x_t	100	0.045	0.010	0.052	0.043	0.043	0.043	0.046	..	0.035
x_{t-1}	8.7	0.001	0.006	..	..	..	..	..	..	0.013
x_{t-2}	8.2	-0.001	0.005	..	..	..	..	..	..	..
x_{t-3}	5.8	0.000	0.004	..	..	..	..	..	..	..
x_{t-4}	22.9	-0.005	0.011	..	..	..	-0.021	-0.010	..	..
x_{t-5}	41.3	0.011	0.017	..	0.025	..	0.037	..	..	..
nVar				2	4	3	5	3	..	5
R^2				0.948	0.955	0.951	0.958	0.950	..	0.954
BIC				-122.2	-121.5	-120.9	-120.6	-120.1	..	-116.3
Post. Prob				0.140	0.098	0.073	0.062	0.050	..	0.007

Note: The estimation window is from 1999Q4 until 2010Q3. A symmetric Occam's window with a ratio of 20 is applied. 35 models were selected in Occam's window; only five models with the highest posterior probability and the model with the lowest posterior probability are reported. The model ensemble is used to nowcast employment growth in 2010Q4. "EV" and "SD" stand for expected value and standard deviation of the posterior distribution, respectively.

Table 4. **Outcome of BMA applied to the AR model in Equation (3)**

	Inclusion frequency	Posterior distribution		Model						
		EV	SD	1	2	3	4	5	..	10
Incpt	100	0.178	0.086	0.204	0.154	0.156	0.189	0.195	..	0.165
y_{t-1}	100	1.176	0.110	1.116	1.181	1.226	1.218	1.146	..	1.211
y_{t-2}	16.1	-0.032	0.126	..	..	..	-0.146	..	..	..
y_{t-3}	28.4	-0.074	0.157	..	..	-0.362	..	-0.085	..	-0.331
y_{t-4}	78.5	-0.284	0.220	-0.291	-0.553	..	-0.236	-0.230	..	..
y_{t-5}	27.7	0.060	0.133	..	0.238	..	..	..	..	-0.023
nVar				2	3	2	3	3	..	3
R^2				0.923	0.927	0.919	0.924	0.923	..	0.919
BIC				-105.2	-103.7	-103.2	-101.9	-101.6	..	-99.5
Post. prob.				0.389	0.190	0.144	0.077	0.064	..	0.022

Note: The estimation window is from 1999Q4 until 2010Q3. A symmetric Occam's window with a ratio of 20 is applied. 10 models are selected in Occam's window; only five models with the highest posterior probability and the model with the lowest posterior probability are reported. The model ensemble is used to nowcast employment growth in 2010Q4. "EV" and "SD" stand for expected value and standard deviation of the posterior distribution, respectively.

information is reported in the columns of the tables: the column *Inclusion frequency* indicates the inclusion frequency of a regressor in model specifications select in Occam's window, and the columns EV and SD report the expected value and standard deviation of the posterior distributions of the coefficients.

For nowcasting equation (1), altogether 35 models were selected in Occam's window. The model with only two regressors – the first lag of the dependent variable, y_{t-1} , and the contemporaneous value of the KOFEI, x_t , – has the highest posterior probability of 0.140 and is nested in the remaining 34 models. The value of the Bayesian Information Criterion (BIC) is -122.2. The model ranked second yields a posterior probability of 0.098.

It is informative to compare the results reported for the ARDL model with those for the AR model in Table 4. Ten models were selected in Occam's window. The model with the regressors Y_{t-1} and Y_{t-4} has the highest posterior probability of 0.389 and the value of the BIC of -105.2. Observe that the reported posterior probabilities in Tables 3 and 4 are specific to BMA model ensembles, and cannot therefore be directly compared. Nevertheless, a comparison of most likely models with and without the KOF Employment indicator can be made in terms of the reported BIC values, yielding a difference of 17. According to the Raftery (1995, Table 6 on p. 139), such a difference in the BIC values indicates "very strong" evidence in favour of the model with the KOFEI.

We replicated the comparison of most likely models with and without the KOFEI for the forecasting equations used for one-quarter ahead predictions. Based on the estimation results reported in Tables 5 and 6 for the ARDL and AR models, we notice that the BIC value for the most likely model with the KOFEI is -96.6, which should be compared with a BIC value of -60.5 for the most likely univariate model. The corresponding difference in terms of the BIC values is 36.1. Referring to Table 6 in Raftery (1995), we again conclude that there is "very strong" evidence in favour of the model with the KOFEI.⁸ In sum, the results based on in-sample estimation strongly suggest that there is a large informative content in business tendency surveys for predicting employment growth both in the current and the next quarter. It remains to be seen whether this encouraging conclusion will hold also for predictions made out of sample.⁹

Table 5. **Outcome of BMA applied to the ARDL model in Equation (2)**

	Inclusion frequency	Posterior distribution		Model						
		EV	SD	1	2	3	4	5	..	34
Incpt	100	0.754	0.145	0.665	0.698	0.794	0.898	0.827	..	0.831
y_{t-2}	92.2	0.477	0.266	0.620	0.212	0.662	0.428	0.599	..	0.641
y_{t-3}	53.9	-0.241	0.283	-0.355	..	-0.553	..	-0.521	..	-0.297
y_{t-4}	29.4	-0.104	0.214	..	..	..	-0.427	..	..	-0.472
y_{t-5}	16.3	0.025	0.096	..	..	..	..	..	..	0.197
x_{t-1}	100	0.085	0.013	0.081	0.094	0.083	0.079	0.081	..	0.081
x_{t-2}	7.6	-0.001	0.006	..	..	..	..	..	..	..
x_{t-3}	8.4	0.000	0.006	..	..	..	..	..	..	..
x_{t-4}	19.9	0.004	0.011	..	..	..	..	0.023	..	..
x_{t-5}	30.2	0.007	0.014	..	..	0.020	0.032	..	..	0.028
nVar				3	2	4	4	4	..	6
R^2				0.914	0.904	0.919	0.919	0.919	..	0.924
BIC				-96.6	-95.6	-95.5	-95.3	-95.2	..	-90.9
Post. prob.				0.142	0.083	0.082	0.075	0.070	..	0.008

