DOCUMENT

DE TRAVAIL

N° 473

NEW ESTIMATE OF THE MIBA FORECASTING MODEL.

MODELING FIRST-RELEASE GDP USING THE BANQUE DE FRANCE'S MONTHLY BUSINESS SURVEY AND THE "BLOCKING" APPROACH.

Matteo Mogliani, Véronique Brunhes-Lesage, Olivier Darné and Bertrand Pluyaud

January 2014

BANQUE DE FRANCE

EUROSYSTÈME

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

NEW ESTIMATE OF THE MIBA FORECASTING MODEL.

MODELING FIRST-RELEASE GDP USING THE BANQUE DE FRANCE'S MONTHLY BUSINESS SURVEY AND THE "BLOCKING" APPROACH.

Matteo Mogliani, Véronique Brunhes-Lesage, Olivier Darné and Bertrand Pluyaud

January 2014

New estimate of the MIBA forecasting model

Modeling first-release GDP using the Banque de France's Monthly Business Survey and the "blocking" approach

Matteo Mogliani[1]

Véronique Brunhes-Lesage[2]

Olivier Darné[3]

Bertrand Pluyaud[4]

Acknowledgments: We are indebted to Pierre Sicsic, Marie Diron, Thomas Ferrière and the participants in seminars at the Banque de France for helpful comments and discussions. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Banque de France.

[1] *Corresponding author.* Banque de France, DGEI - DCPM - DIACONJ. Email: matteo.mogliani@banque-france.fr

[2] Banque de France, DGS - DSMF - SESOF. Email: veronique.brunhes-lesage@banque-france.fr

[3] University of Nantes and Banque de France. Email: olivier.darne@univ-nantes.fr

[4] Banque de France, DGEI - DEMS - SEPS. Email: bertrand.pluyaud@banque-france.fr

Abstract: This paper introduces the new Monthly Index of Business Activity (MIBA) model of the Banque de France for forecasting France's GDP. As the previous versions, the model relies exclusively on data from the monthly business survey (EMC) conducted by the Banque de France. However, several major changes have been implemented in the present version, such as the shift from a model based on factors to a model based on survey opinions, the explicit targeting of first-release GDP, and the use of the "blocking" approach to deal with mixed frequencies and missing observations. The selected monthly equations are consistent with the time frame of real-time forecasting exercises: the first month equation is dominated by data on expected evolution of the economic activity across the coincident quarter, while for the second and third month equations data on observed economic activity become more important and forward-looking information is progressively discarded. Finally, out-of-sample results suggest that the new MIBA model broadly outperforms several competing models, such as the previous version of MIBA and models based on alternative specifications.

Keywords: GDP nowcasting; Real-time data; Mixed-frequency data.
JEL classification: C22, C52, C53, E37.

Résumé : Cette étude présente des nouvelles équations d'étalonnage pour l'Indicateur Synthétique Mensuel d'Activité (ISMA) de la Banque de France, modèle de prévision de la croissance trimestrielle du PIB de la France. Comme les versions précédentes, le modèle repose exclusivement sur les données de l'enquête mensuelle de conjoncture (EMC) de la Banque de France. Toutefois, plusieurs changements majeurs ont été mis en oeuvre dans cette nouvelle version, comme, par exemple, le passage d'un modèle basé sur des facteurs à un modèle basé sur les soldes d'enquête, le ciblage explicite du PIB publié par l'INSEE dans les Premiers Résultats des comptes trimestriels, et l'utilisation de la méthode du "blocking" pour traiter le problème des fréquences mixtes et des observations manquantes. Les équations mensuelles sélectionnées sont compatibles avec le calendrier de prévision des exercices en temps réel : l'équation du premier mois dépend principalement des données relatives aux perspectives de l'activité économique pour le trimestre coïncident, alors que pour les étalonnages du deuxième et troisième mois l'information contemporaine sur l'activité économique devient plus importante et celle relative aux perspectives est progressivement négligée. Enfin, les résultats hors-échantillon suggèrent que le nouveau modèle ISMA fournit des prévisions plus précises que celles obtenues à partir de plusieurs modèles concurrents, comme, par exemple, la version précédente de l'ISMA ou des modèles basés sur des spécifications alternatives.

Mots-Clés : Prévisions du PIB ; Données en temps réel ; Données à fréquence-mixte.
Classification JEL : C22, C52, C53, E37.

1 Introduction

This paper introduces a new version of the model used to determine the Monthly Index of Business Activity (MIBA; in French, *Indicateur Synthétique Mensuel d'Activité*, ISMA) published in the Overview of the Banque de France monthly business survey (EMC, *Enquête Mensuelle de Conjoncture*) since January 2000. This model is designed for predicting the quarterly growth rate of France's GDP at the current quarter horizon (*nowcasting*), based exclusively on information stemming from the EMC on manufacturing industry and services. Forecasts are updated monthly on the basis of new EMC data inflows.

Forecasting models relying on the predictive information stemming from qualitative surveys are commonly used to forecast GDP growth rates. Business survey data ("soft data"), which usually display high correlation with GDP growth, have the main advantage of being rapidly available (usually at the end of the month covered by the survey), while monthly macroeconomic indicators, such as the industrial production index (IPI) or the households' consumption of goods ("hard data"), have a longer publication lag (40 days, for instance, in the case of IPI). Data from surveys also have the advantage of being mostly unrevised.

The previous version of the MIBA model (Darné and Brunhes-Lesage, 2007) was estimated before the 2008-2009 economic crisis. Three distinct equations were simultaneously estimated each month across the quarter of interest. The MIBA forecast was, in practice, selected among the predictions resulting from these three equations, which were mainly based on the first factor from a principal component analysis (PCA) carried out on balances of opinion on the whole industry. The predictive content of other variables, while not insignificant, was less substantial. In addition to the first factor on industry, the first equation included other factors from the same PCA. The second equation included PCA factors on sectoral decomposition of industry and the third equation included the balance of opinion on past activity in services. Accounting for a survey opinion on services was one of the main innovations of the model estimated in 2007, along with the implementation of an algorithm for automatic model selection (GETS, see Krolzig and Hendry, 2001).

The rise in recent years of a systematic upward predictive bias in the MIBA forecasts represents one of the main reasons underlying a re-assessment of the model. Several innovations, presented in the remainder of this article, are also introduced with the aim of improving the predictive performance of the model, such as: *i)* reconsidering the use of PCA factors as explanatory variables, *ii)* investigating the role played by real-time data in forecasting explicitly first-release GDP values, and *iii)* implementing the "blocking" approach in order to address the issue of missing observations in the quarter to be forecast by taking into account the information available at the time of the forecast exercise. Finally, unlike the 2007 version, the new MIBA model does not rely on the service sector survey, as no balances of opinion were selected through the automatic model selection procedure.

Estimation results suggest that the selected monthly equations are consistent with the time frame of the real-time forecast exercise. For the first month equation, only partial information on the coincident quarter is available, so that data on expected evolution of the economic activity across the quarter dominates data on coincident economic activity. However, for the second and third month equations, salient information on the coincident quarter becomes available, so that data on observed economic activity outpace forward-looking indicators. Out-of-sample evaluation suggests that the new MIBA model is reasonably well accurate over a large evaluation period, in particular after the 2008-2009 economic crisis. Indeed, the model is not able to capture the second consecutive strong contraction of GDP in Q1 2009, but it can track extremely well the recovery and the moderate growth observed in 2012. The predictive accuracy of the new MIBA model is tested against several alternative models, such as the previous version of MIBA and a model based on final-release GDP. Benchmark exercises reveal that our model broadly outperforms the competing specifications, mainly over the more recent period. In particular, the predictive gain over the previous version of the MIBA model ranges between 5% and 50%, according to MAE, depending on the evaluation period considered. Similarly, the predictive gain over a model based on final-release GDP can reach 50%, when forecasts are evaluated with respect to first-release GDP data. As expected, when final-release data are targeted, results suggest that the model based on final-release GDP slightly outperforms the new MIBA.

The remainder of the paper is organized as follows. Section 2 discusses in greater detail the main features of the new MIBA model. In Section 3, we describe the data, the empirical strategy and the implemented model selection approach. Section 4 presents and discusses the estimation results. Section 5 reports forecast evaluation, while Section 6 compares the predictive accuracy of the new MIBA model with that of alternative specifications. Finally, Section 7 concludes.

2 Main features of the new MIBA model

2.1 From a model based on the business sentiment indicator to a model based on survey balances

The previous version of the MIBA model was mainly based on the econometric relationship between GDP growth rates and the first PCA factor computed on the monthly survey on industry. By and large, this approach consisted in relating GDP to the Banque de France's Business Sentiment Indicator in industry (BSI; in French *Indicateur du Climat des Affaires*, ICA), because the first PCA factor broadly coincides with this indicator.[1]

However, this empirical strategy has been proven to be progressively disappointing, because the forecasting model systematically displayed an upward predictive bias requiring a manual correction at every monthly forecasting exercise. Preliminary analysis suggests that the bias could be induced, at least in part, by the presence in the first PCA factor of the whole set of balances of opinion, including those apparently not relevant for modeling and forecasting GDP. This factor is indeed a synthetic indicator encompassing business managers' opinions on output, sales and orders, inventories and prices, so that only a sub-set of this information, based

[1] The Banque de France's Business Sentiment Indicator in industry is proportional to the first PCA factor computed over the 14 balances of opinion constituting the core of the monthly business survey on industry. However, series are smoothed (three months moving average) prior to the implementation of the PCA analysis. In addition, this factor is normalized to have mean 100 and standard deviation 10.

specifically on the productive activity of firms, could be actually useful for modeling GDP. We test this assumption by comparing a regression model of GDP based on the first PCA factor in industry, *i.e.*, an equation very close to the set of three equations constituting the former MIBA, to a second model in which the factor is broken down into several components, each including a sub-set of relatively homogenous indicators (*e.g.*, order books, production, prices, inventories and the capacity utilization rate) combined using the weights obtained from the same PCA (eigenvectors). The results of this test, presented in the Appendix (Table A1), confirm the intuition: only the group of indicators that include the survey balances on productive activity (*Production*) appears statistically significant. Further, the inclusion of additional groups of indicators leads to a substantial deterioration in the out-of-sample performance of the model (Table A2 in the Appendix).

A visual representation of the BSI and its *Production* sub-component is helpful in understanding this finding (Figure A1). Before the 2000s, the BSI seems to underestimate the economic activity, while since the start of the decade it has tended to overestimate it. Although slight, this gap appears remarkably upward trending, and could explain the rise of an upward predictive bias in the MIBA forecasts.

2.2 First vs final GDP Estimates

Unlike surveys, GDP undergoes numerous revisions, because first releases are based on partial information, available at the time of the publication of first estimates, which is progressively updated for the construction of final GDP releases. Hence, the French National Statistical Agency (INSEE) releases an initial estimate of the quarterly accounts 45 days after the end of the quarter, "provisional" annual accounts for the year n (on quarterly frequency) in May $n+1$, "semi-final" accounts in May $n+2$, and, lastly, "final" annual accounts in May $n+3$. As suggested by Croushore (2011), GDP revisions can have a significant impact on forecasts. Therefore, the optimal forecasting model should take into account this feature. Revisions affect the input of the forecasting model, and the revised data may, in turn, modify the coefficients as well as the specification of the model itself.[2] Forecasts are therefore inevitably affected by the modeling decisions undertaken *ex ante* by the forecaster, involving both the inputs and the target of the forecasting model. What GDP series should therefore be used and predicted? First-release or final-release? From the point of view of economic analysis, final estimates represent a reliable picture of the true macroeconomic outlook and should be hence preferred. However, there is a large evidence on the effect of earlier announcements of economic news on markets, so that the predictive accuracy of a forecasting model is often assessed when the first official data are released. This is an issue for the real-time forecaster, such as the MIBA forecaster. In his

[2]We can discern between two kinds of revisions: adding news (information) or reducing the statistical noise (of first estimates). If revisions contain news, first estimates are optimal forecasts of final estimates, but the revisions are not predictable. However, if revisions reduce statistical noise, first estimates are the difference between final estimates and a measurement error, and the revisions are hence predictable (Croushore, 2011).

conclusion, Croushore (2011, p.96) strongly argues in favor of first-release data: "If you want to analyze policy or forecasts you *must* use real-time data, or your results are irrelevant".[3]

In this paper, we therefore choose to forecast first-release GDP. For this purpose, a time series of first GDP estimates is reconstructed from Q1 1992 on by copying first-releases from paper archives at the INSEE library. Let us assume the existence over time of a constant relationship between EMC survey data and the data used by INSEE to compute first GDP estimates, and then let us also assume the existence of a second relationship between EMC survey data and the data used to compute final GDP estimates. If these assumptions hold, there might be a break in the residuals of a regression of final GDP estimates on EMC data, for the period covering the "non-final" results. In practice, non-zero mean errors, which may be partly explained by this feature, have been actually observed in MIBA forecasts over the recent period. A priori, we should address this issue by using first-release GDP data, although in practice we have no clear-cut evidence of the existence of a constant relationship between first-release data and monthly business survey data. Indeed, first-release data may be affected by methodological changes and, in particular, by adjustments made by INSEE when base years are changed.

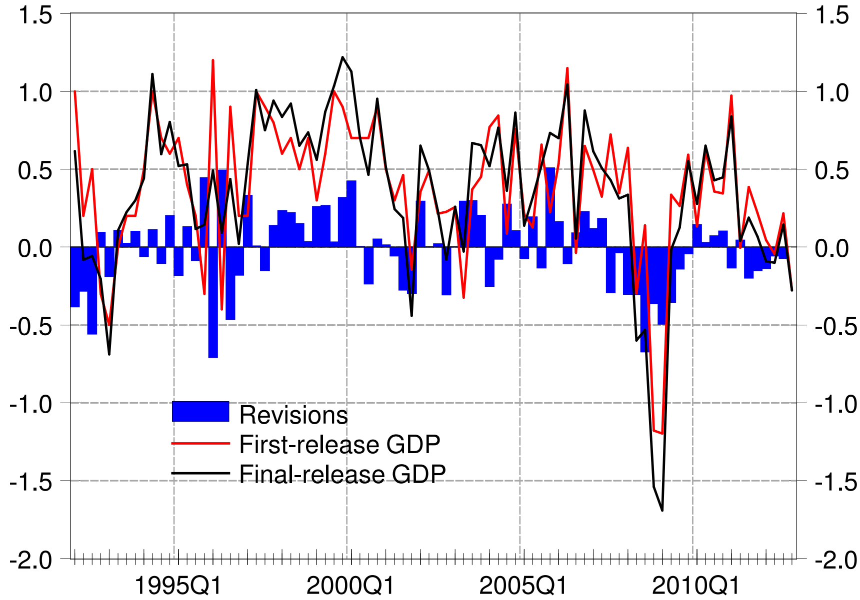
Recently, using French data, Minodier (2010) compared forecasting models using either final or first GDP estimates as the targeted variable and INSEE survey data as explanatory variables. She shows that the predictive accuracy is significantly improved by the use of first-release data, when the official target is the first-release GDP. More specifically, for the Q1 1992 - Q4 2005 period, Minodier (2010) shows that a regression model involving first-release GDP and INSEE surveys on industry leads to an in-sample performance that is statistically equivalent to a regression model involving instead revised GDP series (R^2 of 0.63 and 0.66, respectively; RMSE of 0.23 and 0.20 respectively, therefore higher RMSE with first-release data). However, based on a real-time forecasting exercise targeting the first-release GDP values released between Q1 2000 and Q4 2008, out-of-sample results appear to favor the model using first-release GDP (Table 3 p.17, out-of-sample RMSE of 0.26 with the equation based on first-release data, and 0.31 with the equation based on revised GDP). Similar results were obtained by Koenig et al. (2003) and Clements and Galvão (2013) on US data, in a similar context of real-time forecasting exercises.[4]

Out-of-sample results reported in the present study show that models based on first-release GDP lead to more accurate forecasts: for the period spanning from Q2 2009 to Q4 2012, RMSE are between 0.14 and 0.23 for the models using first-release data, and between 0.23 and 0.33 for the models using revised GDP values, so that predictive gains amount to about 30-40% (see Section 6.4, Panel D of Table 4). Finally, the inclusion of non-industry survey data in the forecasting models appears to depend on the targeted GDP. For instance, Bessec (2010, p.84

[3]On page 90 he writes: "No forecaster today is going to modify her forecasts to account for the possibility five years hence; nor should anyone do so. Thus, evaluations of forecasts should usually focus on early releases of the data, or the last vintage of the data after a forecast is made but prior to a benchmark revision that changes base years or redefines variables. Still, most evaluations of forecast exercises are based on latest-available data for convenience, even though they may provide a distorted view of forecastability".

