Centre de Recherche en Économie et Statistique

CREST

Série des Documents de Travail

n° 2013-21

Nowcasting French GDP in Real-Time from Survey Opinions : Information or Forecast Combinations ?

F. BEC[1] – M. MOGLIANI[2]

December, 2013

[1] THEMA, Université de Cergy-Pontoise and CREST–LMA. Email : bec@ensae.fr

[2] *Corresponding author :* Banque de France, 46-1383 DGEI-DCPM-DIACONJ, 31 rue Croix des Petits Champs, 75049 Paris Cedex 01 (France). Email : matteo.mogliani@banque-france.fr

Nowcasting French GDP in Real-Time from Survey Opinions: Information or Forecast Combinations?*

Frédérique Bec[1]

Matteo Mogliani[2]

December 3, 2013

Abstract: This paper investigates the predictive accuracy of two alternative forecasting strategies, namely the forecast and information combinations. Theoretically, there should be no role for forecast combinations in a world where information sets can be instantaneously and costlessly combined. However, following some recent works which claim that this result holds in population but not necessarily in small samples, our paper questions this postulate empirically in a real-time and mixed-frequency framework. An application to the quarterly growth rate of French GDP reveals that, given a set of predictive models involving coincident indicators, a simple average of individual forecasts outperforms the individual forecasts, as long as no individual model encompasses the others. Furthermore, the simple average of individual forecasts outperforms, or it is statistically equivalent to, more sophisticated forecast combination schemes. However, when a predictive encompassing model is obtained by combining information sets, this model outperforms the most accurate forecast combination strategy.

Keywords: Forecast Combinations; Pooling Information; Macroeconomic Nowcasting; Real-time data; Mixed-frequency data.

JEL classification: C22, C52, E37.

*We are indebted to Anindya Banerjee, Pierre Sicsic, Michel Juillard, Bertrand Pluyaud and the participants in the 13th OxMetrics User Conference (Aarhus, 5-6 September 2013) and 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]THEMA, Université de Cergy-Pontoise and CREST-LMA. Email: bec@ensae.fr

[2]*Corresponding author*, Banque de France, 46-1383 DGEI-DCPM-DIACONJ, 31 Rue Croix des Petits Champs, 75049 Paris Cedex 01 (France). Email: matteo.mogliani@banque-france.fr.

1 Introduction

Since the influential work by Bates and Granger (1969), forecasters are aware that combining forecasts obtained from two or more models can yield more accurate forecasts, in the sense that the forecast error variance of the combined forecasts is not larger than the smaller variance of the individual forecasts.[1] Nevertheless, as claimed by Diebold (1989), there is no role for forecast combinations in a world where information sets can be instantaneously and costlessly combined. Furthermore, under such assumptions "it is always optimal to combine information sets rather than forecasts" (Diebold, 1989, p. 590). Recently, this claim has been challenged by Huang and Lee (2010), who find some simulation evidence of forecast combinations superiority, compared to the pooled information strategy, whether the model combining the information is correctly specified or not. One rationale for this result can be found in the bias-variance trade-off between small and large models in finite samples.[2] Yet, the individual forecasting models used in forecast combinations are usually parsimoniously specified, whereas the model combining all the available and relevant information is typically large. Further, as noted by Diebold (1989, p. 591), "when the user of the forecasts is in fact the model builder, the possibilities for combination of information sets — as opposed to forecasts — are greatly enhanced". Hence, the debate about combining either forecasts or information sets might well be still open within forecasting institutions, such as Central banks and National statistical institutes, which are often both model builders and data collectors.

The forecast combination literature is strongly related to the encompassing paradigm described in Mizon (1984), Mizon and Richard (1986) and Hendry and Richard (1989), among others. As summarized in Chong and Hendry (1986, p. 677), "...the composite artificial model which might be considered for forecast encompassing essentially coincides with the 'pooling of forecasts' formula... Note that the *need* to pool forecasts is *prima facie* evidence of a failure to encompass, and if H_1 is an econometric model and H_2 a univariate time series model (say) then if H_1 does not encompass H_2 it seems highly suggestive of the possibility that H_1 is dynamically misspecified...". As noted by Diebold (1989), there is little room for forecast combinations within this encompassing paradigm. Basically, given an exhaustive information set, this approach aims at detecting the encompassing model in a way that is akin to the "general-to-specific" selection approach described in Krolzig and Hendry (2001): if a first-guess model (a general-unrestricted-model, for instance) turns out to be misspecified, the selection process is iterated following a reduction algorithm until a satisfactory in-sample alternative is found, from which optimal forecasts follow directly. Hence, the information combination strategy should always dominate the

[1] This result is derived by Bates and Granger (1969) under the assumptions that the individual forecast errors are stationary and that the forecasts are unbiased, and provided that they are not too strongly correlated. Moreover, the weights used in the combination are chosen so as to minimize the overall variance of errors in the combined forecasts. For comprehensive surveys on forecast combinations, the reader is referred to Clemen (1989) and Timmermann (2006).