Note: The estimation window is from 1999Q3 until 2010Q2. A symmetric Occam's window with a ratio of 20 is applied. 34 models are selected in Occam's window; only five models with the highest posterior probability and the model with the lowest posterior probability are reported. The model ensemble is used to make the one-quarter ahead forecast of employment growth in 2010Q4. "EV" and "SD" stand for expected value and standard deviation of the posterior distribution, respectively.

5.2. Out-of-sample forecast accuracy evaluation

The results of out-of-sample forecast accuracy evaluation are summarised in Table 7. As discussed in Section 4.2, we evaluate the accuracy of point forecasts using the unconditional test of equal predictive ability of Giacomini and White (2006), and by means

Table 6. **Outcome of BMA applied to the AR model in Equation (4)**

	Inclusion frequency	Posterior distribution		Model					
		EV	SD	1	2	3	4	5	6
Incpt	100	0.369	0.128	0.388	0.307	0.366	0.377	0.373	0.435
y_{t-2}	100	1.472	0.280	1.324	1.834	1.571	1.352	1.618	1.072
y_{t-3}	40.8	-0.345	0.502	..	-1.129	-0.477	..	-0.767	..
y_{t-4}	68.9	-0.449	0.354	-0.693	..	-0.441	-0.784	..	..
y_{t-5}	18.8	-0.029	0.142	..	..	..	0.075	-0.206	-0.489
nVar				2	2	3	3	3	2
R^2				0.787	0.778	0.793	0.787	0.787	0.762
BIC				-60.5	-58.7	-58.0	-56.7	-56.7	-55.5
Post. prob.				0.477	0.198	0.137	0.075	0.072	0.041

Note: The estimation window is from 1999Q3 until 2010Q2. A symmetric Occam's window with a ratio of 20 is applied. six models are selected in Occam's window. The model ensemble is used to make the one-quarter ahead forecast of employment growth in 2010Q4. "EV" and "SD" stand for expected value and standard deviation of the posterior distribution, respectively.

Table 7. **Outcome of BMA applied to the AR model in Equation (4)**

Giacomini and White (2006) unconditional test of equal predictive ability				
	ARDL	AR	ARDL/AR	GW (2006)
	RMSFE		Ratio	P-value
Nowcast ^a	0.257	0.332	0.773	0.124
Forecast ^b	0.328	0.649	0.505	0.007
Forecast encompassing test				
	H_0 : ARDL enc. AR H_0 : AR enc. ARDL			
	P-value			
Nowcast	0.197	0.004		
Forecast	0.798	0.004		
Mitchell and Hall (2005) test of equal density forecast accuracy				
	ARDL	AR	ARDL/AR	MH (2005)
	Average log score		Difference	P-value
Nowcast	-0.099	-0.348	0.249	0.026
Forecast	-0.371	-0.985	0.614	0.000

a. Prediction samples are for 2004Q4-2010Q4.

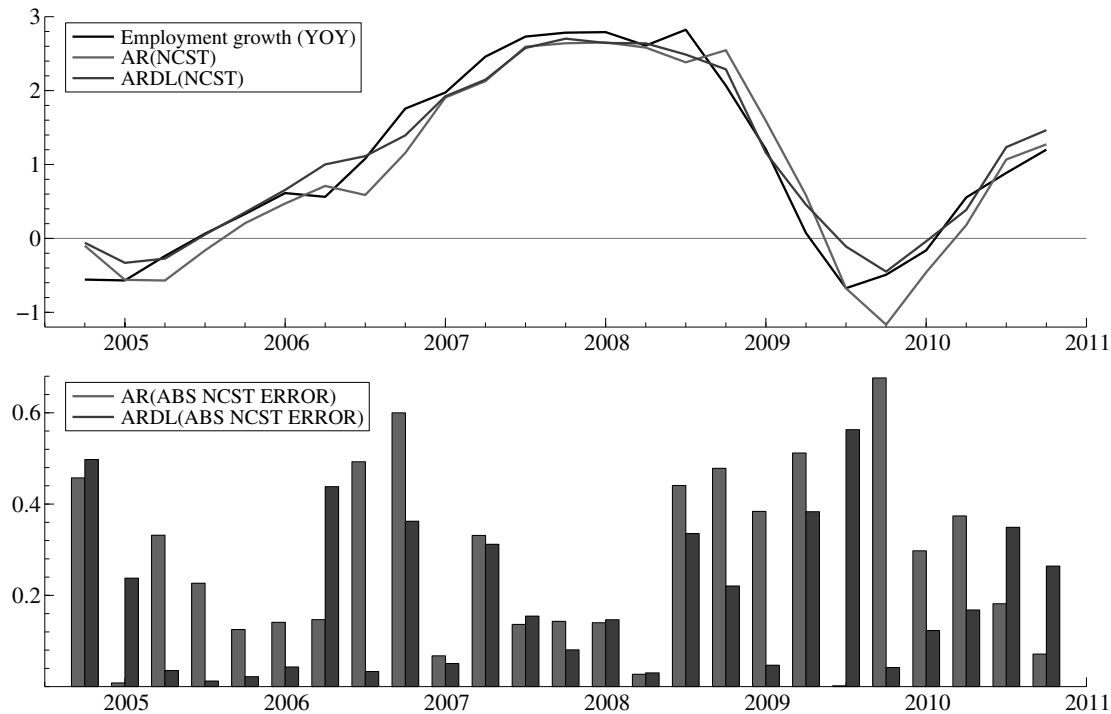
b. Prediction samples are for 2005Q1-2010Q4.