[4]With respect to the theory, the literature focuses on the predictive bias arising from regression models when the target is first-release GDP (or, more generally, a revised series) while the model is based on latest available GDP series (a mix of fully revised and unrevised observations).

Figure 1: Revisions, first-release and final-release GDP (Q1 1992 - Q4 2012)



and p.91) points out that surveys in services and construction are useful for forecasting final GDP values, but not first-release values.

2.2.1 A comment on GDP revisions

Even if the choice was made here to consider first-release GDP, we do not intend to deny any interest to the prediction of final GDP estimates. Related contributions to the present work could hence be either to focus on forecasting revised GDP values, or alternately considering the predictability of GDP revisions (see Figure 1). Although addressing these issues is clearly beyond the scope of this paper, it is worth considering a few remarks. GDP is revised for various reasons, such as seasonal adjustment refining, progressive inclusion in the National Accounts of indicators previously not available but only extrapolated, and changes in the base-year, which also represents the occasion for a more general reshaping of the accounting methodology. Assessing the predictability of GDP revisions, as well as tracking the factors affecting revisions, may hence be a very complicated exercise.

On the one hand, a simple forecast-rationality test based on a regression of GDP revisions, computed as final-release GDP *minus* first-release GDP, on first-release GDP (see, *e.g.*, Faust et al., 2005) points to a strong rejection of the hypothesis of predictability of revisions ($H0: \alpha = \beta = 0$ is rejected in the left Panel of Table A3 in the Appendix). Hence, INSEE seems to efficiently incorporate all the available information into preliminary estimates of GDP, since revisions appear unbiased ($\alpha = 0$) and uncorrelated with first-release values ($\beta = 0$). However, on the other hand, estimated parameters from a regression of GDP revisions on final-release

GDP point to a statistically significant positive correlation between the two variables ($\beta = 0.25$ in the right Panel of Table A3 in the Appendix). In particular, revisions have been significantly positive, by 0.12 pp on average, over the 1997Q1-2001Q1 period, marked by a strong GDP growth and a large diffusion of new technologies of information and communication, but very negative over the Great Recession episode spanning from 2008Q2 to 2009Q3 (-0.39 pp on average). A source of explanation could be found in the fact that the models used in National Accounts to extrapolate indicators unavailable at the time of early GDP estimates, were unable to capture these unexpected large swings.

2.3 Dealing with missing monthly observations

Due to the quarterly frequency of GDP series, monthly business indicators must be converted into quarterly data. An intuitive way to achieve this goal is to time-aggregate monthly data through a simple average. However, an issue arises when survey data covering the quarter to be predicted are partially available at the time of the forecasting exercise.[5]

An alternative method, commonly used to address this issue, is to fill missing observations by forecasting explanatory variables through the so-called "bridge" models (autoregressive models are often preferred for their simplicity).[6]

However, besides the inconvenience of having to forecast explanatory variables, quarterly averages of monthly indicators have the additional drawback of representing a non-optimal treatment of the available information. For instance, let us consider a forward-looking survey indicator, such as expected production or order books. The information available at the beginning of the quarter may be more useful for forecasting GDP over a given quarter, rather than a quarterly average which would include some information about the following quarter. The econometric approach implemented in the present paper allows us to address these two issues. This approach, known as "blocking", described and implemented in the case of France by Dubois and Michaux (2006), Bessec (2010) and Minodier (2010), consists in splitting the high frequency information into multiple low frequency time series, which means, in our case, distributing the monthly survey data into three quarterly series: the first one collects observations from the first months ($m1$) of each quarter (January, April, July and October), the second one ($m2$) collects observations from the second months (February, May, August and November), while the last one ($m3$) collects the remaining observations from the third months (March, June, September and December). This approach allows the forecaster to take into account observations that are effectively available at the time of the forecasting exercise, with no need to extrapolate the missing information: for example, for a Q1 forecast based on the February survey, coincident $m1$ and $m2$ series are available for modeling purposes, while $m3$ series can only be used with a lag of at least one quarter (series ending in December of the previous year). This approach also enables the forecaster to use the most relevant information conditional on the month in which

[5]In the economic literature, this problem is commonly called "ragged edge data problem", as popularized by Wallis (1986).

[6]Three "bridge" methods were implemented in the previous version of the MIBA model: extrapolation using autoregressive models, simple average of the observations available over the quarter, and moving average over a window of the last three observations.

the forecast is conducted. It is possible to use different equations for each of the three months of the quarter, with the expected result that forward-looking balances of opinion are mainly selected at the beginning of the quarter, while balances referring to past economic activity are preferred at the end of the quarter. That said, this method may have a normative drawback if the equations are different: it would be more difficult to justify forecast revisions across the quarter when a different equation, rather than a single and identical equation, is used each month, as was the case hitherto for the MIBA model. However, in Section 4 we shall see that the selected equations are sequentially quite similar, so that it is relatively easy to track and explain the source of revisions. This point will be further explored in Section 5.2.

3 Estimating the new MIBA model

3.1 Model selection

In this study, the set of explanatory variables entering the forecasting equations are selected in an automatic fashion, following the "general-to-specific" approach (GETS) popularized by Krolzig and Hendry (2001) (see also Hoover and Perez, 1999). Starting from a general unconstrained model that captures the underlying features of data, this approach consists in discarding sequentially those variables appearing statistically non-significant, provided that a certain number of specification tests are passed at each step. The aim is to obtain a model that is both adequately specified, *i.e.*, no relevant explanatory variable are omitted, and parsimonious, *i.e.*, only redundant variables are excluded. Whenever this approach leads to several competing models, very close in terms of in-sample fit, the final model is selected on the basis of the Schwarz selection criterion. It is nonetheless possible to keep these competing models in order to test their predictive accuracy. This approach allows the econometrician to automatically test for a large number of potentially relevant models in a very short time, with no need to estimate manually all the possible models (2^N, where N is the number of explanatory variables). It is therefore possible to reduce the uncertainty surrounding the choice of regressors and their lags. Latest computational development in this automatic approach can allow the researcher to deal with a large number of explanatory variables, compared to the size of the sample, while this number was very limited at the time of the estimate of the previous version of the MIBA model. In the present study, the GETS approach is implemented by using the *Autometrics* algorithm (Doornik, 2009; Hendry and Doornik, 2009).

However, the automatic selection approach is here backed by the "expert opinion" of the forecaster. Indeed, judgmental arguments can be advocated in order to fine-tune the automatically selected equations, especially in case the outcome of the *Autometrics* algorithm does not reflect reasonable priors on the stability and the interpretation of the selected models.[7] With this aim in mind, we proceed as follows. *First*, variables entering the equations are supposed to display some parameter stability and statistical significance over time. Model selection is

[7] As pointed out by Ciccone and Jarocinski (2010), *inter alia*, the variables chosen by the automatic selection procedure can be very sensitive to the particular time span of the sample.

Table 1: Balances of opinion on manufacturing sector and market services

Balance of opinion	Sector	Question	Reference period
EVPRO	Manufacturing	Changes in production	M/M-1
EVLIV	Manufacturing	Changes in deliveries	M/M-1
EVCOM	Manufacturing	Changes in overall orders	M/M-1
EVCOME	Manufacturing	Changes in foreign orders	M/M-1
EVPRMP	Manufacturing	Changes in commodity prices	M/M-1
EVPRPF	Manufacturing	Changes in prices of finished goods	M/M-1
EVSTPF	Manufacturing	Changes in inventories of finished goods	M/M-1
ETCC	Manufacturing	Order books	M/"norm"
STPF	Manufacturing	Inventories of finished goods	M/"norm"
STMP	Manufacturing	Inventories of commodities	M/"norm"
CSEMA	Manufacturing	Weekly order levels	M/M-1
TUC	Manufacturing	Average capacity utilisation rate	M/M-1
PREVPRO	Manufacturing	Expected changes in production	M+1/M
PREVSTPF	Manufacturing	Expected changes in inventories of finished goods	M+1/M
EVACT	Services	Changes in activity	M/M-1
EVPRIX	Services	Changes in prices	M/M-1
EVEFF	Services	Changes in staff levels	M/M-1
NIVTRES	Services	Levels of cash flows	M
PREVACT	Services	Expected changes in activity	M+1/M
PREVPRIX	Services	Expected changes in prices	M+1/M
PREVEFF	Services	Expected changes in staff levels	M+1/M

Notes: M denotes the current month. M/M-1 denotes the current month compared to the previous month. M+1/M denotes expectations over the next month compared to the current month. M/"norm" denotes the current month compared to normal level.

initially conducted over a sample spanning from Q1 1992 to Q4 2012. The obtained equations are then re-estimated over a shorter sample that excludes the crisis (Q1 1992 - Q4 2007), in order to discard those variables selected because of a good fit over the crisis period but resulting not statistically significant over the pre-crisis period.[8] Once non-significant variables are discarded, the resulting equation is again estimated over the Q1 1992 - Q4 2012 sample. Similarly, non-significant variables are again discarded, in order to obtain the final equation. *Second*, models are supposed to embed the most recent information available on the current quarter, whenever possible. For instance, the second-month equation is supposed to include balances of opinion on the second month of the quarter (according to the *blocking* approach), possibly in addition to past information. *Third*, models are supposed to not include indicators displaying odd and unexpected signs, or linear combinations of them. The way these conditions affected the selection of final models is described in Section 4.

[8]Since the crisis, it has been frequently noted that models selected through automatic procedures tend to include linear combinations of variables that lead to a better fit over few observations (Q4 2008 and Q1 2009 in particular), where GDP variations have been very strong and where forecasting models have often reported substantial forecast errors, but that are difficult to interpret economically.

3.2 Survey data

The full dataset included in the first step of the selection process is drawn from the manufacturing industry and services sections of the EMC.[9] Survey data used in our equations are expressed on a monthly frequency and have been collected since 1987 and 1989 for industry and services, respectively. Fourteen balances of opinion relate to total manufacturing industry, which covers the four main manufacturing sub-sectors according to the INSEE's NAF rev.2 nomenclature (agri-food industry, capital goods, transport equipment and other manufactured goods; only the low-weighted sub-sector "Manufacture of coke and refined petroleum products" is not covered by the survey). Seven balances of opinion are selected for total services. The questions asked on industry and services refer to the recent past (for example, changes in production, deliveries or activity compared to the previous month) as well as to business managers' expectations (changes in production expected in the following month). Overall, a total of 21 $(14 + 7)$ balances can be used in the analysis. See Table 1 for more details.[10]

As the number of survey balances described above is quite high, principal component analysis (PCA) is also implemented to resume the largest amount of information in a limited number of factors, as was already the case in the previous version of the MIBA model (see Darné and Brunhes-Lesage, 2007). PCA is implemented distinctly to balances of opinion on both the manufacturing industry and services sections of the monthly business survey. Five and three factors are retained, respectively, for industry and services, covering 70% to 90% of the variance in each survey. PCA is also implemented by sub-sectors, so that synthetic information stemming from each sub-sector of the manufacturing industry and services is available. Overall, the number of usable factors in this second analysis amounts to 8 $(5 + 3)$ for total industry and services and to 50 $(5 \times 4 + 3 \times 10)$ for sub-sectors.

3.3 Specifications

The estimation period spans from Q1 1992 to Q4 2012, *i.e.*, 84 observations. The targeted variable is first-release GDP growth rate (Y_t). General unrestricted forecasting equations relate GDP to its lagged value (Y_{t-1}), balances of opinion or factors (contemporaneous and lagged, aggregated or non-aggregated sectors) on the reference month ($X_t^{(mj)}$, with $j = 1, 2, 3$), a constant, and an error term, *i.e.*:

[9] Survey data on the construction sector and on retail/wholesale trade are also available from the Banque de France. However, the former are only available on a quarterly basis, while the latter start in 1993 and in 1996, respectively. Further, data on wholesale trade are only available on a quarterly basis. Yet, the use of trade indicators may appear not suitable for modeling GDP from a supply-side perspective, as in the case of the MIBA model.

[10] As for the manufacturing sector, four additional questions are asked in the survey, but have shorter historical records than the fourteen questions presented in Table 1. As for the market services sector, survey data are collected over ten sub-sectors of activity: hotels, temporary employment, computer engineering, technical engineering, car rental, business and management counseling services, agencies, advertising, cleaning services, car repair and road freight transport.

- Month 1 Equation (M1)

$$Y_t = \alpha + \beta Y_{t-1} + \sum_{k=1}^{K} \left(\delta_{1,k} X_{k,t}^{(m1)} + \delta_{2,k} X_{k,t-1}^{(m3)} + \delta_{3,k} X_{k,t-1}^{(m2)} + \delta_{4,k} X_{k,t-1}^{(m1)} \right) + \varepsilon_{1,t} \tag{1}$$

- Month 2 Equation (M2)

$$Y_t = \alpha + \beta Y_{t-1} + \sum_{k=1}^{K} \left(\delta_{1,k} X_{k,t}^{(m2)} + \delta_{2,k} X_{k,t}^{(m1)} + \delta_{3,k} X_{k,t-1}^{(m3)} + \delta_{4,k} X_{k,t-1}^{(m2)} \right) + \varepsilon_{2,t} \tag{2}$$

- Month 3 Equation (M3)

$$Y_t = \alpha + \beta Y_{t-1} + \sum_{k=1}^{K} \left(\delta_{1,k} X_{k,t}^{(m3)} + \delta_{2,k} X_{k,t}^{(m2)} + \delta_{3,k} X_{k,t}^{(m1)} + \delta_{4,k} X_{k,t-1}^{(m3)} \right) + \varepsilon_{3,t} \tag{3}$$

where K represents the number of balances of opinion or factors.[11] The inclusion of a lag of GDP as a regressor may rise two types of issues, especially in a context of real-time forecasting exercises. *First*, the 45-day lag in the publication of GDP means that the lagged value is not known when the first forecast of a quarter is made. *Second*, the presence of an autoregressive term, with, potentially, a negative and statistically significant coefficient, introduces a correction mechanism that operates systematically in each forecasting exercise. However, in practice, we may suppose that this term essentially reflects the way the growth rate reverts to a "norm" after a one-off shock to GDP (for example, a sharp falloff in construction production due to adverse weather conditions, to be offset in the next quarter by a recovery in construction output) and it may then be expected that a quarter-on-quarter correction would not be relevant in the event of a strong increase or a sharp drop in GDP related to a cyclical phase of acceleration or slowdown of the economic activity (a recession, for example).[12] A simple solution to the first issue is presented in Section 5.1. The second issue may be arbitrarily addressed by dropping the lag of GDP from equations, but this would lead mechanically to a strong autocorrelation of residuals. To check the relevance of this problem, MAX(1) (Moving Average with eXogenous inputs) models, *i.e.*, regressions involving a moving average term of order 1 (ε_{t-1}), as well as models that incorporate a Cochrane-Orcutt type correction for the first-order autocorrelation of residuals, were studied. However, out-of-sample performances of these models have proven to be generally moderate compared to models including a lag of GDP. As for the surveys, we assume that balances of opinion are unrevised. This means that, unlike with GDP, we should not make the distinction between first-release data and revised data, so that the most recent series available can be used.[13]

[11] Dubois and Michaux (2006) only consider coincident balances of opinion, as well as month-on-month changes in balances over the last three months.