[2] In econometrics, the well-known bias-variance *dilemma* refers to the trade-off arising from the inconclusive attempt of simultaneously reducing two estimation errors: the error due to the difference between the expected predictions of a model and the correct values (bias), and the error due to the variability of a model prediction for a given observation (variance).

forecast combination strategy. However, this would be true in the absence of the bias-variance trade-off between small and large models affecting the estimator in finite samples.[3]

This paper aims at comparing the two approaches, namely the forecast and the information combinations, by illustrating the particular case of nowcasting the *first release* of French GDP. A similar question is raised and explored empirically by Clements and Galvão (2006) for the US. These authors combine forecasts from simple models, each one including a single explanatory variable selected among the leading indicators composing the Conference Board Leading Economic Index, which is a mix of hard- and soft-data. Beside, they combine the whole information set by implementing a relatively simple model selection strategy. Their empirical findings suggest that pooling forecasts from single-indicator models leads to more accurate forecasts of the GDP than pooling the available information, although the results on the predictability of macroeconomic recessions appear less conclusive. Our approach departs from theirs along two dimensions. First, we are interested in predicting the first release of French GDP by using soft-data only. Indeed, the peculiarity of the French case is that, in addition to the Markit Purchasing Managers Index (PMI) survey, the National statistical institute (National Institute of Statistics and Economic Studies, INSEE hereafter) and the Central Bank of France (BDF hereafter) also collect their own survey data on business conditions in the manufacturing sector (Monthly business surveys). Yet, surveys are usually the earliest monthly-released data conveying information on the current quarter's economic activity, and thus they have often proven to be useful for nowcasting (Banbura et al., 2013). Second, in this paper we consider only a bunch of partial-information forecasting models, namely a model for each survey database. This strategy amounts to estimating only three restricted forecasting models (using either the INSEE, or the BDF, or the PMI data), plus a super-model exploiting all the available information conveyed by the three surveys. Nevertheless, the restricted forecasting models are specified in the same way as the full-information model, *i.e.*, by implementing the "general-to-specific" selection approach described above. As a result, careful attention is paid to the construction of the individual forecasting models within a reasonably restricted information environment. It follows that, by reducing the probability of selecting misspecified models, we deliberately set a strong prior in favor of the forecast combination strategy (*i.e.*, the restricted models are not penalized, compared to the full-information model), so that very strong results would be obtained in the event this strategy is dominated, in terms of predictive accuracy, by the information combination approach.[4]

[3]Boivin and Ng (2006) show that the forecasting performance of dynamic factor models deteriorates when highly correlated additional series are included in the information pool. These findings, based on simulation experiments, suggest to pay careful attention to the characteristic of the data used to build the information pool, and seem to postulate against the theoretical result in Diebold (1989). In the present contribution, we assume that each survey data does not provide an exhaustive information, so that there should be some gain in pooling them. However, we verify this assumption by selecting our full-information models accurately through the "general-to-specific" approach. In practice, this strategy should be akin to the one recommended by Boivin and Ng (2006), because the model selection is consistent with the exclusion of redundant information. It follows that the proposition in Diebold (1989) holds in our empirical application if (well-specified) full-information models outperform partial-information models in terms of predictive accuracy.

[4]By contrast, other studies, such as Clements and Galvão, 2006, and Huang and Lee, 2010, retain quite crude small forecasting models.

Preliminary results are broadly in line with the theoretical findings reported in Chong and Hendry (1986): when the individual models fail to encompass, there is a forecasting accuracy gain in combining them. Further, we provide additional support to the widely documented empirical evidence on the outperformance of simple forecast combination methods over more sophisticated ones, although we can find some noticeable exceptions. However, our main findings are clear-cut regarding the comparison of the two competing combination strategies: the full-information models encompass the restricted-information models, and the former outperform the latter in terms of forecast accuracy, whatever predictions are either taken individually or in combination.

The paper is organized as follows. Section 2 discusses various forecast combination approaches, with emphasis on the ones used in the application. In Section 3 we describe the real-time analysis and we explain how the the mixed-frequency issue is handled. Section 4 presents the empirical application on nowcasting French GDP: it describes the data and the modeling strategy, and reports empirical results for partial- and full-information models. Section 5 reports results on forecast encompassing tests and compares the predictive performance of the forecast combination *vs* the information combination strategy. Section 6 concludes.

2 Forecast combination schemes

The rationale for combining a set of forecasts relies on the existence of sizeable diversification gains. These gains are expected to become significant when a predictive model generating smaller forecast errors than its competitors cannot be found, as well as when forecast errors cannot be hedged by other models' forecast errors. This is usually the case when all the information sets used to compute the individual forecasts cannot be instantaneously and costlessly combined or when, due to the finite sample nature of the setup, the forecasting models may be considered as local approximations, subject to misspecification bias, so that it is unlikely that the same model outperforms all others at all points in time. Using the same notation as in Timmermann (2006), let us assume that at time t we wish to forecast the h-period ahead value of some variable denoted y, with $y \in \mathbb{R}$. Let us also assume that N different forecasts are available at time t, so that the information set available at that time includes $\hat{\mathbf{y}}_{t+h,t} = (\hat{y}_{t+h,t,1}, \hat{y}_{t+h,t,2}, \ldots, \hat{y}_{t+h,t,N})'$ on top of the history of these forecasts and of the realizations of y up to time t, as well as a set of additional information variables, $\mathbf{x}_t$, so that $\mathcal{F}_t = \{\hat{\mathbf{y}}_{t+1,1}, \ldots, \hat{\mathbf{y}}_{t+h,t}, y_1, \ldots, y_t, \mathbf{x}_t\}$. As emphasized in Timmermann (2006), most studies on forecast combinations deal with point forecasts, which amounts here to focusing on $\hat{y}^c_{t+h,t} = C(\hat{\mathbf{y}}_{t+h,t}; \boldsymbol{\omega}_{t+h,t})$, *i.e.*, the combined point forecast as a function of the N individual forecasts $\hat{\mathbf{y}}_{t+h,t}$ and the combination weights $\boldsymbol{\omega}_{t+h,t} \in \mathcal{W}_t$, a compact subset of $\mathbb{R}^N$:

$$\hat{y}^c_{t+h,t} = \sum_{i=1}^{N} \omega_{t+h,t,i} \hat{y}_{t+h,t,i}. \tag{1}$$

Note that the vector of weights $\boldsymbol{\omega}_{t+h,t}$ can be time-varying. Many forecast combination methods coexist, each one defining a different $\boldsymbol{\omega}_{t+h,t}$. Only a few of them is implemented in our empirical analysis. In what follows, we describe the methods which assume constant weights, and the we explain how this assumption can be relaxed.

2.1 Constant weights

The most simple weighting scheme assigns the same weight to each individual model, so that $\hat{y}^c_{t+h,t}$ is the simple average of the N available individual forecasts:

$$\hat{y}^c_{t+h,t} = \frac{1}{N}\sum_{j=1}^{N}\hat{y}_{t+h,t,j}, \tag{2}$$

This Naïve combination of individual forecasts ($\omega_{i,\text{Naïve}} = 1/N$) is often found in the empirical literature to outperform more sophisticated combination schemes. As discussed in Smith and Wallis (2009), *inter alia*, the so-called forecast combination puzzle stems from the finite-sample estimation error of combination weights. For this reason, the Naïve approach of forecast combinations is considered in the following empirical analysis, along with two additional simple combination methods belonging to the Mean Squared Forecast Errors (MSE)-based class of combination weights proposed by Bates and Granger (1969). The first one posits the combination weights as the inverse of relative MSE of the forecasts (Bates and Granger, 1969; Stock and Watson, 1999):[5]

$$\omega_{i,\text{IMSE}} = \frac{\sigma_i^{-2}}{\sum_j^N \sigma_j^{-2}} = \frac{\text{MSE}_i^{-1}}{\sum_j^N \text{MSE}_j^{-1}}. \tag{3}$$

The second one posits the weights as the inverse of the models' rank (Aiolfi and Timmerman, 2006):

$$\omega_{i,\text{IR}} = \frac{\mathcal{R}_i^{-1}}{\sum_j^N \mathcal{R}_j^{-1}}, \tag{4}$$

where $\mathcal{R}_i$ is the rank of model i as obtained in terms of absolute MSE.

Another widely used approach is the one proposed by Granger and Ramanathan (1984), which consists in estimating the combination weights by ordinary or constrained Least Squares. Basically, these authors propose three linear regressions to achieve this goal:

$$y_{t+h} = \omega_{0h} + \boldsymbol{\omega}'_h\hat{\mathbf{y}}_{t+h,t} + \varepsilon_{t+h}, \tag{5a}$$

$$y_{t+h} = \boldsymbol{\omega}'_h\hat{\mathbf{y}}_{t+h,t} + \varepsilon_{t+h}, \tag{5b}$$

$$y_{t+h} = \boldsymbol{\omega}'_h\hat{\mathbf{y}}_{t+h,t} + \varepsilon_{t+h}, \qquad \text{s.t.} \quad \sum_{j=1}^{N}\omega_{h,j} = 1. \tag{5c}$$

[5]The reader is referred to the Appendix for a short review of this approach.

These regressions differ by the way they account or not for a bias in the individual forecasts. Equation (5a) allows for biased forecasts, since the bias is accounted for by the intercept ω_{0h}. By contrast, Equations (5b) and (5c) assume the unbiasedness of individual forecasts, either spontaneously or after correction for the bias, if any. Equations (5a) and (5b) can be thus estimated by OLS, while Equation (5c) requires the constrained least squares estimator, due to the constraint that the weights sum to one. This constraint imposes the desirable property that the combined forecast is also unbiased, but it may lead to efficiency losses since the orthogonality property between the regressors in $\hat{\mathbf{y}}_{t+h,t}$ and the residuals ε_{t+h} may not hold. Nevertheless, as shown by Diebold (1988), relaxing this assumption as in Equation (5b) generates a combined forecast that is serially correlated and hence predictable. For these reasons, only Equation (5c), using bias-corrected individual forecasts, and Equation (5a) are considered in the empirical analysis below. Since they lead to very similar outcomes, only combination weights obtained from the constrained least squares estimate of Equation (5c) are used in practice ($\omega_{i,\mathrm{OLS}}$).