of the forecast encompassing test. The density forecasts are evaluated by means of the test of Mitchell and Hall (2005) based on averages of log scores of the competing models.

In the upper panel of Table 7, the forecast accuracy of the ARDL and AR models is compared in terms of the Root Mean Squared Forecast Errors (RMSFE). Both for the nowcasts and forecasts the ARDL model produces lower RMSFE than the AR model. The improvement in RMSFE is about 22% and 49% for nowcasts and forecasts, respectively. According to the test of Giacomini and White (2006) the null hypothesis of equal unconditional predictive ability can be rejected only at the 15% level for nowcasts. On the contrary, for one-quarter ahead forecasts this test decisively rejects the null hypothesis at the 1% significance level. The point nowcasts and one-quarter ahead forecasts together

with the actual outturns of employment growth are shown in the upper panels of Figure 2 and 3, respectively. In the lower panels the absolute nowcast and one-quarter-ahead forecast errors from the competing models are displayed.

Figure 2. **Nowcast accuracy comparison: Point forecasts**



The results of the forecast encompassing test are reported in the middle panel of Table 7. Both for the nowcasts and forecasts we cannot reject the null hypothesis that the ARDL model forecast encompasses the AR model at the usual significance levels. This is consistent with the superior performance of the ARDL model both in sample and out of sample as noted above. Next we test the hypothesis that the AR model forecasts encompass those of the ARDL model. The null of encompassing in this direction can be rejected at the 1% significance level. Consequently, we conclude that encompassing holds in one direction only, with the ARDL model forecasts encompassing the AR model forecasts.

The results of the forecast density evaluation tests are presented in the lower panel of Table 7. The null hypothesis of equal accuracy is likewise rejected by the test of Mitchell and Hall (2005) at the 5% and 1% significance levels in favour of the ARDL model for nowcasts and one-quarter ahead forecasts, respectively. The corresponding log scores for every observation in the forecast samples are displayed in Figure 4; for nowcasts in the upper panel and for one-quarter ahead forecasts in the lower panel. An example of the forecast densities for 2010Q4 is displayed in Figure 5. The vertical line corresponds to the outturn of 1.2% growth in total employment compared to the same quarter one year earlier.