[12] Forecasting models based on final GDP data does not bring out a statistically significant autoregressive coefficient. See Sections 6.3 and 6.4 for a discussion.

[13] In practice, surveys may be revised in the month following their first release, in order to take into account late responses and, to a lesser extent, seasonal adjustments. However, these revisions are usually minor.

Table 2: New MIBA estimates (Q1 1992 - Q4 2012)

M1			M2			M3		
Variable	Coefficient	t-stat	Variable	Coefficient	t-stat	Variable	Coefficient	t-stat
Constant	16.47	3.76	*Constant*	8.19	1.92	*Constant*	9.11	2.09
	(4.38)	[0.00]		(4.25)	[0.06]		(4.37)	[0.04]
Y_{t-1}	-0.41	-5.00	Y_{t-1}	-0.37	-4.90	Y_{t-1}	-0.39	-4.82
	(0.08)	[0.00]		(0.07)	[0.00]		(0.08)	[0.00]
$\text{EVLIV}_t^{(m1)}$	1.91	3.45	$\text{EVLIV}_t^{(m2)}$	2.14	5.32	$\text{EVLIV}_t^{(m3)}$	0.94	2.37
	(0.55)	[0.00]		(0.40)	[0.00]		(0.39)	[0.02]
$\text{PREVPRO}_t^{(m1)}$	4.09	5.84	$\text{PREVPRO}_t^{(m2)}$	2.17	3.27	$\text{EVLIV}_t^{(m2)}$	2.22	5.35
	(0.70)	[0.00]		(0.66)	[0.00]		(0.41)	[0.00]
			$\text{EVLIV}_t^{(m1)}$	1.93	3.84	$\text{EVLIV}_t^{(m1)}$	2.54	5.58
				(0.50)	[0.00]		(0.46)	[0.00]
$dummy_{09Q1}$	-103.2	-3.64	$dummy_{09Q1}$	-82.7	-3.09	$dummy_{09Q1}$	-96.4	-3.52
	(28.4)	[0.00]		(26.8)	[0.00]		(27.4)	[0.00]
Adj-R^2	0.67		Adj-R^2	0.71		Adj-R^2	0.69	
σ_ϵ	25.73		σ_ϵ	23.92		σ_ϵ	24.65	
SIC	9.54		SIC	9.43		SIC	0.86	
Normality	3.24	[0.20]	Normality	0.99	[0.61]	Normality	0.61	[0.73]
AR(4)	0.59	[0.67]	AR(4)	3.24	[0.02]	AR(4)	1.56	[0.19]
Het	0.42	[0.79]	Het	0.25	[0.94]	Het	0.29	[0.91]

Notes: GDP growth is measured in basis points. Standard errors in parentheses. p-values in brackets. σ_ϵ is the regression standard error. SIC is the Schwarz information criterion. "Normality" denotes the Bera-Jarque test for residual normal distribution. AR(4) denotes the Breusch-Godfrey test for residual serial correlation up to order $p = 4$. "Het" denotes the Breusch-Pagan-Godfrey test for heteroskedasticity.

4 Results

This section presents the new MIBA equations, one for each month of the quarter, selected according to their in-sample statistical properties and out-of-sample performance (see Table 2 and Figures 2, 3 and 4). Final specifications are chosen by comparing several alternative specifications, *i.e.*, equations incorporating either balances of opinion or factors built on balances of opinion. In general, the former have proven to outperform the competing models, and they are therefore selected and presented in greater detail hereunder. However, a slightly better specification was obtained for the M1 equation by using factors issued from a PCA on the manufacturing sector. We nevertheless preferred to retain the specification involving the balances, in order to keep some degree of homogeneity across specifications (*i.e.*, with respect to M2 and M3 equations), as well as to avoid the problem of interpreting the results obtained through a factor model. After selecting the models, we observed that the introduction of a dummy variable for Q1 2009 (taking value 1 at Q1 2009, and 0 elsewhere) proved to be necessary to neutralize parameter instability, or more specifically, a jump in the constant term and the coefficient of lagged GDP, observed from the first quarter of 2009 on. This date coincides with the second consecutive quarter of deep recession in France, which is usually poorly predicted by our equations. While it has little effect on the evaluation of forecasts over the entire evaluation period, the introduction of this dummy variable has an impact on equations through a stabilization of estimated coefficients, and therefore necessarily on forecasts since Q2 2009. We shall discuss the stability of coefficients below.

Figure 2: M1 equation: in-sample fit and residuals

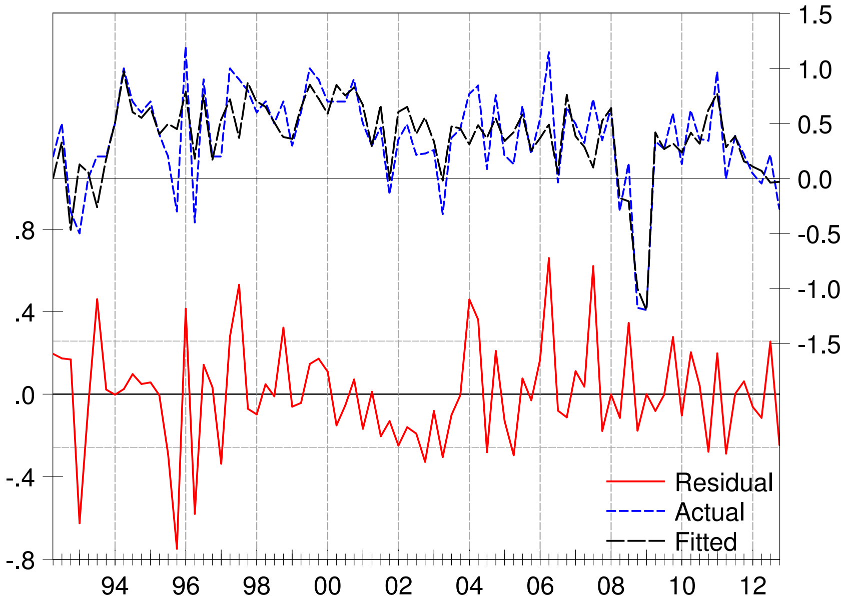
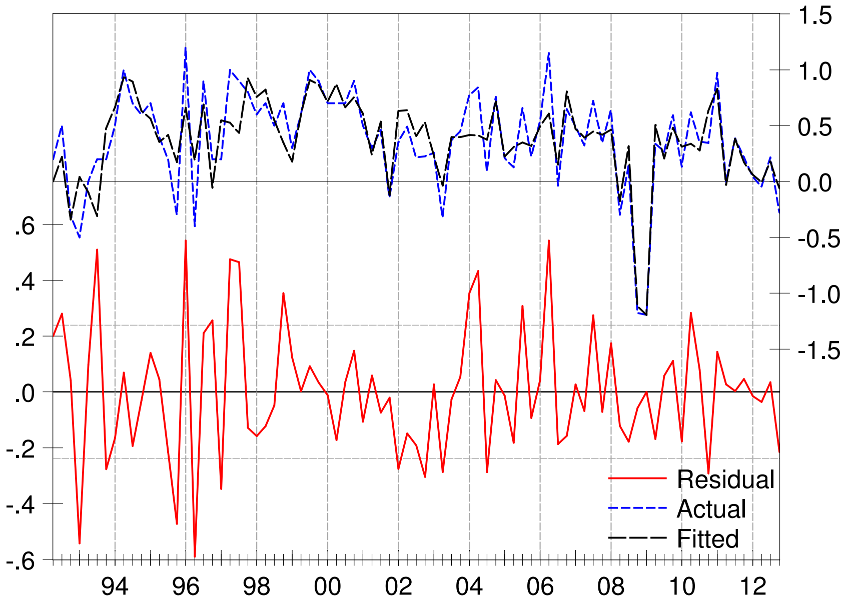


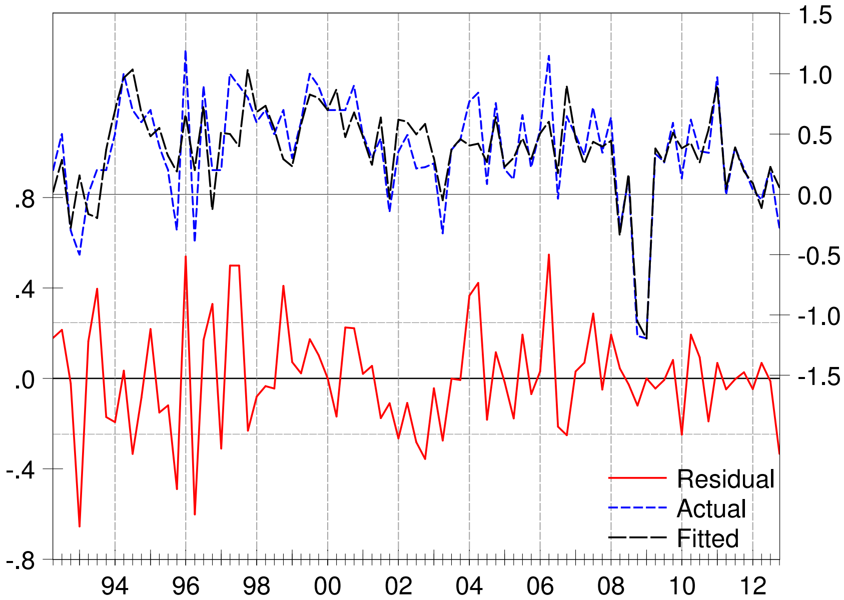
Figure 3: M2 equation: in-sample fit and residuals



4.1 Month 1 equation (M1)

In the forecasting equation associated with the first month of the quarter (M1), GDP growth depends on its lagged term (Y_{t-1}), the balance on changes in deliveries in the first month of

Figure 4: M3 equation: in-sample fit and residuals



the quarter ($\text{EVLIV}_t^{(m1)}$), and the balance on expected changes in production in the first month ($\text{PREVPRO}_t^{(m1)}$) (see below for the issue of lagged GDP, which is unknown at the end of the first month of the quarter). This equation is therefore partially forward-looking (expected changes in production), which is consistent with the time frame of the estimation (first month of each quarter). The coefficient of the expected changes in production being twice as large as that of changes in deliveries, we can therefore interpret the equation as a simple average of deliveries of the first month and projected deliveries for the next two months. The automatic procedure initially selected two additional balances: changes in deliveries in the first month of the previous quarter ($\text{EVLIV}_{t-1}^{(m1)}$) and changes in overall orders in the first month of the previous quarter ($\text{EVCOM}_{t-1}^{(m1)}$). The two variables being highly pairwise correlated (0.94), the size and sign of their estimated coefficients (-0.03 and +0.04, respectively) suggest that a linear combination of these balances was actually selected by the algorithm. Indeed, the exclusion of one balance implied the fall in the statistical significance of the other, which means that their individual contribution to the forecasting equation is negligible. For a matter of parsimony, we hence excluded these two balances from the final M1 equation.

4.2 Month 2 equation (M2)

The forecasting equation associated with the second month of the quarter (M2) includes a lagged term for GDP, the balances on changes in deliveries in the first and second month and expected changes in production in the second month. This equation presents an obvious similarity with the previous equation: negative autoregressive term and the same balances of opinion. Estimated coefficients are very close to each other (approximately 2 in magnitude), as well as close to those of the previous equation, if we assume that the balance on expected changes in production in the first month, entering the M1 equation with an estimated coefficient around 4, is a proxy for the balances on changes in deliveries in the second month and expected changes in production in the second month, both entering the M2 equation with estimated coefficients around 2. As in the case of the M1 equation, the automatic procedure initially selected two additional balances: average capacity utilization rate in the second month of the quarter ($\text{TUC}_t^{(m2)}$) and average capacity utilization rate in the second month of the previous quarter ($\text{TUC}_{t-1}^{(m2)}$). Given the statistical properties of this series (a deep through in correspondence to the Great Recession episode), its first difference is obviously selected by the algorithm in order to accommodate the large swings in GDP growth observed between 2008 and 2009 (estimated level coefficients are +0.10 and -0.08, respectively). Again, for a matter of parsimony, we excluded these two balances from the final M2 equation.

4.3 Month 3 equation (M3)

Finally, the forecasting equation associated with the third month of the quarter (M3) includes the balance on changes in deliveries in the three months, in addition to the lagged value of GDP. It is worth noticing that the estimated coefficient for the third month is roughly half the size of those for the first two months. This result suggests that survey data collected over the third month of the quarter provides much less valuable information about the current economic activity than survey data collected over the previous months. In fact, balances for the last month of the quarter were not included in the M3 equation initially selected by the automatic procedure, which preferred the $\text{PREVPRO}_t^{(m2)}$ balance. We nevertheless preferred to include $\text{EVLIV}_t^{(m3)}$, which, although it may provide less information, remains statistically significant when it is introduced into the equation. Further, the two competing equations display comparable in-sample fits and out-of-sample performances (see Table 3 in Section 5.1). This strategy represents a way to overcome the problem of computing the third forecast based on an equation that presents exactly the same variables as the equation for the second month, disregarding therefore the information stemming from survey data collected over the third month of the quarter (even if we clearly observe that, in practice, moderate new information is provided).

4.4 Discussion on estimation results

The statistical properties of the estimated forecasting equations are broadly satisfactory. In particular, the hypothesis of absence of autocorrelation (order 1-4), homoskedasticity and normality of residuals is not rejected (with the exception of some evidence of residual serial correlation in the M2 equation). Stability of regression coefficients is checked through recursive estimates. This assumption is not rejected for any of the variables and equations, but it is nonetheless worth noticing that in the absence of the dummy variable, the 1-step recursive Chow test reaches a peak in Q1 2009. This justifies the introduction of this variable in our equations.[14] On the basis of estimation results, we can compute the long-term GDP growth rate (the unconditional mean) and compare it to the empirical mean.[15] For the pre-crisis period of Q1 2000 - Q4 2007, the three equations point to a quarterly GDP growth rate of 0.45%, which is consistent with the observed average growth rate (0.44%).

As for the selected variables, we observe that only balances of opinion from surveys on the manufacturing industry are included in forecasting equations. This is due to the fact that manufacturing industry is a sector showing sizable output swings and spillover effects on other sectors, such as services to firms. Over the estimation period, the correlation index between the growth rate of GDP and that of manufacturing output is 0.89. Therefore, unlike with the previous version of MIBA, no balances of opinion on services has been selected in the present study, in spite of the increasingly significant role played by services in the economy. This result is often observed in the economic literature, with the exception of Bessec (2010), where balances from the INSEE survey on services and construction are useful for forecasting final GDP values, while the equations do not include autoregressive terms (see Section 6.4 below). In our study, the exclusion of balances from the survey on services mostly reflects the choices made by the *Autometrics* algorithm in the last round of the automatic procedure, *i.e.*, when "final" models (usually very few, but all having passed the battery of specification tests) are assessed against each other and a unique final model is selected based on information criteria. In other words, models including balances on services would perhaps have underperformed, but only slightly. In practice, we observed that a very limited number of balances of opinion on services (mainly, the balance on expected activity in services instead of the balance on expected production) were eventually not selected due to the slight statistical superiority of balances of opinion on the manufacturing industry over the observed sample. However, this superiority being assessed in-sample by the automatic procedure, we report in Section 6.2 out-of-sample results for alternative models using balances of opinion on services.