A growing number of studies implements Bayesian approaches to estimate the combination weights (see, *e.g.*, Jackson and Karlsson, 2004, and Eklund and Karlsson, 2007). From the Bayesian model averaging literature, forecast combinations can be obtained by averaging over individual forecasts and using posterior model probabilities as weights:

$$\hat{y}^c_{t+h,t} = \sum_{i=1}^{N} \hat{y}_{t+h,t,i} p(\mathcal{M}_i|y), \tag{6}$$

where

$$p(\mathcal{M}_i|y) = \frac{p(y|\mathcal{M}_i)p(\mathcal{M}_i)}{p(y)} = \frac{p(y|\mathcal{M}_i)p(\mathcal{M}_i)}{\sum_{j=1}^{N} p(y|\mathcal{M}_j)p(\mathcal{M}_j)} \tag{7}$$

is the posterior model probability $p(\mathcal{M}_i|y)$, which is proportional to the prior model probability $p(\mathcal{M}_i)$ *times* the marginal likelihood of the model $p(y|\mathcal{M}_i)$ (*i.e.*, the probability of the data given the model $\mathcal{M}_i$, once parameters $\boldsymbol{\theta}_i$ are integrated out),

$$p(y|\mathcal{M}_i) = \int_{\Theta} p(y|\boldsymbol{\theta}_i, \mathcal{M}_i)p(\theta_i|\mathcal{M}_i)\mathrm{d}\boldsymbol{\theta}_i. \tag{8}$$

In this work, we focus on the predictive likelihood $p(\tilde{y}|y^*, \mathcal{M}_i)$ of each individual model for the computation of the posterior model probability $p(\mathcal{M}_i|y)$. The predictive likelihood can be obtained by splitting the data into two parts, namely a hold-in (y^*) and a hold-out ($\tilde{y}$) sub-samples, and computing the posterior predictive density:

$$p(\tilde{y}|y^*, \mathcal{M}_i) = \int_{\Theta} p(\tilde{y}|y^*, \boldsymbol{\theta}_i, \mathcal{M}_i)p(\theta_i|y^*, \mathcal{M}_i)\mathrm{d}\boldsymbol{\theta}_i. \tag{9}$$

The predictive density in (9) is the distribution of future observations $\tilde{y}$ conditional on the observed sample y^*, and indicates how well the model predicts the realized observations. By replacing the marginal likelihood in (7) with the predictive likelihood in (9), we have a posterior model probability calibrated on the predictive performance of the models. The combination weights are then computed as:

$$\omega_{i,\text{PLIK}} = \frac{p(\tilde{y}|y^*, \mathcal{M}_i)p(\mathcal{M}_i)}{\sum_{j=1}^{N} p(\tilde{y}|y^*, \mathcal{M}_j)p(\mathcal{M}_j)}, \tag{10}$$

where the prior model probability is set to $p(\mathcal{M}_i) = 1/N$.[6]

2.2 Time-varying weights

As noted by Bates and Granger (1969) and Newbold and Granger (1974), one desirable property of the combination weights is that they should adapt quickly to new values if, for instance, there is a lasting change in the success of one of the forecasts. The simplest way to obtain adapting weights is proposed by Bates and Granger (1969). It basically consists in computing weights according to the inverse of MSE (IMSE) as in system (3) above, but instead of using the whole sample, we restrict it to the last υ observations. The resulting rolling window weights are given by:

$$\omega_{t,t-h,i}^{Rol} = \frac{\left(\sum_{\tau=t-\upsilon+1}^{t} e_{\tau,\tau-h,i}^2\right)^{-1}}{\sum_{j=1}^{N}\left(\sum_{\tau=t-\upsilon+1}^{t} e_{\tau,\tau-h,i}^2\right)^{-1}}. \tag{11}$$

An additional adapting weighting scheme can be obtained from Equation (11) by setting υ to t. In this case, an expanding window is used at each date and the corresponding recursive sample estimates of the weights are given by:

$$\omega_{t,t-h,i}^{Rec} = \frac{\left(\sum_{\tau=1}^{t} e_{\tau,\tau-h,i}^2\right)^{-1}}{\sum_{j=1}^{N}\left(\sum_{\tau=1}^{t} e_{\tau,\tau-h,i}^2\right)^{-1}}. \tag{12}$$

These adapting strategies can of course be applied to other weighting schemes. In particular, it is straightforward to compute rolling and recursive weights for the inverse rank in the MSE (IR), the OLS regression (OLS), and the Bayesian predictive likelihood methods (PLIK). So, instead of restraining the analysis to one particular *ad hoc* sample, three types of weighting samples are considered in this study:

- the full sample (denoted FS), which assumes that the weights are fixed at the same value for all forecasts,
- a sample rolling (denoted *Rol*) with $\upsilon = 8$ quarters in Equation (11), and
- a recursive sample (denoted *Rec*).

[6] In the present paper, the posterior distribution of parameters is estimated using the independent Normal-Gamma prior and the Gibbs sampler. We set informative priors on model parameters and run the Gibbs sampler 10'000 times, after discarding 1'000 burn-in replications.