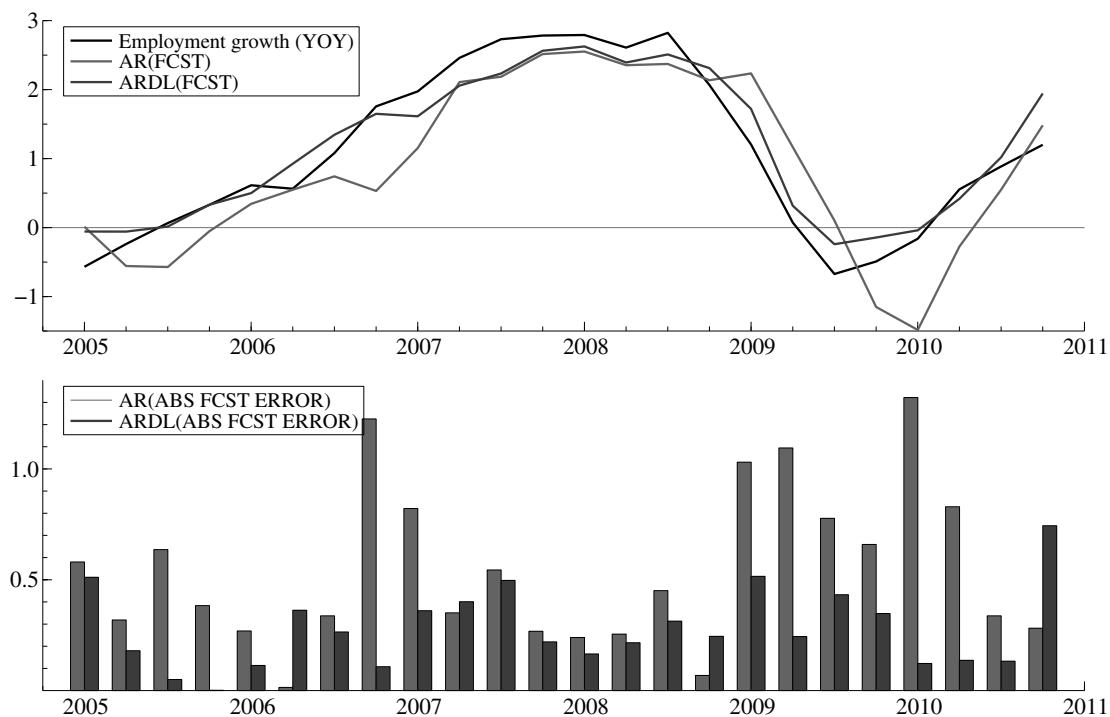
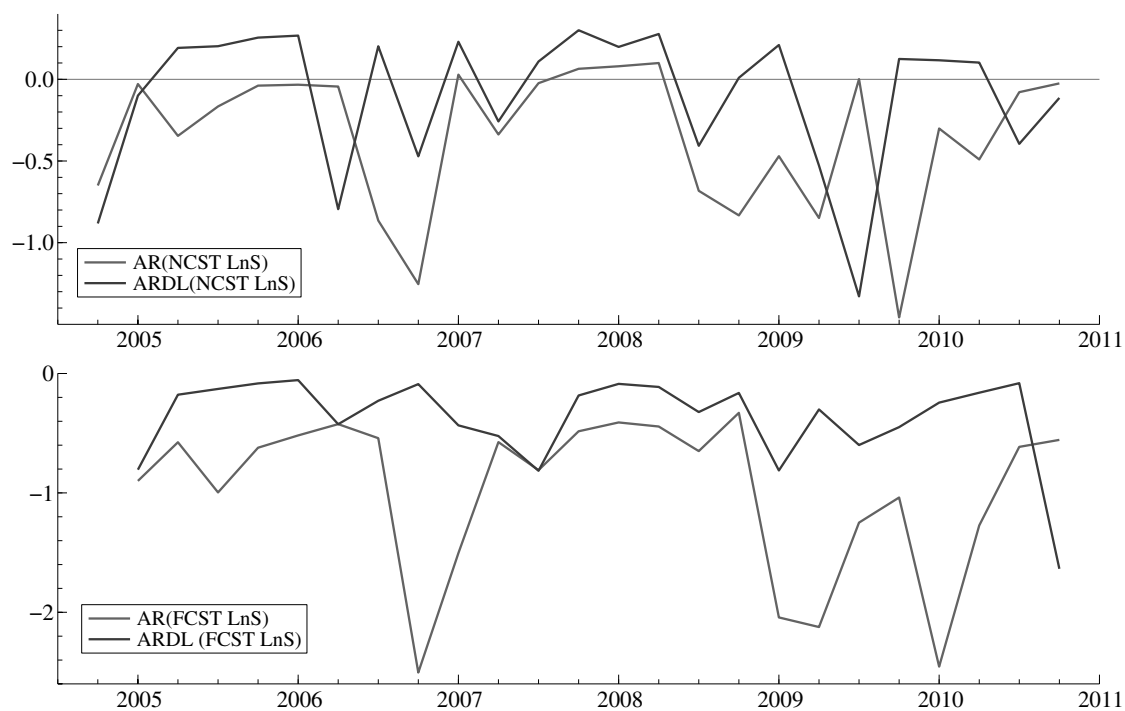
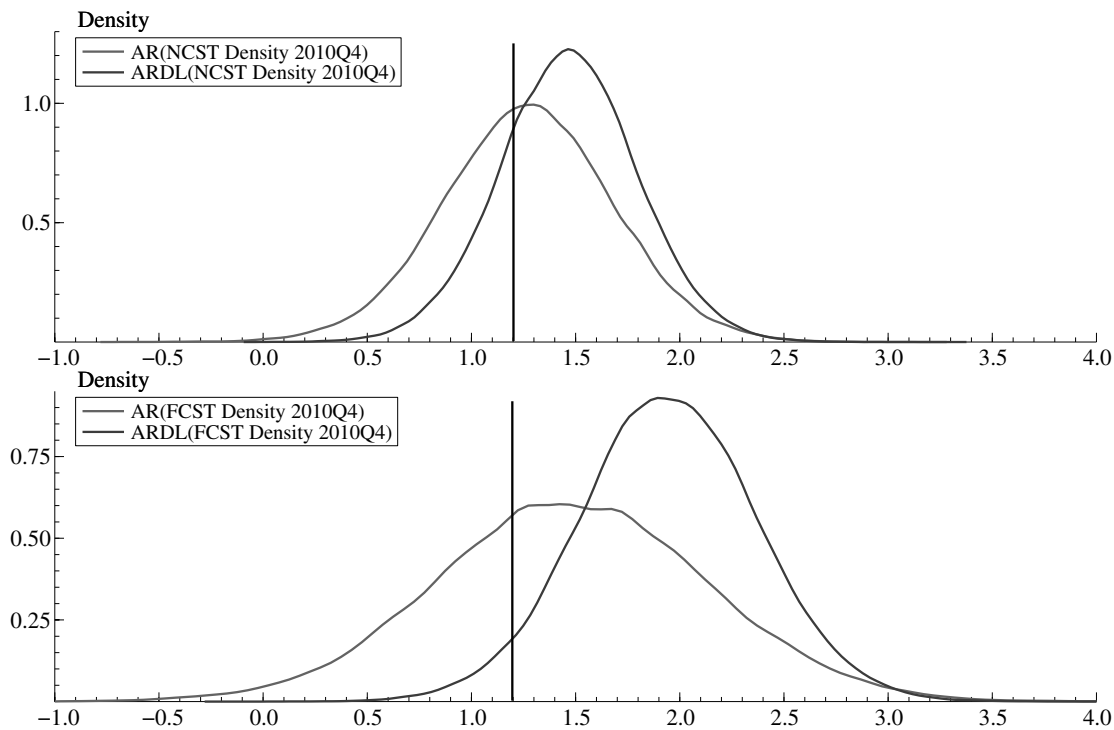
Figure 3. **One-quarter ahead forecast accuracy comparison: Point forecasts**Figure 4. **Forecast density accuracy comparison: Log scores**

Figure 5. **Nowcast and one-quarter ahead forecast densities for 2010Q4**

Note: The vertical line denotes the outturn of 1.2% annual growth in that quarter.