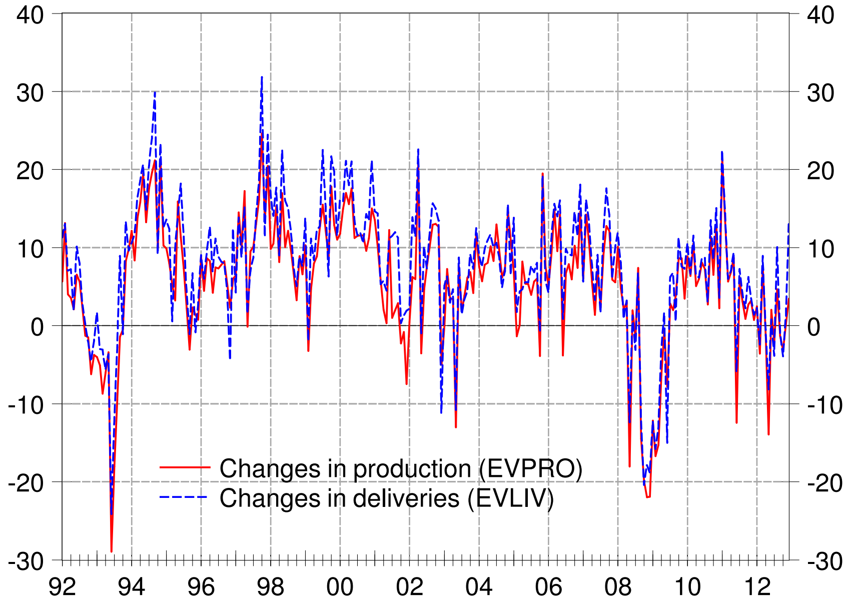
Conversely, what is much less clear is the contribution of the 50 factors constructed from balances of opinion on sub-sectors of activity (4 sub-sectors in the manufacturing industry with 5 factors each, and 10 sub-sectors in services with 3 factors each), which did not provide conclusive results and were not selected for that reason. The same applied to the combination of balances of opinion (14 for the manufacturing industry and 7 for services) with PCA factors.

The selection described above points to the high consistency between models: the information used in the first month focuses on contemporaneous information (changes in deliveries) and information on growth prospects (expected changes in production). From the second month on, the contemporaneous information becomes progressively more substantive, while forward-looking information carries less weight (the estimated coefficient is halved), until it is definitely discarded from the forecasting model associated with the third month of the quarter.

[14]Recursive parameter estimates are presented graphically in Figures A3, A4 and A5 in the Appendix. Visual inspection of these figures suggests that parameters are broadly stable over the sample considered (Q1 1995 - Q4 2012).

[15]The unconditional mean μ_Y is computed as $\mu_Y = (1-\hat{\beta})^{-1}[\hat{\alpha} + \hat{\boldsymbol{\delta}}'\boldsymbol{\mu}_X]$, where $\boldsymbol{\mu}_X$ is the vector of empirical means of regressors X_t, while $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\boldsymbol{\delta}}$ are the estimated parameters defined in Equations (1), (2) and (3).

Figure 5: Comparison of balances - past activity (M1 1992 - M12 2012)



A remark may be made on the presence of the balance on changes in deliveries in all the equations, rather than the balance of opinion on, say, changes in production (EVPRO). Since it refers to actual output, the use of the latter would appear, intuitively, more consistent with the aim of modeling and forecasting a measure of aggregated output such as GDP. In practice, these two balances are strongly correlated (the correlation index is about 0.95) and both track cyclical GDP growth relatively well. However, the balance on changes in deliveries has the statistical and economic advantage (at least for modeling GDP) of appearing slightly less volatile than the balance of opinion on changes in production when there is a slowdown in economic activity (see Figure 5), which also means fewer false alarm of a drop in output and GDP. A more normative explanation of this statistical feature relates to the structure of the survey and to accounting data available to business managers when information is being processed and released. For instance, according to the Banque de France's Surveys and Sectoral Statistics Directorate, which conducts these surveys, the balance on changes in deliveries essentially reflects changes in firms' sales, *i.e.*, a fairly accurate quantification of monthly evolution of activity, while the balance on changes in production incorporates a higher level of uncertainty, due to the absence of accurate and readily available information on current output levels.

Table 3: Forecast results (Q1 2002 - Q4 2012)

	Q1 2002 - Q4 2012		Q1 2002 - Q1 2007		Q2 2007 - Q4 2012		Q2 2009 - Q4 2012	
Equations	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
M1	0.23 (0.03)	0.30 (0.04)	0.22 (0.04)	0.27 (0.05)	0.24 (0.04)	0.32 (0.06)	0.19 (0.05)	0.23 (0.06)
M2	0.19 (0.02)	0.25 (0.03)	0.21 (0.03)	0.27 (0.04)	0.16 (0.03)	0.24 (0.04)	0.12 (0.04)	0.16 (0.05)
M3	0.17 (0.02)	0.25 (0.03)	0.20 (0.04)	0.26 (0.04)	0.15 (0.04)	0.25 (0.05)	0.10 (0.04)	0.14 (0.05)
M1 w/ known Y_{t-1}	0.21 (0.02)	0.29 (0.03)	0.20 (0.03)	0.26 (0.04)	0.22 (0.03)	0.31 (0.05)	0.15 (0.04)	0.19 (0.05)

Notes: Standard errors (in parentheses) are computed through non-parametric bootstrap with 10,000 draws.

5 Forecast evaluation: Q1 2002 - Q4 2012

5.1 Out-of-sample performance

Out-of-sample performance are assessed over the period spanning from Q1 2002 to Q4 2012 (44 observations). The evaluation period is split into three sub-periods (Q1 2002 - Q1 2007, Q2 2007 - Q4 2012 and Q2 2009 - Q4 2012), in order to account for potential underperformance due, for instances, to the transition of quarterly accounts from constant prices to chain-linked prices in Q1 2007 and to the effects on GDP of the financial crisis of 2007-2008. Forecasts are obtained recursively and predictive accuracy is assessed according to MAE (mean absolute error) and RMSE (root mean squared error) criteria. Bootstrap standard errors for these criteria are also computed. Results are presented in Table 3.

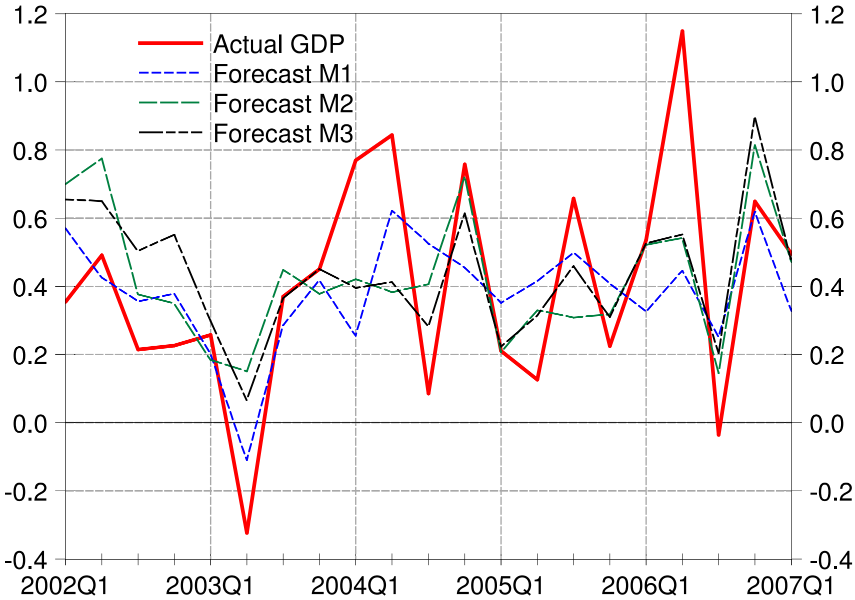
In the context of a real-time analysis, the first month equation raises the issue of the availability of the previous quarter's GDP. Indeed, GDP figures are released roughly 10 to 15 days after the publication of the monthly business survey.[16] To overcome this problem, the solution adopted here (due notably to a better out-of-sample performance, compared to alternative solutions explored preliminarily, such as an equation without lagged GDP and/or with a moving-average term) consists in using the GDP growth predicted by the M3 equation of the previous quarter in the M1 equation.[17] Forecast errors are therefore expected to be fairly larger for the first projection, because the model embeds the prediction of the previous quarter's GDP. This intuition is not confirmed for the Q1 2002 - Q1 2007 sub-period, as mean (absolute and squared) errors are quantitatively the same across forecasting equations. Further, forecast evaluation of a *pseudo*-M1 equation, incorporating the actual information about the previous quarter's GDP (which is not known in a real-time exercise), point to an overall moderate predictive gain compared to the M1 equation (see the last row in Table 3). These findings suggest that the solution adopted in the present paper (*i.e.*, using previous estimates of GDP growth) does not introduce a significant noise in M1 predictions.

[16] We remind the reader that MIBA forecasts and Banque de France's monthly business surveys are supposed to be contemporaneously released to the public by the 10th working-day of the following month.

[17] As to the computation of standard errors, we maintained the real-time constraint by first resampling (with replacement) residuals from the M1 equation and the response variable, and then using the obtained $Y_t^\star$ to estimate the M3 equation and produce forecasts to be embedded into the M1 equation.

Figure 6: Forecasts and forecast errors (Q1 2002 - Q1 2007)

(a) Forecasts



(b) Forecast errors

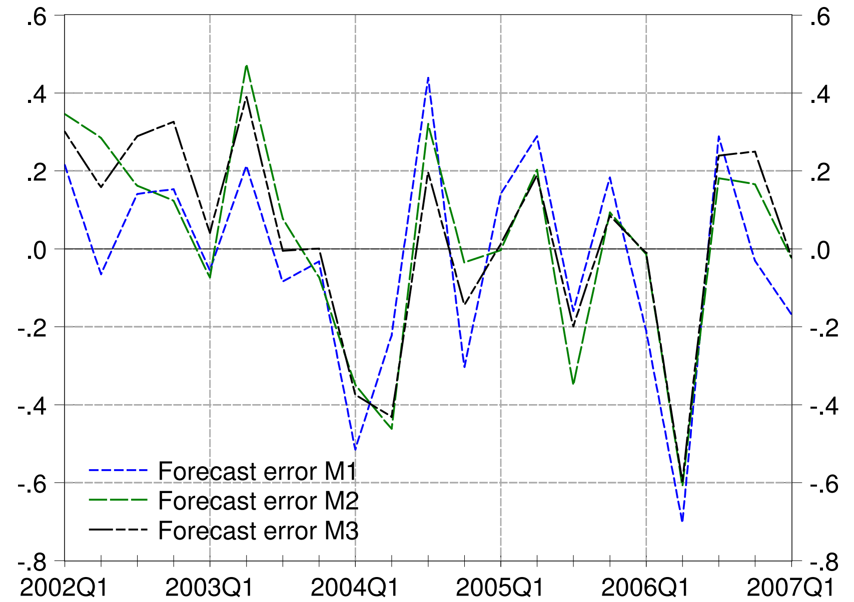
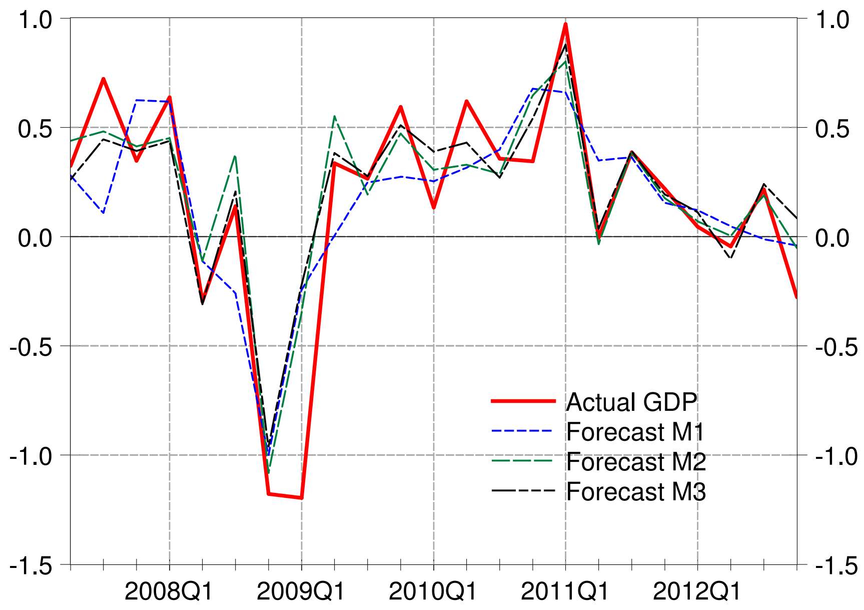
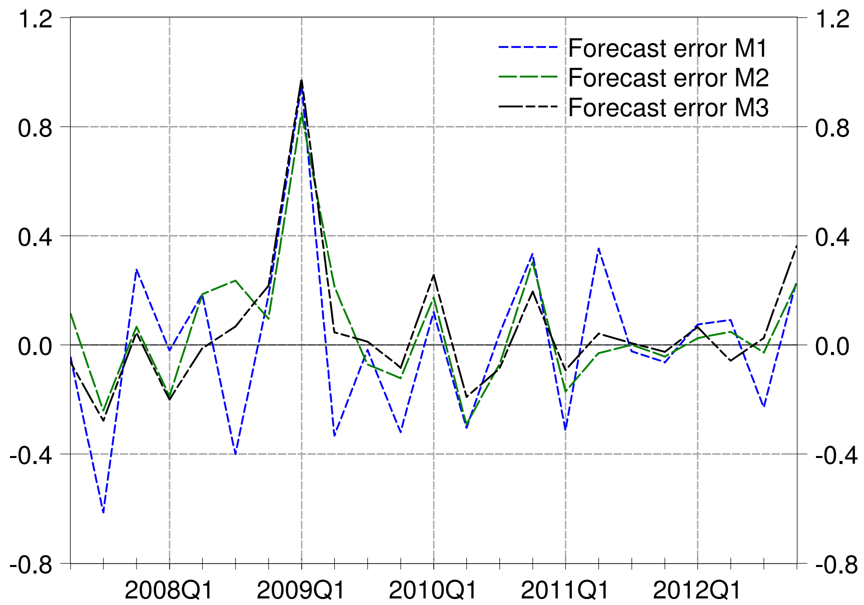


Figure 7: Forecasts and forecast errors (Q2 2007 - Q4 2012)

(a) Forecasts



(b) Forecast errors



Forecasts and forecast errors are graphically presented in Figures 6 and 7, while both empirical and theoretical error densities are presented in Figure A2 in the Appendix. Figures 6a and 6b show a long sequence of over-predictions from 2002 up to mid-2003, which are unexpectedly greater for M2 and M3 projections than for M1. The strong growth rates observed in the first and second quarters of 2004 are instead under-predicted. From mid-2004 up to the remarkable growth rate observed in Q2 2006, forecast errors are small and show an uneven pattern around zero. In 2007, M1 forecasts post a GDP growth rate below 0.4% per quarter, while the first releases hover around 0.5%.

Over the period from 2008 on (Figures 7a and 7b), the model is extremely accurate in predicting the sharp drop observed in Q4 2008, but it broadly over-predicts the growth rate in Q1 2009, partly because of the correction induced by the negative autoregressive term. From the second quarter on, forecast errors from M2 and M3 equations are less significant than those from M1 equation. As expected, the contribution of $\textsc{evliv}_t^{(m3)}$ to the third month forecast is very small, which reflects its lower weight compared to survey data collected over the first two months of the quarter. However, in Q4 2012, the sharp rise in deliveries observed in December broadly contributes to the over-prediction posted by the M3 model (M1 and M2 models predict a slight negative GDP growth rate).