3 Real-time forecasting of GDP from survey opinions

3.1 Dealing with revisions: Forecasting the first-release GDP

It is well known that historical GDP data are usually affected by frequent and sizeable revisions, usually up to two/three years after their first release. This implies that a “pseudo-true value” for the GDP at time t, and released at time $t+1$, is only observed in a farther vintage estimate, say, $t+\ell$. However, even though policy makers are ideally interested in this true value for policy purposes, the performance of short-term forecasting models is *de facto* evaluated by comparison with the corresponding first release of actual data. Further, as claimed by Croushore (2011, p. 90), “forecasts are affected by data revisions because the revisions change the data that are input into the model, the change in the data affects the estimated coefficients, and the model itself may change”.[7] Hence, if the aim of the forecaster is to predict first-release outcomes, efficient estimation of model parameters, leading to the optimal solution of the classic problem of minimizing the standard squared-error loss function, can be achieved by using the real-time matrix of GDP releases (*i.e.*, the revision triangle Y_t^{t+1}). This is the so-called Real-Time Vintage (RTV) estimation approach (Koenig et al., 2003; Clements and Galvão, 2013), which consists in matching early-released data (or revisions) by using the vector of observations from the main diagonal (y_t^{t+1}) as dependent variable, and the p adjacent diagonals ($y_{t-1}^t, \dots, y_{t-p}^t$) as lagged regressors, if any. This framework should ensure that predictions are optimal and unbiased, even in small samples.[8]

The models considered here aim at predicting the current quarter of the first-release GDP growth rate by using monthly information available up to the end of the quarter. This can be considered as nowcasting, rather than forecasting, because monthly data used here are released and available for econometric purposes about the end of each reference month. For instance, let us suppose we observe at the beginning of April 2012 the first release of the GDP growth rate for 2012Q1, and we are interested in predicting the first release of the GDP growth for 2012Q2.[9] For this aim, the latest and coincident news on the current quarter, progressively released at the end of April, May and June 2012 by different surveyors, is particularly valuable for predicting the target variable, as pointed out by Banbura et al. (2013). Indeed, survey data represent the earliest (monthly) information available on the current quarter's economic activity, when hard-data are usually not yet available to the forecaster.[10] Further, survey data are tipically subject to small and lump-sum revisions only, occurring at the time of the first revision, so that they can be considered as unrevised indicators, without loss of generality. As a result, an analysis

[7] We can assume the existence of two statistical relationships: one between the data used by the forecaster at time t and the data used by the National statistical agencies to estimate the first release of GDP at time $t+1$, and another one between the data used by the forecaster at time t and the data used by the National statistical agencies to compute the true value of GDP at time $t+\ell$.

[8] The RTV approach is opposed to the standard End-of-Sample estimation (EOS), which involves the most recent release of the GDP, *i.e.*, the latest column of the real-time matrix. Although frequently implemented by forecasters, predictions with the EOS approach can be non-optimal (in a squared-error loss sense) and biased (see Clements and Galvão, 2013).

[9] In practice, first estimates of the quarterly GDP are usually released about 45 days after the end of the quarter, depending on statistical agencies.

[10] See Banbura et al. (2011) for a review of the literature on this particular point.

combining the RTV estimation approach with survey data is consistent, by and large, with the actual real-time nowcasting exercise.

3.2 Dealing with mixed-frequency data: The blocking approach

The use of monthly survey data to nowcast quarterly real GDP in real-time raises two well-known issues: the mixed-frequency and the ragged edge data problems. The former refers to the use of variables sampled at different frequencies in the same econometric model, while the latter refers to the lack of complete information on predictors over the nowcasting period. These issues are clearly related in our case. Since we target a quarterly variable (the GDP) using predictors sampled at monthly frequency and released with almost no delay (the survey data), nowcasts can be released on a monthly basis. However, these predictors are progressively released by statistical agencies, so that at the beginning of each quarter we have partial information stemming from survey data, while complete information becomes available at the end of the quarter.

Various approaches have been proposed to deal with these issues. An intuitive solution to the mixed-frequency problem is to time-aggregate higher frequency data in order to match the sampling rate of lower-frequency data. Although easy to implement, this approach has the drawback of assigning the same weight to high-frequency observations across the low-frequency window, which could be non-optimal compared to a different weighting scheme. In addition, when the relevant information is not released simultaneously, time-aggregation does not solve the ragged edge data problem. With this respect, approaches such as MIDAS models (Clements and Galvão, 2008, 2009; Kuzin et al., 2011), bridge models (Rünstler and Sédillot, 2003; Baffigi et al., 2004; Diron, 2008), and the Kalman filter (Giannone et al., 2008) have been successfully implemented for nowcasting and forecasting quarterly GDP growth.

In this paper, we depart from these approaches and implement instead the so-called *blocking*, a technique originating from the engineering literature of signal processing and multirate sampled-data systems (see Chen, Anderson, Deistler, and Filler, 2012, for a recent review), but so far disregarded by the economic literature.[11] This method 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 of each quarter (January, April, July and October), the second one collects observations from the second months (February, May, August and November), while the last one collects the remaining observations from the third months (March, June, September and December). Compared to the approaches cited above, the *blocking* solution to mixed-frequency and ragged edge data problems has many advantages. First, the regression model with the distributed series is linear, which is convenient for the estimation of the predictive model by standard OLS techniques and the implementation of general-to-specific algorithms for model selection. Second, this approach allows the nowcaster to directly exploit the partially available

[11] See Carriero, Clark, and Marcellino (2012) for a very recent example. However, it is worth noticing that applications using the blocking approach have circulated in non-academic French circles since a few years. The (french-speaking) reader is referred to a few contributions issued by French national institutions (INSEE and the Ministry of Finance), such as Dubois and Michaux (2006), Bessec (2010), and Minodier (2010).

data at any time, with no need to extrapolate the missing information, unlike bridge models. Third, the *blocking* approach allows the forecaster to readily evaluate and interpret the signals provided by changes in the current economic activity in terms of GDP growth.