5.3. Robustness check

A referee suggested to check the performance of the BMA procedure not only with the univariate benchmark model but also with the “single-best” model characterised by the highest posterior probability. The results of such a comparison are reported in Table 8. For the AR model the single best model selection strategy produces slightly smaller RMSFE values and higher average log scores. However, the differences in forecast accuracy both for point and density forecasts is insignificant according to the results of the tests of Giacomini and White (2006), and Mitchell and Hall (2005). In contrast, for the ARDL model the BMA procedure resulted in slightly more accurate point and density forecasts. It is interesting to note that the small improvement in RMSFE for the one-quarter ahead forecasts turned out to be statistically significant at the 5% level.

Table 8. **Forecast accuracy evaluation: BMA versus Single Best (SB) model**

		BMA	SB	BMA/SB	GW (2006)	BMA	SB	BMA/SB	MH (2005)
		RMSFE		Ratio	P-value	Average	Log score	Difference	P-value
AR	Nowcast ¹	0.332	0.327	1.016	0.309	-0.348	-0.331	-0.017	0.139
AR	Forecast ²	0.649	0.644	1.007	0.403	-0.985	-0.980	-0.005	0.723
ARDL	Nowcast	0.257	0.261	0.984	0.580	-0.099	-0.123	0.024	0.333
ARDL	Forecast	0.328	0.341	0.961	0.050	-0.371	-0.385	0.013	0.403

1. Prediction samples are for 2004Q4-2010Q4.

2. Prediction samples are for 2005Q1-2010Q4.

To conclude, on the one hand the application of BMA did not bring about a sizable improvement in forecast accuracy in the present context. However, on the other hand it did not induce any additional deterioration in forecast accuracy. It did, however, result in a systematic, yet small, improvement in one-quarter ahead point forecasts for the ARDL model. The information contained in Tables 3-6 helps to assess model selection uncertainty by reporting the distance between the models in terms of the Bayesian Information Criterion. The fact that for out-of-sample forecasting purposes we used only the most likely models, which were selected into Occam's window, helped us to reduce the computational costs significantly compared to the situation where the full set of models is considered for Bayesian model averaging.

6. Conclusion

In this paper we evaluate the usefulness of the KOF Employment Indicator, derived from business tendency surveys collected at the KOF Swiss Economic Institute, for out-of-sample prediction of the annual growth rate of (full-time equivalent) total employment in Switzerland. We capitalise on the fact that the business tendency surveys, reflecting current and future employment expectations of firms in the current quarter, are available at least two months ahead of the first official release by the Swiss Federal Statistical Office.

We conduct our exercise in real time, *i.e.* at each forecast origin we strictly only use information that was available to a forecaster at that time. For this purpose we constructed a real-time database that contains all the official employment data vintages from 2004Q4 until 2010Q4. The data vintages cover both total employment figures as well as employment broken down by sectors. We use these sectoral employment data as the weights in constructing the KOF Employment Indicator in real time.

We produce current- and next-quarter forecasts of employment growth using autoregressive distributed lag (ARDL) models cast in a Bayesian Model Averaging framework. In doing so, we integrate model selection uncertainty in making out-of-sample predictions. We compare the predictive accuracy of the ARDL model that includes both the KOF Employment Indicator as well as the lags of the dependent variable with that of the benchmark autoregressive (AR) model from which the KOF Employment Indicator is omitted.

Our main findings are the following. The KOF Employment Indicator has an informative content that can be used to improve the prediction of the growth in total employment in Switzerland. More specifically, inclusion of the KOF Employment Indicator results in a substantial improvement in forecast accuracy compared to the performance of a benchmark AR model. For nowcasts the value of the RMSFE criterion recorded for the ARDL model is 22% lower than the value of the benchmark autoregressive model. For one-quarter ahead forecasts the improvement is much more pronounced – the corresponding reduction in the RMSFE criterion is about 49%. The results of forecast encompassing tests indicate that encompassing holds only in one direction, with the ARDL model forecasts encompassing the AR model forecasts. The ARDL model produces not only more accurate point forecasts but also more accurate density forecasts, as evaluated by means of the test of Mitchell and Hall (2005).

There are a number of relevant extensions of the current research. For example, the fact that the KOF Employment Indicator is based on industry-specific components allows us to forecast not only changes in aggregate employment but also changes in employment

in the industries in question. Another extension of the paper is to consider whether predicting aggregate employment is better done using a direct method, as it is currently done, or indirectly by summing up the industry-specific forecasts, using the approach advocated in Hendry and Hubrich (2011).