5.2 The source of forecast revisions

As pointed out in the previous paragraph, forecasts are subject to revisions across the same quarter. Revisions can be decomposed in two factors: changes in the model and new data inflows. The former is a natural consequence of the *blocking* approach, since under this framework monthly equations are supposed to evolve according to the availability of data. The latter is instead a consequence of the nowcasting design of the MIBA model, which is supposed to account for real-time data inflows. These two factors can be quantified by approximating forecast revisions as the contribution of new information to each monthly forecast (Dubois and Michaux, 2006; Banbura et al., 2011).

Let us consider two consecutive forecasting equations (such as equations M1 and M2), so that we also have two, possibly different, consecutive predictions. Hence, our aim is to show how much of the revision observed between the first prediction and the second prediction can be explained by the innovation contained in most recent data inflows. One way to achieve this goal is to estimate this innovation, *i.e.*, the non-redundant information contained in the balances of opinion entering the second forecasting equation only, with respect to both the common information and the information used by the first forecasting equation only (see the formal proof presented in Appendix A.1). Results are presented graphically in Figure 8.

From Figure 8a, forecast revisions occurring between M1 and M2 are quite large ($\sigma = 0.2$), but can be mostly explained (about 91%) by innovations, which can in turn be decomposed in survey innovations and GDP innovations. The former reflect the inclusion of $m2$ variables ($\textsc{evliv}_t^{(m2)}$ and $\textsc{prevpro}_t^{(m2)}$) in the M2 equation, while the latter reflects the update of GDP for quarter $t-1$, from an estimated value used in the M1 forecast exercise to an official value used in the M2 forecast exercise (see Section 5.1 for a discussion on publication lags). It is nevertheless worth noticing that most of the total innovation is explained by survey innovation (about 75%). This finding can be interpreted as the moderate role played by GDP updates in the revision of forecasts, and confirms the evaluation results discussed in Section 5.1 and reported in Table 3 (Equation M1 with known GDP). Figure 8b points out the small forecast revisions between M3 and M2 ($\sigma = 0.1$), which are consistent with the evaluation results reported in Table 3. Further, only a moderate share of revisions can be explained by survey innovation (about 20%), which is also consistent with the evidence of poor contribution of $m3$ survey data to the M3 equation.

6 Benchmarking the new MIBA model

6.1 Simple benchmarks

The relative performance of the new MIBA model is assessed by comparison with alternative forecasting models. Following the practice in the forecasting literature, simple benchmarks, such as AR(p) forecasts and *naïve* random-walk predictions, are presented in Panel A of Table 4. The sample spans from Q1 1992 to Q4 2012, and out-of-sample predictions are evaluated for the period spanning from Q1 2002 to Q4 2012. As in the previous paragraph, the evaluation period is split into three sub-periods (Q1 2002 - Q1 2007, Q2 2007 - Q4 2012 and Q2 2009 - Q4 2012). The relative predictive accuracy of the new MIBA model is assessed according to relative MAE and RMSE criteria. In practice, a value below 1 means that the new MIBA is more accurate than its benchmark. As expected, results point to strong predictive gains for the new MIBA model: over the full evaluation window, the predictive gains amount to 30-40% compared to autoregressive models and to 50% compared to random-walk; however, over more recent samples (including the 2008-2009 crisis or not) predictive gains can increase up to 40-50% and 60-70%, respectively. These findings suggest that survey data play a crucial role in improving the accuracy of GDP predictions.

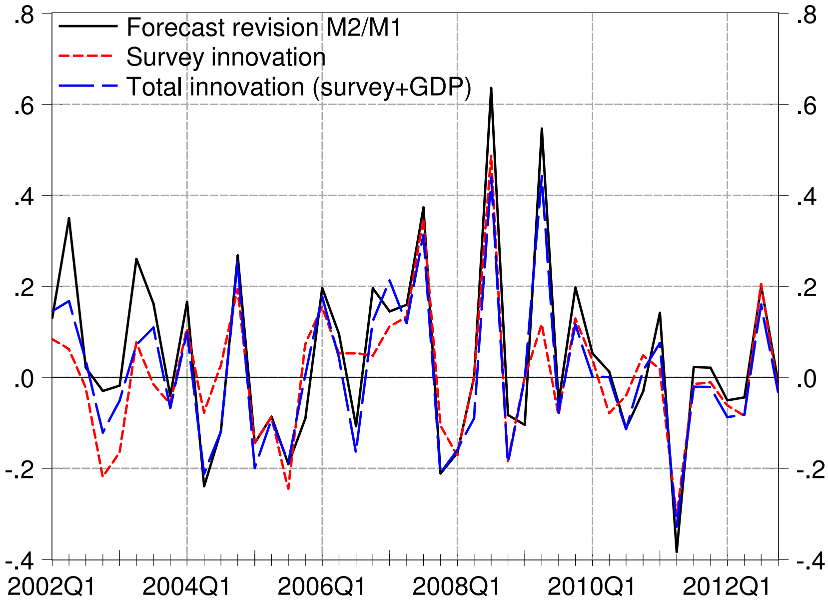
6.2 The new MIBA model including services

In addition to simple benchmarks, we present the relative predictive performance of the new MIBA model with respect to alternative equations accounting for balances of opinion on services. Indeed, as described in Section 4.4, these balances were not selected by the automatic procedure due to the slight statistical superiority of balances of opinion on the manufacturing industry. We then check that the in-sample superiority also translates into out-of-sample outperformance. Results are reported in Panel B of Table 4, while the equations including balances of opinion on services are presented in greater detail in the Appendix (Table A4).[18] Two features are

[18] A visual inspection of in-sample criteria reported in Table A4 (R^2, regression standard error, SIC) reveals a slightly better performance for equations including balances on services, compared to new MIBA equations. However, this result is a modeling artifact induced by the presence of the dummy variable in Q1 2009. In practice, the automatic model selection procedure was implemented prior to the analysis of parameters stability and the inclusion of impulse indicators.

Figure 8: Forecast revisions and innovations (Q1 2002 - Q4 2012)

(a) Revisions and innovations M2/M1



(b) Revisions and innovations M3/M2

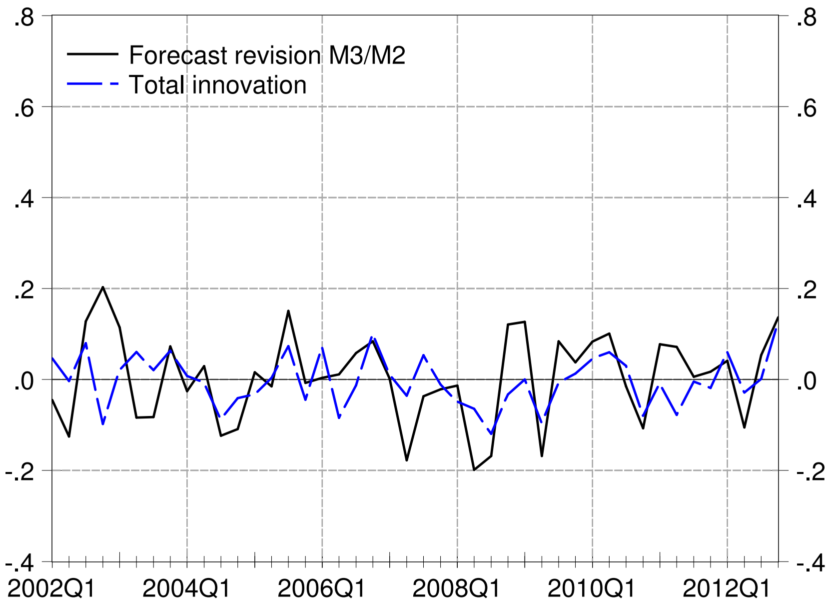


Table 4: Benchmarking the new MIBA forecasts (Q1 2002 - Q4 2012)

Panel A. New MIBA *vs* simple benchmarks

	Q1 2002 - Q4 2012		Q1 2002 - Q1 2007		Q2 2007 - Q4 2012		Q2 2009 - Q4 2012	
Equations	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
AR-M1	0.75 (0.11)	0.71 (0.12)	0.84 (0.18)	0.84 (0.18)	0.69 (0.20)	0.66 (0.21)	0.66 (0.25)	0.63 (0.24)
AR-M2	0.61 (0.10)	0.60 (0.11)	0.81 (0.16)	0.83 (0.16)	0.47 (0.19)	0.49 (0.20)	0.42 (0.22)	0.43 (0.21)
AR-M3	0.57 (0.09)	0.61 (0.10)	0.78 (0.15)	0.80 (0.15)	0.43 (0.17)	0.51 (0.19)	0.36 (0.21)	0.40 (0.20)
RW-M1	0.53 (0.08)	0.54 (0.08)	0.53 (0.13)	0.54 (0.13)	0.54 (0.10)	0.54 (0.10)	0.45 (0.15)	0.40 (0.13)
RW-M2	0.43 (0.06)	0.46 (0.05)	0.51 (0.11)	0.53 (0.10)	0.37 (0.08)	0.40 (0.06)	0.29 (0.08)	0.27 (0.07)
RW-M3	0.40 (0.06)	0.46 (0.05)	0.49 (0.10)	0.51 (0.10)	0.33 (0.08)	0.42 (0.07)	0.24 (0.09)	0.25 (0.07)

Panel B. New MIBA *vs* New MIBA with services

	Q1 2002 - Q4 2012		Q1 2002 - Q1 2007		Q2 2007 - Q4 2012		Q2 2009 - Q4 2012	
Equations	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
M1	0.94 (0.15)	0.91 (0.14)	0.87 (0.27)	0.91 (0.26)	1.00 (0.21)	0.90 (0.19)	1.17 (0.31)	1.07 (0.30)
M2	0.98 (0.13)	0.95 (0.13)	1.00 (0.23)	0.99 (0.22)	0.96 (0.19)	0.91 (0.18)	0.93 (0.31)	1.00 (0.30)

Panel C. New MIBA *vs* previous MIBA

	Q1 2002 - Q4 2012		Q1 2002 - Q1 2007		Q2 2007 - Q4 2012		Q2 2009 - Q4 2012	
Equations	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
M3	0.82 (0.13)	0.93 (0.14)	0.93 (0.23)	0.91 (0.21)	0.71 (0.16)	0.94 (0.21)	0.61 (0.19)	0.66 (0.19)

Panel D. First-release MIBA *vs* Final-release MIBA

	Q1 2002 - Q4 2012		Q1 2002 - Q1 2007		Q2 2007 - Q4 2012		Q2 2009 - Q4 2012	
Equations	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
M1	0.84 (0.13)	0.92 (0.14)	0.88 (0.23)	0.94 (0.23)	0.81 (0.19)	0.91 (0.20)	0.71 (0.22)	0.70 (0.21)
M2	0.89 (0.13)	0.96 (0.14)	0.97 (0.23)	0.92 (0.20)	0.81 (0.18)	1.02 (0.21)	0.69 (0.21)	0.77 (0.22)
M3	0.78 (0.11)	0.89 (0.13)	0.88 (0.21)	0.84 (0.18)	0.68 (0.15)	0.94 (0.20)	0.52 (0.16)	0.62 (0.17)

Notes: Relative MAE and RMSE (<1 means that the new MIBA outperforms the benchmark). AR denotes the autoregressive forecasting model with lags selected by the Schwarz information criterion (max 5 lags). RW denotes the *naïve* random-walk forecast. Standard errors (in parentheses) are computed through non-parametric bootstrap with 10,000 draws.

worth noticing. *First*, only results for M1 and M2 equations are reported in Table 4, since no balances of opinion on services competed against balances on the manufacturing industry for the M3 equation. *Second*, equations only differ for the presence of the balance on expected changes in activity instead of the balance on expected changes in production, so that the models are very similar in terms of economic interpretation. Out-of-sample results tend to confirm the in-sample model selection, because relative RMSE and MAE mainly point to a better, although not distinctly statistically different, predictive accuracy for the new MIBA models. Only for the more recent evaluation period (Q2 2009 - Q4 2012), results seem to suggest that the M1 equation including expected changes in activity outperforms the M1 MIBA equation, by about 10-20%, although standard errors appear quite large (0.3).

6.3 The previous version of the MIBA model

Out-of-sample performance of the new equations are compared here to the historical performance of the previous version of the MIBA model. This model consists of three equations, each using a lag of GDP and the first factor from a PCA on balances of opinion on the manufacturing sector, as well as, respectively, the third factor on the manufacturing sector (first equation), the second factor on the sub-sector "manufacturing of electric, electronic equipment and machines" (second equation), and the "changes in activity" balance of opinion from the monthly business survey on services (third equation). These three equations, whose forecasts are combined using equal weights, are presented in greater detail in the Appendix (Table A5).[19]

As discussed in Section 2, several methods are implemented in the previous version of the MIBA model to deal with missing observations when forecasting GDP in the first and second months of the quarter. To simplify the benchmarking exercise, we only focus on the third month forecast, because the underlying equation is not affected by missing observations. To meet the objective of real-time analysis adopted in this study, and in order to replicate the actual working condition of the forecaster, the GDP series used in previous MIBA equations coincides with the vintage estimate available at the time of each forecasting exercise. For this purpose, we use a real-time dataset.

Compared to the equations described in the previous paragraphs, equations forming the previous version of the MIBA model are characterized by a very low statistical significance of the first-order autoregressive term of GDP. The analysis of several vintage estimates shows that this remarkable difference have only recently arisen, and it can be attributed to the revision of INSEE quarterly accounts, carried out in May 2011 at the time of the transition to the base-year 2005. Indeed, the equations estimated just before this revision report a statistically significant autoregressive term. According to the analysis of quarterly accounts presented by INSEE in its *Note de Conjoncture* (Economic Outlook) of June 2011 (entitled "The quarterly national accounts switch to the 2005 base-year"), this feature could be attributed to revisions of investment data, as well as to an upgrade of seasonal-adjustment procedures.

In practice, as a consequence of this change in the stochastic nature of the series, the autoregressive term of GDP growth rate is no longer statistically significant in regressions using final-release data. However, in order to maintain some consistency with the latest MIBA model on duty, the autoregressive term of GDP growth is kept in the benchmarking exercise. The results are presented in Panel C of Table 4.

It is worth noticing that the new MIBA model is in general more accurate, although not always statistically different, than the previous MIBA model, regardless of the evaluation period considered. However, the accuracy gap between the two models tends to widen in the second part of the sample, achieving the noteworthy predictive gain of roughly 30-40% over the period covering the end of the recession up to Q4 2012, as pointed out by both the RMSE and the MAE criteria.

[19] The equations used in the present exercise correspond to the most recent versions officially on duty, and they differ slightly from the equations presented in Darné and Brunhes-Lesage (2007).

6.4 A model using final-release GDP

The forecasting performance of the new MIBA model, which is based on first-release GDP, is compared here to the performance of an alternative model, still implementing the blocking approach, but based instead on final-release GDP. The sample spans from Q1 1992 to Q4 2012, and relevant indicators for the first, second and third month equations are again selected using the "general-to-specific" *Autometrics* algorithm. As to the comparison with the previous MIBA, this model is estimated in real-time, such that the used GDP series coincides with the vintage estimate available at the time of each forecasting exercise. The estimated autoregressive term is not statistically significant; we nonetheless keep this variable in the equations. Estimation results are presented in the Appendix (Table A6).