More formally, let's consider the following ARDL(p,d) class of models for the quarterly first-release GDP growth rate, denoted y_t^{t+1} hereafter:

$$y_t^{t+1} = \beta_0 + \sum_{j=1}^{p} \beta_j y_{t-j}^t + \sum_{l=0}^{d} \theta_l x_{t-l}^{(q)} + \epsilon_t, \tag{13}$$

where $x_t^{(q)}$ denotes the quarterly observations of the relevant information available at time t. For ease of exposition, only one explanatory variable is introduced in (13), but the subsequent analysis generalizes straightforwardly to the case where $x_t^{(q)}$ is in fact a vector of variables. Assume now that $x_t^{(q)}$ is released at higher frequency, say monthly (x_t), than the target variable y_t^{t+1}: hence, provided that the quarterly dependent variable and the coincident monthly information are linked through a regression model, the nowcast can be released at higher frequencies. In order to match the frequencies, the monthly series x_t is converted into a vector of 3 quarterly series $\mathbf{x}_t^{(m)} = \left(x_{t-\frac{2}{3}}^{(m)\prime}, x_{t-\frac{1}{3}}^{(m)\prime}, x_t^{(m)\prime} \right)'$, collecting respectively observations from the first, second and third month of each quarter. Let's also assume that the quarterly regressor $x_t^{(q)}$ in Equation (13) is an aggregation of the monthly variable x_t and can be written as a weighted average of the vector of *blocked* series, $x_t^{(q)} = \boldsymbol{\omega}'\mathbf{x}_t^{(m)}$, where $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)'$ is the vector of weights. Assuming $d = 0$ without loss of generality, we set the following three predictive equations:

$$y_t^{t+1} = \beta_0 + \sum_{j=1}^{p} \beta_j y_{t-j}^t + \gamma_1 x_{t-\frac{2}{3}}^{(m)} + \epsilon_{1,t} \tag{14a}$$

$$y_t^{t+1} = \beta_0 + \sum_{j=1}^{p} \beta_j y_{t-j}^t + \gamma_2 x_{t-\frac{1}{3}}^{(m)} + \gamma_1 x_{t-\frac{2}{3}}^{(m)} + \epsilon_{2,t} \tag{14b}$$

$$y_t^{t+1} = \beta_0 + \sum_{j=1}^{p} \beta_j y_{t-j}^t + \gamma_3 x_t^{(m)} + \gamma_2 x_{t-\frac{1}{3}}^{(m)} + \gamma_1 x_{t-\frac{2}{3}}^{(m)} + \epsilon_{3,t} \tag{14c}$$

Equations (14a) and (14b) model the dependent variable conditional on the partial information available at the end of the first and the second months of the quarter, while Equation (14c) is equivalent to Equation (13), with population parameters $(\gamma_1^*, \gamma_2^*, \gamma_3^*)' \equiv \boldsymbol{\gamma}^* = \theta_0\boldsymbol{\omega}$. It is nevertheless worth noticing that the missing information in Equations (14a) and (14b), *i.e.*, the second and/or the third month values for x_t, leads the parameters γ_1^* and γ_2^* to be potentially affected by an omitted-variable bias. For instance, let us suppose we dispose of survey information up to the second month of the quarter, $\tilde{\mathbf{x}}_t^{(m)} = \left(x_{t-\frac{2}{3}}^{(m)\prime}, x_{t-\frac{1}{3}}^{(m)\prime} \right)'$. From Equation (14b), population parameters $\tilde{\boldsymbol{\gamma}}^* = (\tilde{\gamma}_1^*, \tilde{\gamma}_2^*)'$ are given by:

$$\tilde{\boldsymbol{\gamma}}^* = \theta_0\tilde{\boldsymbol{\omega}} + \mathbf{b}$$

where $\tilde{\boldsymbol{\omega}} = (\omega_1, \omega_2)$ and $\mathbf{b} = (\tilde{\mathbf{x}}_t^{(m)\prime}\tilde{\mathbf{x}}_t^{(m)})^{-1}(\tilde{\mathbf{x}}_t^{(m)\prime}x_t^{(m)}\gamma_3)$ is the omitted-variable bias. The magnitude of the bias crucially depends on the weighting structure $\boldsymbol{\omega}$, because the more the weight attached to the information at the beginning of the quarter, the less the impact of omitting late monthly information. It follows that $\mathbf{b} \to \mathbf{0}$ when $\omega_3 \to 0$. However, if the forecaster has a diffuse prior on the weighting structure $\boldsymbol{\omega}$, such as equally distributed weights, a simple way to attenuate the bias consists in including in the regression model an additional coincident indicator, say $k_{t-\frac{1}{3}}^{(m)}$, but strongly correlated with $x_t^{(m)}$. This indicator, if available, would play the role of a short-term leading indicator for the monthly survey data and it would partially replace the missing information, leading to a reduction of the bias. We shall see in Section 4.3 how this solution naturally applies to our selected predictive models.