Notes

1. Table 2 describes the composition of the KOF Employment Indicator reflecting the state of the art for 2010Q4 – the last quarter in our forecast sample. Several changes were introduced in the composition of the KOFEI since the second quarter of 2011. Namely, the introduction of monthly indicators for sectors like construction, architects/engineers, banking and insurance. The questions were added to the sectoral questionnaires directly asking firms about the current employment assessment in the construction, architects/engineers and banking sectors and the employment expectations for the next three months in the hotels/restaurants sector.
2. All the employment vintages and the survey time series used for constructing the KOFEI were downloaded from the KOF Database.
3. Other weighting schemes are possible. For example, “optimal pool” weights as proposed in Hall and Mitchell (2007) could be used. Geweke and Amisano (2009) suggest an optimal pooling scheme assuming that none of the models considered is the true model. The combination scheme based on the logarithmic opinion pool represents another alternative. Bjørnland et al. (2011, Section 2.2) provide a discussion of the properties of linear and logarithmic pooling schemes.
4. The point forecasts based on median values produce very similar results.
5. Mitchell and Wallis (2011, p. 1027) provide a discussion of the Mitchell and Hall (2005) test and its relation to the likelihood ratio test of Amisano and Giacomini (2007).
6. The complete set of all estimation results covering all subperiods and models (both ARDL and AR) is not reported here for the sake of saving space but we make it available upon request.
7. Bayesian Model Averaging is carried out using the BMA package for R Estimation of the parameters is carried out using the MCMCpack package for R.
8. The full set of BIC values and their differences for every rolling estimation window is available upon request.
9. Following a suggestion of an anonymous referee we verified that there no significant departures from the assumption of parameter constancy in Equations (1) and (2). The residuals of the nowcasting equation conform to the assumption of being normally distributed and without remaining autocorrelation. However, for the residuals of the next-quarter forecasting equation we find evidence of remaining autocorrelation, which, as pointed out by the referee, could lead to a suboptimal forecasting performance. The presence of the residual autocorrelation is due to the fact that the contemporaneous values of the KOF Employment Indicator, x_t , and the first lag of the dependent variable, y_{t-1} , were omitted in Equation (2) according to the definition of the direct forecasting approach. After a thorough overview of the forecasting literature dealing with direct multi-step (DMS) estimation Chevillon (2007, p. 782) concludes that “... the misspecification of the error process in the case of DMS is not so detrimental to the accuracy of the estimators”. Further on this article we report the results of forecasting exercise without accounting for autocorrelation in the residuals. This provides us with a benchmark against which we can evaluate the influence of residual autocorrelation on forecasting accuracy in our further research.

References

- Abberger, K. (2007a), “Forecasting quarter-on-quarter changes of German GDP with monthly business tendency survey results”, *Ifo Working Paper*, No. 40, Ifo Institute for Economic Research at the University of Munich.
- Abberger, K. (2007b), “Qualitative business surveys and the assessment of employment – A case study for Germany”, *International Journal of Forecasting*, Vol. 23(2), pp. 249-258.
- Amisano, G. and R. Giacomini (2007), “Comparing density forecasts via weighted likelihood ratio tests”, *Journal of Business and Economic Statistics*, Vol. 25, pp. 177-190.

- Balke, N.S. and D. Petersen (2002), "How well does the Beige Book reflect economic activity? Evaluating qualitative information quantitatively", *Journal of Money, Credit and Banking*, Vol. 34(1), pp. 114-136.
- Bergström, R. (1995), "The relationship between manufacturing production and different business survey series in Sweden 1968-1992", *International Journal of Forecasting*, Vol. 11(3), pp. 379-393.
- Bjørnland, H.C., K. Gerdrup, A.S. Jore, C. Smith, and L.A. Thorsrud (2011), "Weights and pools for a Norwegian density combination", *The North American Journal of Economics and Finance*, Vol. 22(1), pp. 61-76.
- Chevillon, G. (2007), "Direct multi-step estimation and forecasting", *Journal of Economic Surveys*, Vol. 21(4), pp. 746-785.
- Croushore, D. (2005), "Do consumer-confidence indexes help forecast consumer spending in real time?", *The North American Journal of Economics and Finance*, Vol. 16(3), pp. 435-450.
- Diebold, F.X. and G.D. Rudebusch (1991), "Forecasting output with the composite leading index: A real-time analysis", *Journal of the American Statistical Association*, Vol. 86(415), pp. 603-610.
- Doornik, J.A. (2007), *Object-Oriented Matrix Programming Using Ox*, London: Timberlake Consultants Press and Oxford: www.doornik.com.
- Drechsel, K. and R. Scheufele (2012), "The performance of short-term forecasts of the German economy before and during the 2008/2009 recession", *International Journal of Forecasting*, Vol. 28(2), pp. 428-445.
- Geweke, J. and G. Amisano (2009), "Optimal prediction pools", *Working Paper Series 1017*, European Central Bank.
- Giacomini, R. and H. White (2006), "Tests of conditional predictive ability", *Econometrica*, Vol. 74 (6), pp. 1545-1578.
- Golinelli, R. and G. Parigi (2008), "Real-time squared: A real-time data set for real-time GDP forecasting", *International Journal of Forecasting*, Vol. 24(3), pp. 368-385.
- Hall, S.G. and J. Mitchell (2007), "Combining density forecasts", *International Journal of Forecasting*, Vol. 23(1), pp. 1-13.
- Hanssens, D.M. and P.M. Vanden Abeele (1987), "A time-series study of the formation and predictive performance of EEC production survey expectations", *Journal of Business and Economic Statistics*, Vol. 5(4), pp. 507-519.
- Hansson, J., P. Jansson, and M. Löf (2005), "Business survey data: Do they help in forecasting GDP growth?", *International Journal of Forecasting*, Vol. 21(2), pp. 377-389.
- Hartle, D. (1958), "Predictions derived from the employment forecast survey", *The Canadian Journal of Economics and Political Science/Revue canadienne d'Economie et de Science politique*, Vol. 24(3), pp. 373-390.
- Harvey, D.I., S.J. Leybourne, and P. Newbold (1998), "Tests for forecast encompassing", *Journal of Business and Economic Statistics*, Vol. 16(2), pp. 254-59.
- Hendry, D.F. and K. Hubrich (2011), "Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate", *Journal of Business and Economic Statistics*, Vol. 29(2), pp. 216-227.
- Hendry, D.F. and H.-M. Krolzig (2001), *Automatic Econometric Model Selection Using PcGets*, London: Timberlake Consultants Ltd.
- Hoeting, J., D.M.A. Raftery, and C. Volinsky (1999), "Bayesian model averaging: A tutorial", *Statistical Science* 14, pp. 382-401.
- Kascha, C. and F. Ravazzolo (2010), "Combining inflation density forecasts", *Journal of Forecasting*, Vol. 29(1-2), pp. 231-250.
- Kauppi, E., J. Lassila, and T. Terasvirta (1996), "Short-term forecasting of industrial production with business survey data: Experience from Finland's great depression 1990-1993", *International Journal of Forecasting*, Vol. 12(3), pp. 373-381.
- Lemmens, A., C. Croux, and M.G. Dekimpe (2005), "On the predictive content of production surveys: A pan-European study", *International Journal of Forecasting*, Vol. 21(2), pp. 363-375.
- Lindström, T. (2000), "Qualitative survey responses and production over the business cycle", *Working Paper Series 116*, Sveriges Riksbank (Central Bank of Sweden).