Unlike specifications based on first-release data, equations based on final GDP data include a balance of opinion on services (the balance on changes in activity in the first month of the quarter, $\text{EVACT}_t^{(m1)}$). As reported by Bessec (2010), survey data on services may therefore be useful for modeling and predicting the final GDP.[20]

We also observe that the introduction of dummy variables in Q3 and Q4 2008 and in Q1 2009, are required to deal with the instability of regression coefficients. Thus, compared to the new MIBA, these equations seem to poorly fit the sharp changes in GDP observed during the 2008 - 2009 economic crisis. However, this feature appears to be largely due to the difficulty of taking into account GDP revisions, extremely substantial over the 2008 - 2009 period (see Figure 7b), carried out over time by INSEE, because the goodness of fit of the model deteriorates as we use more recent vintage data.

Relative MAE and RMSE are reported in Panel D of Table 4. For the 2002 - 2012 period, the results suggest that new MIBA forecasts are generally more accurate than those from equations based on final GDP estimates. With the exception of the more recent period, predictive gains may nevertheless appear moderate and the models statistically equally accurate. This feature can be attributed to the good forecasting performance obtained by the benchmark model during the 2008 - 2009 crisis period, in particular the Q1 2009 GDP forecast.[21] Conversely, the new MIBA model displays a remarkably greater and statistically significant predictive accuracy over the period spanning from the end of the recession up to Q4 2012, displaying predictive gains of roughly 30-50%. These findings are confirmed by a visual inspection of Figure 9a, where we report forecast errors for second month equations (M2 equation of Table 3 and M2b of Table A6) computed with respect to first-release GDP.

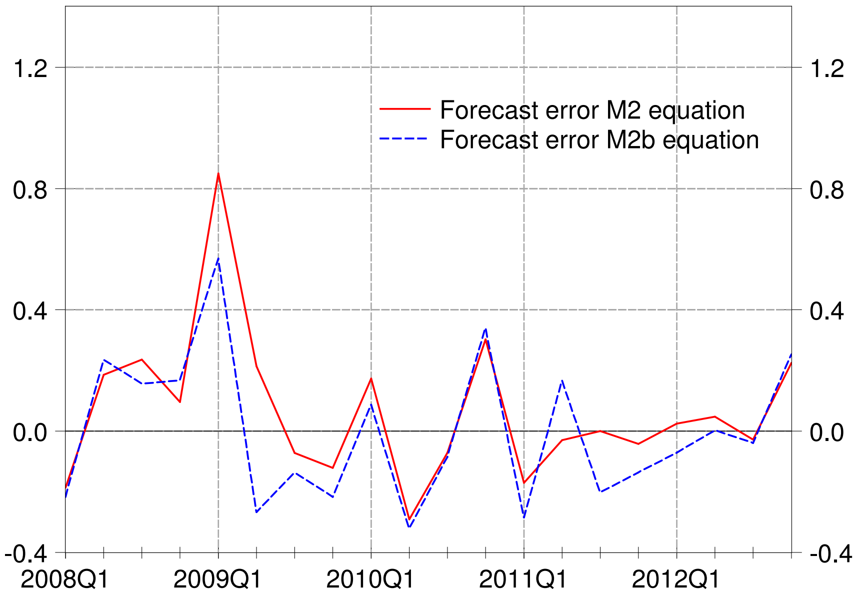
When the forecast errors are computed with respect to the latest available GDP vintage estimate (*i.e.*, the February 2013 vintage, at the time of writing), the model based on final-release data seems to perform slightly better (RMSE of 0.16 since Q2 2009, compared to 0.21 for the model based on first-release data). Survey data on services display some predictive content on long-term revisions (*i.e.*, the difference between first- and final-releases). Revisions

[20]As for the new MIBA equations, the balance on expected changes in activity on services (PREVACT_t) is often in competition with the balance on expected changes in production on the manufacturing sector (PREVPRO_t).

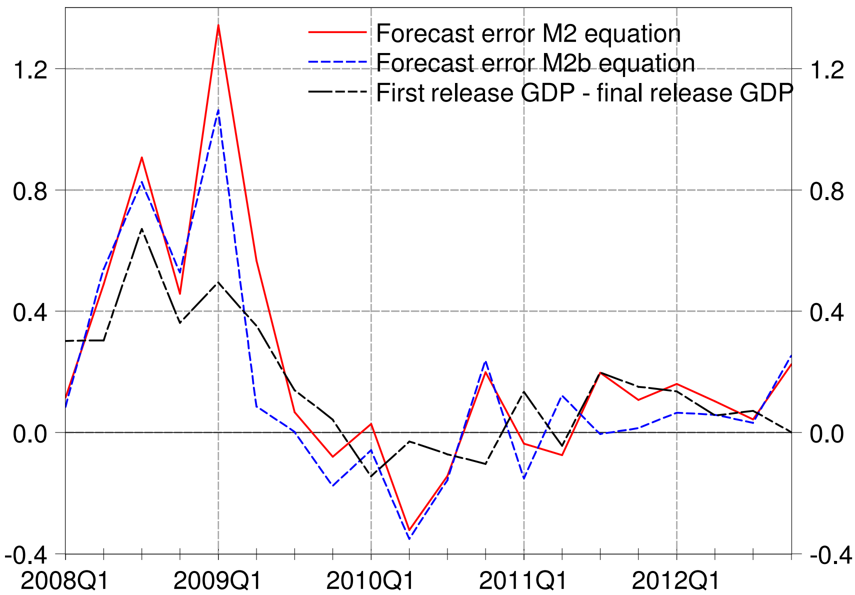
[21]It must be noted however that equations based on final-release GDP require a larger number of dummy variables, covering almost the entire recession period (see Table A6).

Figure 9: Forecast errors with respect to first- and final-release GDP (Q1 2008 - Q4 2012)

(a) Forecast errors and first release GDP



(b) Forecast errors and final release GDP



Notes: Equation M2 is presented in Table 2, while Equation M2b is presented in Table A6.

Table 5: Predictive accuracy of M2 and M3 equations across quarters (Q1 2002 - Q4 2012)

	Q1 2002 - Q4 2012				Q1 2007 - Q4 2012			
	M2		M3		M2		M3	
Quarters	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Q1	0.14	0.18	0.14	0.19	0.12	0.14	0.13	0.15
Q2	0.26	0.32	0.20	0.27	0.15	0.17	0.07	0.09
Q3	0.16	0.19	0.13	0.17	0.11	0.14	0.08	0.12
Q4	0.12	0.14	0.15	0.19	0.14	0.16	0.15	0.18

and forecast errors are particularly strong for the period covering the economic downturn in 2008 and 2009. These errors are already large when computed using first-release data (Figure 9a), but they flare up when computed using final-release data (Figure 9b). The sum of the four quarterly growth rates up to the first quarter of 2009 shifted from -2.5 points in first-release data to -4.4 points in the final-release vintage.

6.5 Impact of months sensitive to seasonal variations on first-release GDP forecasts

The results presented above disregard the problem of measurement errors on predictors. In particular, the equations are based on the assumption that survey data are not affected by recurrent statistical issues related to the monthly processing of data, such as the well-known troublesome seasonal-adjustment in August. Otherwise, this would lead to a systematic predictive bias in Q3 arising from second and third monthly forecasts, which are obtained through equations including balances of opinion collected in August. In order to check for the absence of this bias, forecast evaluation values reported in Table 3 are broken down by quarters. Results, reported in Table 5, do not seem to suggest the presence of a systematic bias in the M2 and M3 forecasts for Q3. The forecasting performance is indeed statistically equivalent across quarters. A similar evidence can be found for forecasts released in other months supposed to be affected by seasonal issues, such as December or May (depending on the year), especially if we consider the recent 2007 - 2012 period. Summing up, while the data collected in some specific months may be affected by seasonal issues, this does not appear to translate into a deterioration of the predictive accuracy of our model.

7 Concluding remarks

In this study we proposed a new MIBA model for forecasting France's GDP. As the previous versions, the model relies exclusively on data from the monthly business survey (EMC) on industry and services conducted by the Banque de France. Several major changes have been introduced in this new model. *First*, the new MIBA model is no longer based on the first factor of a principal component analysis (PCA) on the survey on industry (which provides data for the Business Sentiment Indicator), but rather on balances of opinion from this survey. *Second*, the GDP measure targeted by the new forecasting model is the first-release GDP rather than the final GDP estimate. *Third*, missing data are handled through the "blocking" approach, which makes it possible to take into account data that are actually available at the time of forecast, without having to extrapolate the missing values. This approach also makes it possible to use the most relevant information, given the month in which the forecast is conducted. Model selection was carried out using the general-to-specific approach from the *Autometrics* algorithm, providing different models for each month of the quarter. For the first month of the quarter, the model is a mix of information on past and expected activity (changes in deliveries and expected changes in production), the latter outweighing the former. For the second and third month of the quarter, contemporaneous data are taken into account and forward-looking information is progressively discarded. It is worth noticing that data from the survey on services are eventually not selected for this model. As for the forecasting performance, out-of-sample evaluation on the Q1 2002 - Q4 2012 period showed that the equations used in the new MIBA model lead to forecast errors that are smaller, in terms of RMSE and MAE criteria, compared to various alternative specifications, such as the previous version of the model.

References

Banbura, M., Giannone, D., Reichlin, L., 2011. Nowcasting. In: Clements, M. P., Hendry, D. F. (Eds.), Oxford Handbook on Economic Forecasting. Oxford (UK): Oxford University Press, pp. 63–90.

Bessec, M., 2010. Étalonnages du taux de croissance du PIB français sur la base des enquêtes de conjoncture. Économie et Prévision 193, 77–99.

Ciccone, A., Jarocinski, M., 2010. Determinants of economic growth: Will data tell? American Economic Journal: Macroeconomics 2 (4), 222–246.

Clements, M. P., Galvão, A. B., 2013. Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions. Journal of Applied Econometrics 28 (3), 458–477.

Croushore, D., 2011. Frontiers of real-time data analysis. Journal of Economic Literature 49 (1), 72–100.

Darné, O., Brunhes-Lesage, V., 2007. Líindicateur synthétique mensuel d'activité (ISMA): une révision. Bulletin de la Banque de France 162, 21–36.

Doornik, J. A., 2009. Autometrics. In: Castle, J. L., Shephard, N. (Eds.), The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry. Oxford (UK): Oxford University Press, pp. 88–121.

Dubois, E., Michaux, E., 2006. Étalonnages à l'aide d'enquêtes de conjoncture : de nouveaux résultats. Économie et Prévision 172, 11–28.

Faust, J., Rogers, J. H., Wright, J. H., 2005. News and noise in G-7 GDP announcements. Journal of Money, Credit, and Banking 37 (3), 403–419.

Hendry, D., Doornik, J., 2009. Empirical Econometric Modelling using PcGive. London (UK): Timberlake Consultants Press.

Hoover, K. D., Perez, S. J., 1999. Data mining reconsidered: Encompassing and the general-to-specific approach to specification search. The Econometrics Journal 2 (2), 167–191.

Koenig, E. F., Dolmas, S., Piger, J., 2003. The use and abuse of real-time data in economic forecasting. Review of Economics and Statistics 85 (3), 618–628.

Krolzig, H. M., Hendry, D. F., 2001. Computer automation of general-to-specific model selection procedures. Journal of Economic Dynamics and Control 25 (6-7), 831–866.

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. Document de travail G 2010/01, INSEE.

Wallis, K., 1986. Forecasting with an econometric model: the "ragged edge" problem. Journal of Forecasting 5 (1), 1–13.

Appendix

Table A1: MIBA with BSI and its components (Q1 1992 - Q4 2012)

(*i*)			(*ii*)			(*iii*)		
Variable	Coefficient	t-stat	Variable	Coefficient	t-stat	Variable	Coefficient	t-stat
$Constant$	54.44 (4.52)	12.04 [0.00]	$Constant$	55.11 (4.38)	12.58 [0.00]	$Constant$	12.20 (3.82)	3.19 [0.00]
Y_{t-1}	-0.47 (0.09)	-5.19 [0.00]	Y_{t-1}	-0.48 (0.09)	-5.42 [0.00]	Y_{t-1}	-0.47 (0.08)	-5.69 [0.00]
BSI_t	4.73 (0.41)	11.58 [0.00]	$Production_t$	6.09 (0.73)	8.34 [0.00]	$Production_t$	6.54 (0.51)	12.90 [0.00]
			$Order\ books_t$	-5.59 (4.79)	-1.17 [0.25]			
			$Prices_t$	1.06 (18.7)	0.06 [0.95]			
			$Inventories_t$	3.79 (3.88)	0.98 [0.33]			
			CUR_t	6.28 (7.05)	1.12 [0.26]			
Adj-R^2	0.64		Adj-R^2	0.67		Adj-R^2	0.68	
σ_ϵ	26.80		σ_ϵ	25.37		σ_ϵ	24.99	
SIC	9.54		SIC	9.59		SIC	9.40	
Normality	2.02	[0.36]	Normality	0.18	[0.91]	Normality	0.65	[0.72]
AR(4)	0.77	[0.55]	AR(4)	2.12	[0.09]	AR(4)	2.46	[0.05]
Het	7.05	[0.00]	Het	2.41	[0.03]	Het	3.84	[0.03]

Notes: GDP growth is measured in basis points. BSI is the demeaned "pseudo" Business Sentiment Indicator of the Banque de France (see footnote 1). The sub-group *Order books* includes balances on current order books (expressed in level of opinion and weeks of activity), *Production* includes balances on the evolution of new orders (overall and foreign orders) and output (past and expected output, deliveries), *Prices* includes balances on the evolution of prices (commodities and final goods), *Inventories* includes balances on the level of inventories (past and expected inventories of commodities and final goods), and CUR is the average capacity utilisation rate. All the sub-groups are demeaned. See Table 1 for more details on the balances of opinion. Standard errors in parentheses. p-values in brackets. σ_ϵ is the regression standard error. SIC is the Schwarz information criterion. "Normality" denotes the Bera-Jarque test for residual normal distribution. AR(4) denotes the Breusch-Godfrey test for residual serial correlation up to order $p = 4$. "Het" denotes the Breusch-Pagan-Godfrey test for heteroskedasticity.

Table A2: Forecasting GDP with BSI and its components (Q1 2002 - Q4 2012)

	Q1 2002 - Q4 2012		Q2 2007 - Q4 2012	
Equations Table A1	MAE	RMSE	MAE	RMSE
(*i*)	0.23	0.29	0.23	0.31
(*ii*)	0.23	0.30	0.23	0.31
(*iii*)	0.19	0.25	0.19	0.24

Table A3: Test for forecast efficiency and correlation with revised data (Q1 1992 - Q4 2009)

	Forecast efficiency regression $R_t = \alpha + \beta FirstRelGDP_t + \epsilon_t$		Correlation with revised data $R_t = \alpha + \beta FinRelGDP_t + \epsilon_t$	
Variable	Coefficient	t-stat	Coefficient	t-stat
α	-0.01 (0.03)	-0.41 [0.68]	-0.02 (0.03)	-0.69 [0.49]
β	-0.01 (0.07)	-0.07 [0.94]	0.25 (0.05)	4.76 [0.00]
F-test	0.09 [0.92]		11.46 [0.00]	

Notes: Standard errors in parentheses. p-values in brackets. R_t denotes the revisions series, computed as final-release GDP ($FinRelGDP_t$) *minus* first-release GDP ($FirstRelGDP_t$). Regressors are demeaned for parameters estimation. F-test is the Wald test for the joint null hypothesis $\alpha = \beta = 0$.