4 Nowcasting French real GDP

4.1 Data

In this work, we employ a set of survey data released by three different sources: the French National Statistical Institute (INSEE), the Bank of France (BDF) and the Markit Purchasing Managers Index (PMI). These three surveys share important features: they collect managers and entrepreneurs' opinions about past and expected activity; they sample a large share of representative firms in the manufacturing industry, according to the NAF-NACE rev. 2 nomenclatures (*i.e.*, food products, coke and refined petroleum products, machinery and equipment, transport equipment, and other manufactured products); they are released on a monthly basis. However, they differ on several aspects.

First, monthly release dates do not coincide: by the end of the current month for the INSEE and PMI surveys, and in the first week of the following month for the BDF survey. This means that a fair comparison between predictive models cannot be made until the latest survey (*i.e.*, the BDF survey) is released, which could be an issue for a real-time forecaster. Second, questions entering the survey sheet are usually very close across sources, but remarkable differences are worth noticing. For instance, managers are asked by the INSEE to supply opinions about both the evolution and the expectation of production over the past and following three months, which is a tantamount for a quarter-on-quarter change. On the other hand, both the BDF and PMI surveys are designed to collect opinions about the month-on-month evolution of past and expected activity. A trade-off between horizon and precision of the opinions can be then a relevant issue. Further, the number of answer modalities differs across surveys. For instance, the BDF survey proposes seven modalities, while the INSEE and PMI surveys only allow for three modalities. This feature is not trivial, because the broadest variety of responses can result in a better picture of the current economic activity. Third, the sample of surveyed firms is different across data collectors: around 9000 firms in the BDF survey, 4000 firms in the INSEE survey, and 400 firms in the PMI survey.

A synthetic, or composite, index conveying valuable information for the evaluation of the current business cycle is released by the three surveyors. However, in this work we consider the information stemming directly from disaggregated data (*i.e.*, the pool of opinion balances). This is more attractive than the composite indicators, because the resulting econometric models have a clear economic interpretation in terms of the evolution of current and expected activity, firms' stocking behavior and market tensions. The BDF survey includes 14 balances of opinion, the INSEE survey includes 7 balances, while the PMI survey only includes 4 balances. However, some of these time series are constructed in such a way that they are either unfit for our modeling purpose (potential non-stationarity, trending behaviour of the series) or expected not to convey any useful information about the supply-side of the current economic activity (prices, evolution of employment). For this reason, we consider a subset of main balances of opinions from the BDF survey (6 series) and the INSEE survey (5 series). As for the PMI survey, we do not discard any series. A list and short description of these variables can be found in Table 1.

[Table 1 about here]

4.2 Models selection

Compared to a large model combining all the available information, each small model considered in our empirical application is constrained to be sourced from a restricted information set, *i.e.*, a single survey dataset. This is consistent with the fact that business climate indicators, usually used by analysts for the conjunctural diagnosis of the current state of the economy and short-run forecasts of supply aggregates (for instances, GDP and IPI), are based on individual surveys.[12] Nevertheless, small and large models are selected through the general-to-specific approach discussed in Krolzig and Hendry (2001). It is worth noticing that this selection strategy is at odd with the approach followed in Clements and Galvão (2006), who arbitrarily restrict the information set of the small models and the way this information is chosen: their small models are constrained to include one and only one explanatory variable. As mentioned in the Introduction, we claim that our modeling strategy does not penalize the small models and does favor indeed the forecast combination strategy, so that meaningful conclusions can be drawn from the comparison exercise with the information combination strategy.

The general-to-specific model selection is achieved through the *Autometrics* algorithm (Doornik, 2009). As summed up in Hendry and Nielsen (2007, p. 292), the general, unrestricted model is "tested for misspecification, usually that the residuals are well behaved in terms of being approximately normal, homoskedastic and independent, and that the parameters are constant. If so, the stage of model reduction commences. Since all 2^N paths cannot be investigated in reasonable time for a large N (number of explanatory variables), even on a computer, any algorithm would have to focus on some of these paths only. Such a subselection of paths can be chosen efficiently though. The algorithm checks whether the least significant variable can be eliminated given a critical value c_α and if so, whether the simplified equation remains congruent." The last step is re-iterated until either no additional variables can be eliminated or a diagnostic test rejects the hypothesis of misspecification.

[12] Hence, for the case of France, three business climate indicators, referring to the three surveys, are available on a monthly basis.