- Madigan, D. and A.E. Raftery (1994), "Model selection and accounting for model uncertainty in graphical models using Occam's window", *Journal of the American Statistical Association*, Vol. 89, pp. 1535-1546.
- Mitchell, J. and S.G. Hall (2005), "Evaluating, comparing and combining density forecasts using the KLIC with an application to the Bank of England and NIESR 'fan' charts of inflation", *Oxford Bulletin of Economics and Statistics*, Vol. 67(s1), pp. 995-1033.
- Mitchell, J. and K.F. Wallis (2011), "Evaluating density forecasts: forecast combinations, model mixtures, calibration and sharpness", *Journal of Applied Econometrics*, Vol. 26(6), pp. 1023-1040.
- Öller, L.-E. (1990), "Forecasting the business cycle using survey data", *International Journal of Forecasting*, Vol. 6(4), pp. 453-461.
- Öller, L.-E. and C. Tallbom (1996), "Smooth and timely business cycle indicators for noisy Swedish data", *International Journal of Forecasting*, Vol. 12(3), pp. 389-402.
- Raftery, A.E. (1995), "Bayesian model selection in social research", *Sociological Methodology*, Vol. 25, pp. 111-163.
- Rahiala, M. and T. Teräsvirta (1993), "Business survey data in forecasting the output of Swedish and Finnish metal and engineering industries: A Kalman filter approach", *Journal of Forecasting*, Vol. 12(3-4), pp. 255-271.
- Silverman, B.W. (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.
- Stock, J.H. and M.W. Watson (2002), "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics*, Vol. 20(2), pp. 147-162.
- Teräsvirta, T. (1986), "Model selection using business survey data: Forecasting the output of the Finnish metal and engineering industries", *International Journal of Forecasting*, Vol. 2(2), pp. 191-200.
- Zarnowitz, V. (1973), "A review of cyclical indicators for the United States: Preliminary results", *NBER Working Papers 0006*, National Bureau of Economic Research, Inc.

Referee Acknowledgement

The Editorial Board of the *Journal of Business Cycle Measurement and Analysis* gratefully acknowledges the following individuals for reviewing papers submitted to this journal in 2012 and 2013. Without their devoted and honorary effort, this publication would not be feasible.

Aguiar-Conraria, Luis Francisco	Heer, Burkhard	Pelagatti, Matteo M.
Anas, Jacques	Heij, Christiaan	Piger, Jeremy
Angeloni, Ignazio	Hervai, Saeed	Podgórska, Maria
Arnaud, Franck	Hulagu, Timur	Poser, Günter
Artis, Michael	Issler, João Victor	Próchniak, Mariusz
Assenmacher, Kathrin	Jacobs, Jan	Proietti, Tommaso
Bañbura, Marta	Kappler, Marcus	Radovic-Stojanovic
Bandholz, Harm	Kaya, Hüseyin	Raponi, Valentina
Benczúr, Péter	Keating, John	Rathke, Alexander
Berge, Travis John	Keck, Alexander	Reijer, Ard den
Billio, Monica	Knetsch, Thomas A.	Reijnders, Jan P. G.
Branch, William A.	Knoll, Christian	Richter, Christian R.
Bruzda, Joanna	Kowalczyk, Barbara	Ríos Rull, José Víctor
Buss, Ginters	Lamla, Michael	Rodríguez, Gabriel
Caporale, Guglielmo Maria	Lenart, Lucas	Sattarova, Liliya
Castro, Vítor	Lenza, Michele	Saygil, Hülya
Çelik, Sadullah	Leproux, Solange	Scheiblecker, Marcus
Chow, Hwee Kwan	Lewis, John	Scheufele, Rolf
Conti, Antonio Maria	Love, David R. F.	Schmitt, Elizabeth Dunne
Cosar, Evren Erdogan	Lyziak, Tomasz	Schnatz, Bernd
Croushore, Dean	Malgarini, Marco	Scotti, Chiara
Cubadda, Gianlucca	Marczak, Martyna	Sella, Lisa
Cuche-Curti, Nicolas	Margani, Patrizia	Seymen, Atilim
Cunningham, Rose	Massmann, Michael	Shields, Kalvinder
De Carvalho, Miguel	Maurin, Laurent	Sigalla, Fiona
D'Elia, Enrico	Mayr, Johannes	Silverstovs, Boris
Di Fonzo, Tommaso	Mazzi, Gian Luigi	Silvestrini Andrea
Dijk, Dick van	McElroy, Tucker Sprague	Simonelli, Saverio
Dinu, Marin	Meyler, Aidan	Smirnov, Sergey
Duarte, Augustin	Mitchell, James	Starica, Catalin
El Dakkak, O.	Moauero, Filippo	Stark, Tom
Fichtner, Ferdinand	Modugno, Michele	Strawczynski, Michel
Fonzo, Tommaso Di	Mönch, Emanuel	Sturm, Jan-Egbert
Fräle, Cecilia	Negri, Valeria	Tallman, Ellis
Fritsche, Ulrich	Nieuwenhuyze, Christophe van	Terai, Akira
Friz, Roberta	Norden, Simon van	Terasvirta, Timo
Giannone, Domenico	Notini, Hilton Hostalacio	Veronese, Giovanni Furio
Giedeman, Daniel	Nunes, Luis C.	Vuletin, Guillermo
Groenen, Patrick J. F.	Nymoen, Ragnar	Wegener, Michael
Gyomai, György	Paradiso, Antonio	Wynne, Mark A.
Hall, Pamela	Paries, Matthieu Darracq	Zanetti, Attilio