Table A4: The 2 equations of the new MIBA forecasting model including services

M1s			M2s		
Variable	Coefficient	t-stat	Variable	Coefficient	t-stat
Constant	2.82 (4.87)	0.58 [0.56]	*Constant*	-2.78 (4.50)	-0.62 [0.54]
Y_{t-1}	-0.49 (0.08)	-5.83 [0.00]	Y_{t-1}	-0.44 (0.08)	-5.59 [0.00]
$\text{EVLIV}_t^{(m1)}$	2.84 (0.45)	6.31 [0.00]	$\text{EVLIV}_t^{(m2)}$	1.94 (0.41)	4.78 [0.00]
$\text{PREVACT}_t^{(m1)}$	2.62 (0.42)	6.23 [0.00]	$\text{PREVACT}_t^{(m2)}$	1.54 (0.40)	3.84 [0.00]
			$\text{EVLIV}_t^{(m1)}$	2.29 (0.44)	3.84 [0.00]
Adj-R^2	0.68		Adj-R^2	0.72	
σ_ϵ	25.21		σ_ϵ	23.39	
SIC	9.50		SIC	9.39	
Normality	5.17	[0.08]	Normality	0.62	[0.73]
AR(4)	0.61	[0.66]	AR(4)	2.19	[0.08]
Het	0.29	[0.88]	Het	0.25	[0.94]

Notes: GDP growth is measured in basis points. Standard errors in parentheses. p-values in brackets. σ_ϵ is the regression standard error. SIC is the Schwarz information criterion. "Normality" denotes the Bera-Jarque test for residual normal distribution. AR(4) denotes the Breusch-Godfrey test for residual serial correlation up to order $p = 4$. "Het" denotes the Breusch-Pagan-Godfrey test for heteroskedasticity.

Table A5: The 3 equations of the previous MIBA forecasting model

(*i*) Q2 1987 - Q4 2012			(*ii*) Q2 1987 - Q4 2012			(*iii*) Q1 1989 - Q4 2012		
Variable	Coefficient	t-stat	Variable	Coefficient	t-stat	Variable	Coefficient	t-stat
$Constant$	45.83 (5.02)	9.13 [0.00]	$Constant$	46.12 (4.71)	9.77 [0.00]	$Constant$	17.41 (7.14)	2.44 [0.02]
Y_{t-1}	-0.05 (0.09)	-0.51 [0.61]	Y_{t-1}	-0.05 (0.09)	-0.58 [0.56]	Y_{t-1}	-0.09 (0.09)	-0.97 [0.33]
$F1\ Manuf_t$	43.17 (4.95)	8.72 [0.00]	$F1\ Manuf_t$	44.77 (4.61)	9.71 [0.00]	$F1\ Manuf_t$	38.12 (5.60)	6.80 [0.00]
$F3\ Manuf_{t-1}$	-10.41 (3.29)	-3.16 [0.00]	$F2\ C3_t$	-14.55 (2.95)	-4.93 [0.00]	$F1\ Manuf_{t-2}$	-14.45 (3.66)	-3.95 [0.00]
						EVACT$_t$	2.10 (0.46)	4.54 [0.00]
Adj-R^2	0.67		Adj-R^2	0.71		Adj-R^2	0.71	
σ_ϵ	30.16		σ_ϵ	28.35		σ_ϵ	27.22	
SIC	9.79		SIC	9.67		SIC	9.63	
Normality	0.04	[0.98]	Normality	0.89	[0.64]	Normality	1.20	[0.55]
AR(4)	5.01	[0.00]	AR(4)	2.29	[0.07]	AR(4)	2.46	[0.05]
Het	4.85	[0.00]	Het	1.91	[0.13]	Het	4.43	[0.00]

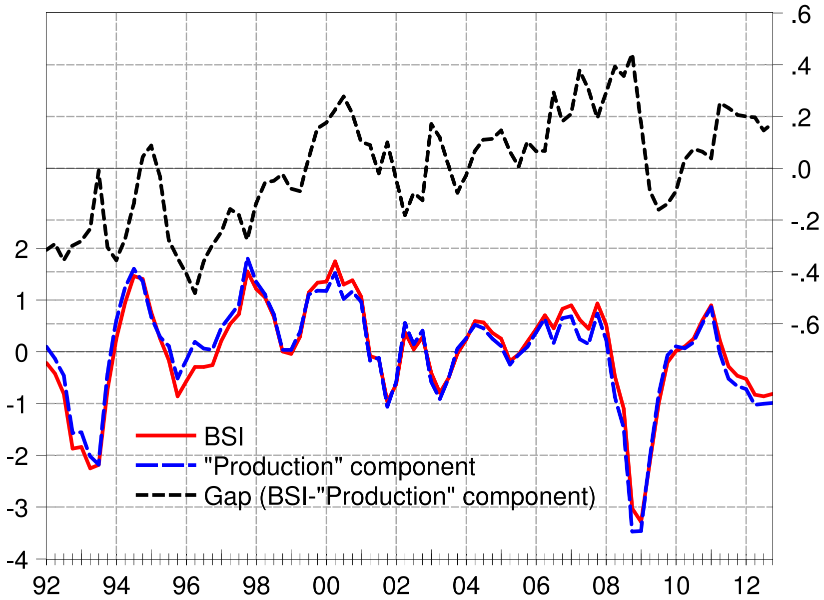
Notes: GDP growth is measured in basis points. $F1\ Manuf$ and $F3\ Manuf$ denote respectively the first and third factors from a PCA on balances of opinion on the manufacturing sector. $F2\ C3$ denotes the second factor from a PCA on balances of opinion on the NAF Rev.2 sub-sector C3 (manufacture of electric and electronic equipment and machines). EVACT denotes the balance of opinion on changes in activity on market services. Standard errors in parentheses. p-values in brackets. σ_ϵ is the regression standard error. SIC is the Schwarz information criterion. "Normality" denotes the Bera-Jarque test for residual normal distribution. AR(4) denotes the Breusch-Godfrey test for residual serial correlation up to order $p = 4$. "Het" denotes the Breusch-Pagan-Godfrey test for heteroskedasticity.

Table A6: MIBA estimates with final-release GDP (Q1 1992 - Q4 2012)

M1b			M2b			M3b		
Variable	Coefficient	t-stat	Variable	Coefficient	t-stat	Variable	Coefficient	t-stat
Constant	13.59	2.56	*Constant*	− 1.85	-0.41	*Constant*	- 0.20	-0.05
	(5.31)	[0.01]		(4.50)	[0.68]		(4.39)	[0.96]
Y^f_{t-1}	- 0.05	-0.53	Y^f_{t-1}	- 0.12	-1.43	Y^f_{t-1}	- 0.14	-1.62
	(0.09)	[0.60]		(0.08)	[0.16]		(0.08)	[0.11]
$\text{EVACT}^{(m1)}_t$	0.99	2.28	$\text{EVLIV}^{(m2)}_t$	2.38	5.68	$\text{EVCOM}^{(m3)}_t$	1.39	3.14
	(0.43)	[0.02]		(0.42)	[0.00]		(0.44)	[0.00]
$\text{PREVPRO}^{(m1)}_t$	4.14	4.99	$\text{PREVPRO}^{(m2)}_t$	2.11	2.89	$\text{EVLIV}^{(m2)}_t$	2.05	4.68
	(0.83)	[0.00]		(0.73)	[0.00]		(0.44)	[0.44]
$\text{EVLIV}^{(m1)}_{t-1}$	- 3.40	-3.29	$\text{EVACT}^{(m1)}_t$	1.20	3.05	$\text{EVACT}^{(m1)}_t$	1.46	4.07
	(1.03)	[0.00]		(0.39)	[0.00]		(0.36)	[0.36]
$\text{EVCOM}^{(m1)}_{t-1}$	2.74	-2.70						
	(1.01)	[0.01]						
			$dummy_{08Q3}$	- 84.9	-3.41	$dummy_{08Q3}$	- 53.2	-2.01
				(24.9)	[0.00]		(26.5)	[0.05]
$dummy_{08Q4}$	-108.1	-3.52	$dummy_{08Q4}$	- 68.7	-2.50	$dummy_{08Q4}$	- 76.5	-2.91
	(30.7)	[0.00]		(27.5)	[0.01]		(26.3)	[0.00]
$dummy_{09Q1}$	- 89.3	-3.09	$dummy_{09Q1}$	-111.9	-3.96	$dummy_{09Q1}$	-127.7	-4.68
	(28.9)	[0.00]		(28.3)	[0.00]		(27.3)	[0.00]
Adj-R^2	0.75		Adj-R^2	0.78		Adj-R^2	0.81	
σ_ϵ	25.53		σ_ϵ	23.79		σ_ϵ	23.59	
SIC	9.64		SIC	9.50		SIC	9.48	
Normality	2.21	[0.33]	Normality	0.77	[0.68]	Normality	0.29	[0.86]
AR(4)	0.89	[0.47]	AR(4)	1.34	[0.26]	AR(4)	0.48	[0.75]
Het	0.38	[0.91]	Het	0.82	[0.57]	Het	0.44	[0.87]

Notes: GDP growth is measured in basis points. Standard errors in parentheses. p-values in brackets. σ_ϵ is the regression standard error. SIC is the Schwarz information criterion. “Normality” denotes the Bera-Jarque test for residual normal distribution. AR(4) denotes the Breusch-Godfrey test for residual serial correlation up to order $p = 4$. “Het” denotes the Breusch-Pagan-Godfrey test for heteroskedasticity.

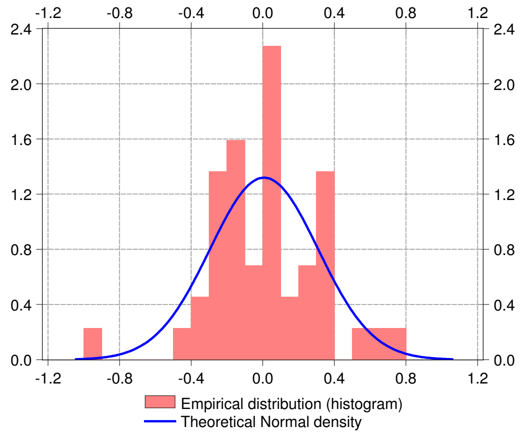
Figure A1: BSI and the *Production* component: levels and gap (Q1 1992 - Q4 2012)



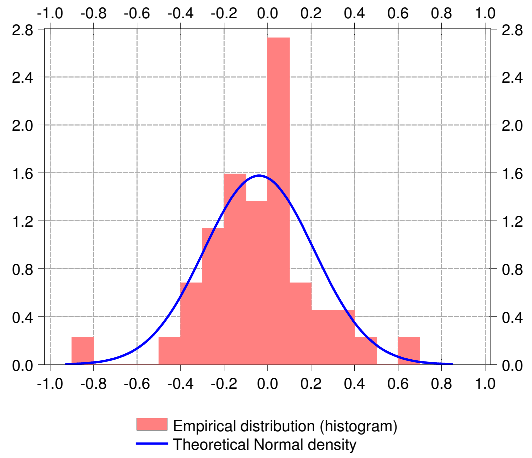
Notes: Series are standardised (mean=0 and standard deviation=1).

Figure A2: Distribution of forecast errors (Q1 2002 - Q4 2012)

(a) MIBA M1 forecast errors distribution



(b) MIBA M2 forecast errors distribution



(c) MIBA M3 forecast errors distribution

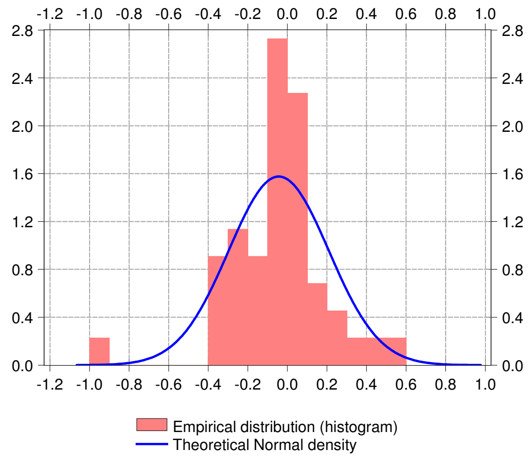
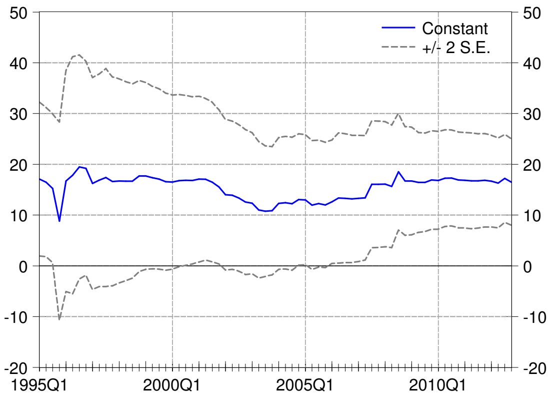
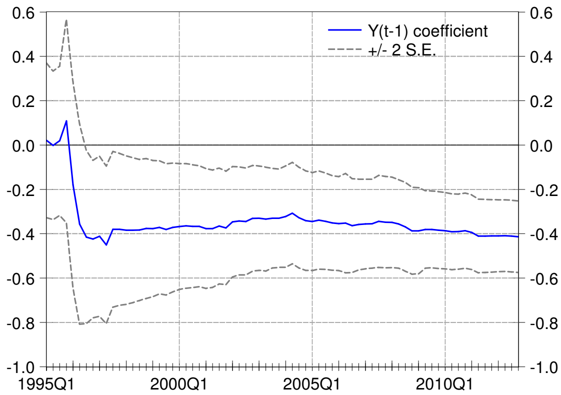


Figure A3: Recursive estimates of new MIBA M1 parameters (Q1 1995 - Q4 2012)

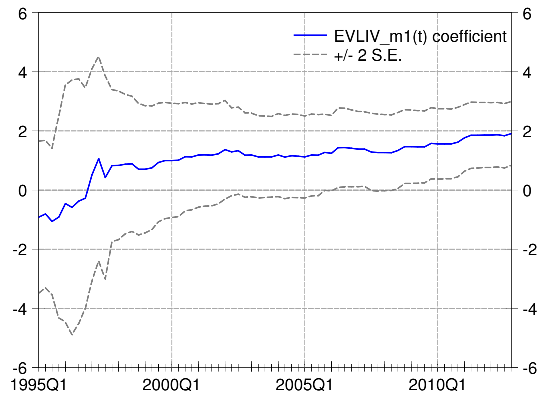
(a) Constant



(b) Y_{t-1}



(c) $\text{EVLIV}_t^{(m1)}$



(d) $\text{PREVPRO}_t^{(m1)}$

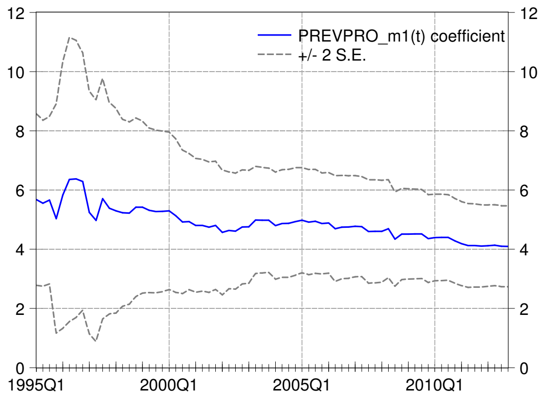
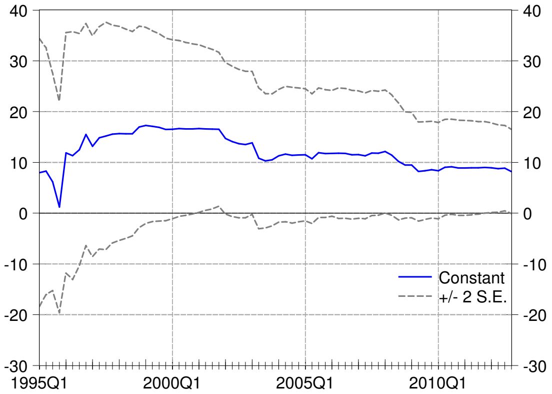
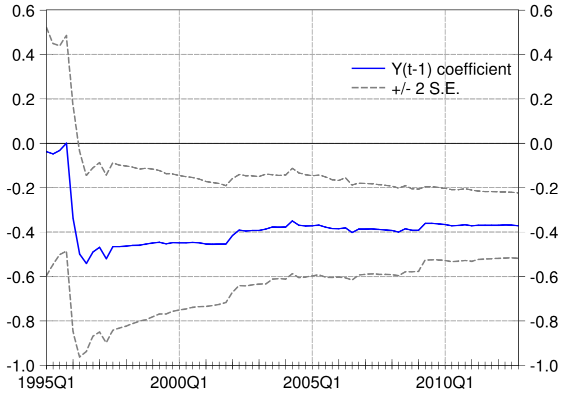


Figure A4: Recursive estimates of new MIBA M2 parameters (Q1 1995 - Q4 2012)

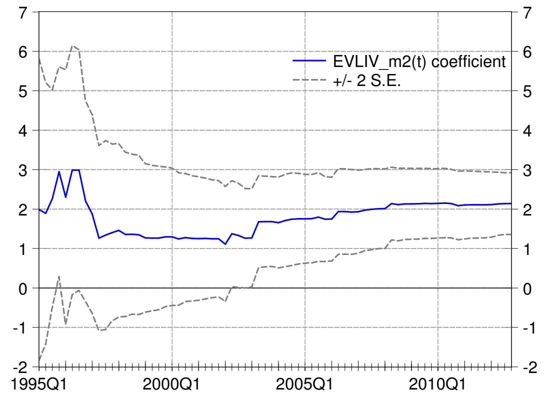
(a) Constant



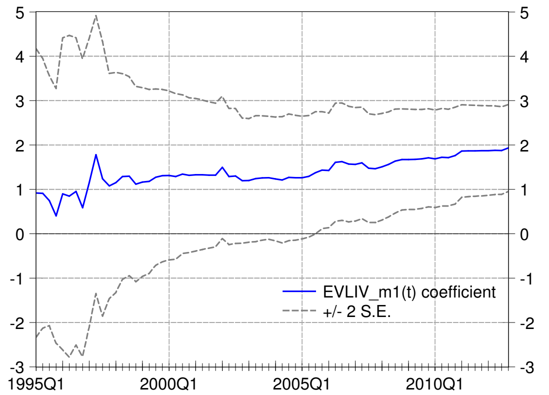
(b) Y_{t-1}



(c) $\text{EVLIV}_t^{(m2)}$



(d) $\text{EVLIV}_t^{(m1)}$



(e) $\text{PREVPRO}_t^{(m2)}$

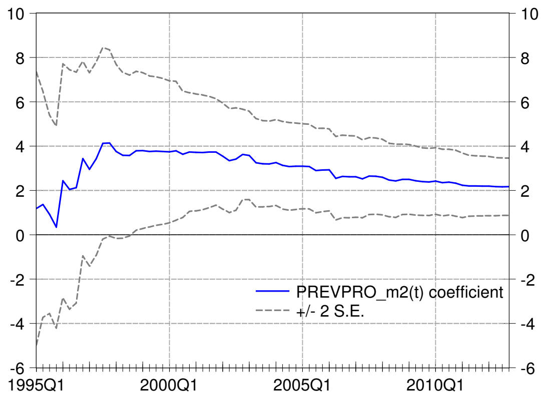
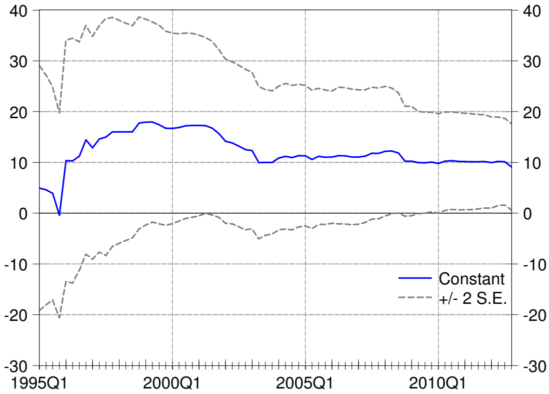
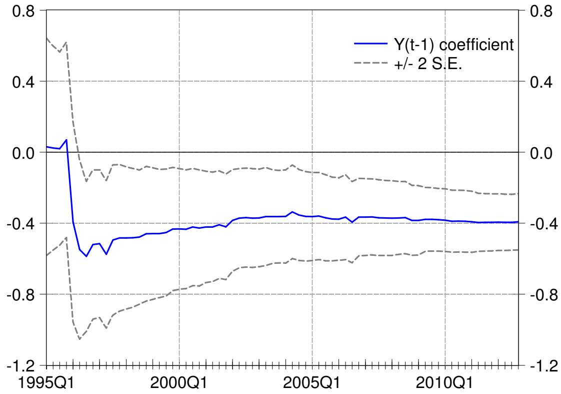


Figure A5: Recursive estimates of new MIBA M3 parameters (Q1 1995 - Q4 2012)

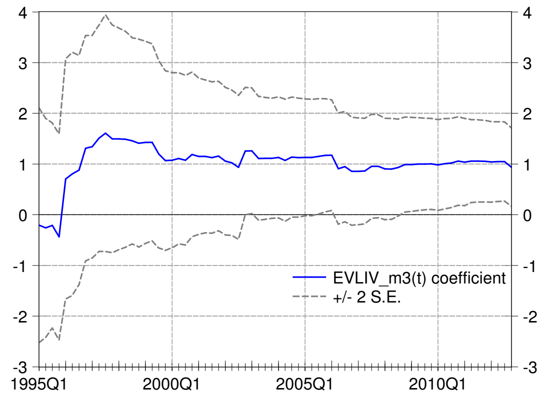
(a) Constant



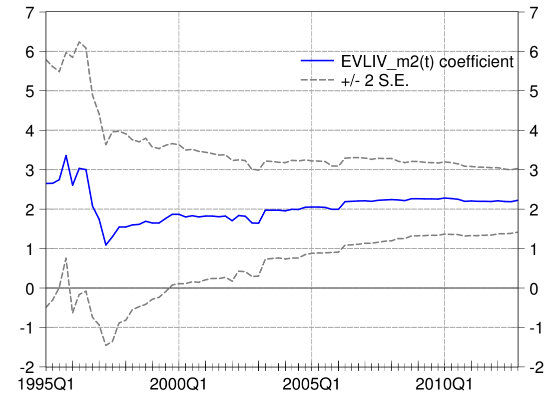
(b) Y_{t-1}



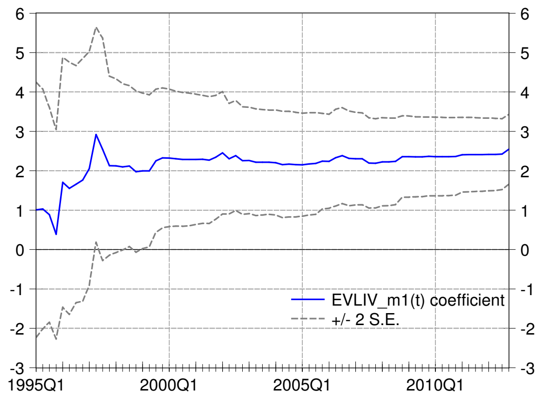
(c) $\text{EVLIV}_t^{(m3)}$



(d) $\text{EVLIV}_t^{(m2)}$



(e) $\text{EVLIV}_t^{(m1)}$



A.1 Data innovation and forecast revisions

This proof follows directly from Dubois and Michaux (2006). Suppose we have the following two consecutive forecasting equations:

$$Y_t = Z_t\boldsymbol{\gamma}_1 + X_t^{(1)}\boldsymbol{\delta}_1 + \varepsilon_t^{(1)} \tag{A-1}$$

$$Y_t = Z_t\boldsymbol{\gamma}_2 + X_t^{(2)}\boldsymbol{\delta}_2 + \varepsilon_t^{(2)} \tag{A-2}$$

where $Z_t = (1, Y_{t-1}, W_t)'$ is a vector of variables common to both equations, $X_t^{(1)} = (x_{1,t}^{(1)}, \ldots, x_{n,t}^{(1)})'$ is a vector of variables entering the first equation only, and $X_t^{(2)} = (x_{1,t}^{(2)}, \ldots, x_{m,t}^{(2)})'$ is a vector of variables entering the second equation only. Let us define $\hat{Y}_{1,t+h}$ the forecast of Y_{t+h} provided by the first equation and $\hat{Y}_{2,t+h}$ the forecast provided by the second equation. If $\hat{Y}_{2,t+h} - \hat{Y}_{1,t+h} \neq 0$, forecasts are revised, and the revision can be explained by both a change in the model (the absence of $X_t^{(1)}$ in Equation (A-2)) and the inflow of new information (the updated information represented by $X_t^{(2)}$). In order to disentangle these two factors, let us run the following auxiliary regression:

$$x_{j,t}^{(2)} = Z_t\mathbf{a} + X_t^{(1)}\mathbf{b} + \epsilon_{j,t}^{(2)}, \tag{A-3}$$

for $j = 1, \ldots, m$. The non-redundant information of each variable $x_{j,t}^{(2)}$ with respect to common variables Z_t and variables belonging exclusively to the first equation, $X_t^{(1)}$, is:

$$x_{j,t}^{(2)\perp} = x_{j,t}^{(2)} - \hat{x}_{j,t}^{(2)} = \hat{\epsilon}_{j,t}^{(2)}. \tag{A-4}$$

It is worth noticing that $X_t^{(2)\perp} = \left(x_{1,t}^{(2)\perp}, \ldots, x_{m,t}^{(2)\perp}\right)'$ is orthogonal to Z_t and $X_t^{(1)}$, by construction. It can then be shown that:

$$\begin{aligned} \hat{Y}_{1,t+h} &= Z_{t+h}\hat{\boldsymbol{\gamma}}_1 + X_{t+h}^{(1)}\hat{\boldsymbol{\delta}}_1 \\ &\approx Z_{t+h}\hat{\boldsymbol{\gamma}}_2 + (X_{t+h}^{(2)} - X_{t+h}^{(2)\perp})\hat{\boldsymbol{\delta}}_2. \end{aligned} \tag{A-5}$$

Hence, since $\hat{Y}_{2,t+h} = Z_{t+h}\hat{\boldsymbol{\gamma}}_2 + X_{t+h}^{(2)}\hat{\boldsymbol{\delta}}_2$, forecast revisions can be approximated by the following:

$$\begin{aligned} \hat{Y}_{2,t+h} - \hat{Y}_{1,t+h} &\approx X_{t+h}^{(2)}\hat{\boldsymbol{\delta}}_2 - (X_{t+h}^{(2)} - X_{t+h}^{(2)\perp})\hat{\boldsymbol{\delta}}_2 \\ &\approx X_{t+h}^{(2)\perp}\hat{\boldsymbol{\delta}}_2. \end{aligned} \tag{A-6}$$

This approximation represents the "inflow of new information" factor, while the residual unexplained revisions can be attributed to the "change in the model" factor.

Now, suppose that Y_{t+h-1} in Z_{t+h} is not observed when forecasting with the first equation, while it is when forecasting with the second equation. However, suppose that its estimate $\tilde{Y}_{t+h-1}$ (*e.g.*, a previous forecast) is available. This means that forecast revisions can be additionally explained by the update of $\tilde{Y}_{t+h-1}$ between the first and the second equation. Conditional expectations from Equations (A-1) and (A-2) can be written as:

$$\hat{Y}_{1,t+h} = \tilde{Z}_{t+h}\hat{\boldsymbol{\gamma}}_1 + X^{(1)}_{t+h}\hat{\boldsymbol{\delta}}_1 \tag{A-7}$$

$$\hat{Y}_{2,t+h} = \tilde{Z}_{t+h}\hat{\boldsymbol{\gamma}}_2 + X^{(2)}_{t+h}\hat{\boldsymbol{\delta}}_2 + (Y_{t+h-1} - \tilde{Y}_{t+h-1})\hat{\gamma}_{y,2} \tag{A-8}$$

where $\tilde{Z}_{t+h} = (1, \tilde{Y}_{t+h-1}, W_{t+h})'$. Following the results presented above, we have that:

$$\hat{Y}_{1,t+h} \approx \tilde{Z}_{t+h}\hat{\boldsymbol{\gamma}}_2 + (X^{(2)}_{t+h} - X^{(2)\perp}_{t+h})\hat{\boldsymbol{\delta}}_2 + (K_{t+h} - K^{\perp}_{t+h})\hat{\gamma}_{y,2}, \tag{A-9}$$

where $K_{t+h} = (Y_{t+h-1} - \tilde{Y}_{t+h-1})$. Hence, forecast revisions can be now approximated by the following:

$$\hat{Y}_{2,t+h} - \hat{Y}_{1,t+h} \approx X^{(2)\perp}_{t+h}\hat{\boldsymbol{\delta}}_2 + K^{\perp}_{t+h}\hat{\gamma}_{y,2}. \tag{A-10}$$

Documents de Travail

460. J. Carluccio and M. Bas, "The Impact of Worker Bargaining Power on the Organization of Global Firms," November 2013

461. F. Canova, F. Ferroni and C. Matthes, "Choosing the variables to estimate singular DSGE models," November 2013

462. S. Avouyi-Dovi, G. Horny and P. Sevestre, "The dynamics of bank loans short-term interest rates in the Euro area: what lessons can we draw from the current crisis?," November 2013

463. D. Fougère, C. Golfier, G. Horny et E. Kremp, "Quel a été l'impact de la crise de 2008 sur la défaillance des entreprises ?," Novembre 2013

464. T. Duprey, "Heterogeneous Banking Efficiency: Allocative Distortions and Lending Fluctuations," November 2013

465. L. Arrondel, M. Debbich and F. Savignac, "Financial Literacy and Financial Planning in France," December 2013

466. P. Antipa, "Fiscal Sustainability and the Value of Money: Lessons from the British Paper Pound, 1797-1821," December 2013

467. D. Bellas and Vincent Vicard, "French firms exports during downturns: evidence from past crises," December 2013

468. V. Chouard, D. Fuentes Castro, D. Irac and M. Lemoine, "Assessing the losses in euro area potential productivity due to the financial crisis," December 2013

469. G. Cléaud, M. Lemoine and P-A. Pionnier, "Which size and evolution of the government expenditure multiplier in France (1980-2010)?," December 2013

470. D. Irac and J. Lopez, "Euro area structural convergence? A multi-criterion cluster analysis," December 2013

471. L. De Benedictis, S. Nenci, G. Santoni, L. Tajoli and C. Vicarelli, "Network Analysis of World Trade using the BACI-CEPII dataset," December 2013

472. G. Gaulier, G. Santoni, D. Taglioni and S. Zignago, "Market Shares in the Wake of the Global Crisis: the Quarterly Export Competitiveness Database," December 2013

473. M. Mogliani, V. Brunhes-Lesage, O. Darné and B. Pluyaud, **"**New estimate of the MIBA forecasting model. Modeling first-release GDP using the Banque de France's Monthly Business Survey and the "blocking" approach," January 2014

Pour accéder à la liste complète des Documents de Travail publiés par la Banque de France veuillez consulter le site : www.banque-france.fr

For a complete list of Working Papers published by the Banque de France, please visit the website: www.banque-france.fr

Pour tous commentaires ou demandes sur les Documents de Travail, contacter la bibliothèque de la Direction Générale des Études et des Relations Internationales à l'adresse suivante :

For any comment or enquiries on the Working Papers, contact the library of the Directorate General Economics and International Relations at the following address :

BANQUE DE FRANCE
49- 1404 Labolog
75049 Paris Cedex 01
tél : 0033 (0)1 42 97 77 24 ou 01 42 92 63 40 ou 48 90 ou 69 81
email : 1404-ut@banque-france.fr