OECD AND CIRET NEWS 2013/2

CIRET Conference

Hangzhou, People's Republic of China

9-11 October 2014

The CIRET biennial conference is the meeting place for top economic scientists from all over the world. The conference successfully integrates the theoretical aspects with the perspectives of those who are primarily engaged in applying survey data.

The objectives of the conference are to:

- develop consistent and applicable explanatory patterns for cyclical fluctuations, and;
- overcome the divergence of empirical, theoretical and policy activities.

The focus theme will be on a topical issue in conducting economic surveys.

Hosted by a CIRET member country, this year the conference will be organised by the local co-ordinators in Hangzhou, the capital of Zhejiang Province in Eastern China.

Participants will include experts and researchers working in the field of socio-economic surveys.

Conference papers will be made available to both CIRET members as well as authors and participants at the conference.

Authors are kindly invited to submit a ready-to-publish version of their paper to the joint OECD and CIRET *Journal of Business Cycle Measurement and Analysis* (JBCMA). For further information on paper submissions, visit www.ciret.org/jbcma.

For selected papers from conferences pre-2000 (including the 25th CIRET Conference in Paris, 2000), they can be found in the conference proceedings available at CIRET Publications www.ciret.org/idc/publications.

Aims and scope

The *Journal of Business Cycle Measurement and Analysis* is jointly published by the OECD and CIRET to promote the exchange of knowledge and information on theoretical and operational aspects of economic cycle research, involving both measurement and analysis.

The OECD has developed various cyclical indicators using quantitative and qualitative information from its member countries. Information about the work of OECD in the field of short-term economic statistics may be found at www.oecd.org/std.

CIRET, the Centre for International Research on Economic Tendency Surveys, is a global forum for economists and institutions, which conduct and analyse business and consumer surveys. CIRET holds international conferences every other year. Information about CIRET may be found at www.ciret.org.

The scope of the *Journal of Business Cycle Measurement and Analysis* encompasses:

- Analysis of cyclical fluctuations.
- Business cycle specification, definition and classification.
- Statistical approaches to the development of short-term economic statistics and indicators.
- Business tendency, investment and consumer surveys.
- Use of survey data or cyclical indicators for business cycle analysis.

Researchers in the field of business cycle measurement and analysis are encouraged to submit their papers to the Editor-in-Chief, Michael Graff (editor-in-chief@ciret.org). All papers undergo a double-blind review process.

Under the heading “Report” additional contributions, including topics such as sample design, data collection and validation, dissemination and evaluation of survey results, are published. Reports are reviewed by members of the Editorial Board.

Other correspondence related to the journal may also be addressed to Pierre-Alain PIONNIER, at the OECD (Pierre-alain.PIONNIER@oecd.org).

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Commission takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD Journal

Journal of Business Cycle Measurement and Analysis

The *Journal of Business Cycle Measurement and Analysis* is jointly published by the OECD and the Centre for International Research on Economic Tendency Surveys (CIRET) to promote the exchange of knowledge and information on theoretical and operational aspects of economic cycle research, involving both measurement and analysis.

Contents

General introduction: Short-term forecasting methods, joint issue with *Économie et Prévision*

Hélène Erkel-Rousse and Michael Graff

Short-term forecasting of French GDP growth using dynamic factor models

Marie Bessec and Catherine Doz

New indicators for tracking growth in real time

Troy D. Matheson

Dynamic factor models: A review of the literature

Karim Barhoumi, Olivier Darné and Laurent Ferrara

Constructing a conditional GDP fan chart with an application to French business survey data

Matthieu Cornec

Do business tendency surveys help in forecasting employment? A real-time evidence for Switzerland

Boriss Siliverstovs

Referee acknowledgement

OECD and CIRET News 2013/2

www.oecd.org

www.ciret.org

www.kof.ethz.ch

This issue is published on the OECD iLibrary at: <http://dx.doi.org/10.1787/jbcma-v2013-2-en>.

The OECD iLibrary gathers all OECD books, periodicals and statistical databases. Visit www.oecd-ilibrary.org, and do not hesitate to contact us for more information.

Volume 2013/2

OECD publishing
www.oecd.org/publishing

ISSN 1995-2880
VOLUME 2013
(2 ISSUES)
33 2013 02 1 P