We only impose three important constraints to the model selection approach. *First*, the selected models should be economically, rather than purely statistically, interpretable. We are hence allowed to appeal to the "expert opinion" of the forecaster on the consistency of the automatically selected equations, mainly when the selected indicators enter the models with odd and unexpected signs, or when their economic interpretation appears not trivial. *Second*, each model should embed some indicators conveying the most recent information available on the current quarter, whenever possible. For instance, the second-month equations are supposed to include balances of opinions on the second month of the quarter, among others. This constraint reflects the preferences of both professional forecasters and conjunctural analysts, who are obviously reluctant to discard coincident information over the quarter under review. We nevertheless relax this restriction if no coincident indicators are found statistically significant in our regression models, or when honoring the constraint comes into significant detriment of both in-sample and out-of-sample performances. *Third*, variables entering the regressions should display some parameter stability and statistical significance over time. We checked for this requirement through a battery of recursive regressions and Chow tests over a window spanning from 2005 to the end of the sample. When a balance of opinion is found to accomodate a particular feature of the regression models, we test whether the presence of this series in our models is robust to the introduction of simple dummy variables, supposed to capture the same feature. If it is not the case, we replace the series with the dummy variables, which have clearly no impact on the forecast results.[13] In other words, we limit the automatic selection of *ad hoc* variables by penalizing those indicators which appear strongly correlated with few outlier observations.

4.3 The restricted- and full-information models

The variables retained in each final restricted-information model, as well as in the full-information models, are reported in Tables 2 to 5. Looking at the estimation results, it can be noticed that the right-hand side variables for the first- and second-month equations (M1 and M2, hereafter) are often related to expected activity indicators, along with past activity indicators, in contrast with the third-month equations (M3, herafter). This is particularly the case for the restricted-information models using BDF and INSEE survey data, although the M3 equation using INSEE data also includes expectations (personal and general) surveyed during the quarter (but not during the third month). As for the PMI models, coincident variables hardly enter the selected equations, and the M2 model appears the best specification also for the third-month equation. These findings may be explained by the potential omitted-variable bias issue in the first- and second-month equations raised in Section 3.2: the expected activity variables may be considered as leading indicators for the second- and/or third-month survey data and hence their presence amongst the explanatory variables may attenuate the possible bias. Together with the finding that coincident third-month survey data hardly contribute to explain the GDP growth rate in the third-month equations, this result gives support to the use of the *blocking* approach,

[13] Additional dummy variables can also reveal necessary to obtain well specified regressions, in terms of constant parameters and good residuals diagnostic. The joint selection of indicators and dummy variables is carried out through the saturated regression approach described by Santos, Hendry, and Johansen (2008).

compared to the simpler (equally weighted monthly releases) time-aggregation solution to the mixed-frequency issue. Quite surprisingly, the fit of the models does not generally improve as we get closer to the end of the quarter. The restricted-information model relying on INSEE survey data is the only one displaying a R^2 (resp. σ_ϵ) steadily increasing (resp. decreasing) across monthly equations. As for the full-information models, the selected equations include at least one indicator from each survey, with a rather dominant presence of BDF and INSEE balances of opinions. It can be noticed again that the right-hand side variables for the first- and second-month equations are often related to expected activity indicators, and that this feature disappears for the third-month equation. It is also worth noticing that most of the variables entering the restricted-information models are recurrently selected for the full-information models, meaning that only a few balances of opinions convey fundamental information on the current state of the economy. This can also be interpreted as some evidence of complementarity of the survey data used here, rather than their substitutability, which justifies the information combination approach.

[Tables 2 to 5 about here]

The root mean squared forecast error (RMSE) and mean average forecast error (MAE) criteria are chosen to evaluate the predictive accuracy of our models. To compute these criteria, one-step ahead forecasts are obtained from a real-time analysis based on recursive regressions over a hold-out sample spanning from T_1=2003Q1 to T_f=2012Q4. Actual forecasts evaluation starts in T_e=2005Q1, while a burn-in sample of the first 8 predictions is used for the initialization of rolling and recursive combination weights (see Section 2). Then, for all $t \in \{T_0, ..., T_f - 1\}$, we estimate the models over a hold-in sample spanning from the initial observation (T_i=1992Q1 for the BDF and INSEE models and T_i=1998Q2 for the PMI and the full-information models) to t. We then use these estimates to compute one-step-ahead forecasts (nowcasts), leading to eight years of quarterly real-time predictions for out-of-sample evaluation (32 observations). Further, to determine whether the predictive accuracy of the full-information models is statistically equivalent to the accuracy of the restricted-information models, we compute bootstrap standard errors for the sample RMSE and MAE (see also Stock and Watson, 2002, 2007, and Clark and McCracken, 2006).[14]

[Table 6 about here]

Results are reported in Table 6. As expected, the full-information models systematically outperforms the most accurate restricted-information models (the M2 BDF model, and the

[14]Since the diagnostic tests reported in the bottom panels of Tables 2 to 5 do not point out major residual misbehaviors, the bootstrap algorithm implemented here is quite simple (see White, 2000, for an illustration of alternative bootstrap algorithms based on block resampling). We first estimate the forecasting models over the full sample. Estimated parameters and resampled residuals are then used to simulate $B = 10,000$ artificial samples. For each artificial sample b_i, with $i \in (1, \ldots, B)$, we compute recursive forecasts and prediction errors. Finally, we compute the empirical standard errors of the B simulated RMSE and MAE criteria.

M1 and M3 INSEE model, according to the RMSE criterion).[15] Indeed, based on the RMSE criterion, the maximum average predictive gains obtained within the set of restricted-information models range between 2% and 6%, while the predictive gains from the full-information models, compared to the best restricted-information models, range between 11% and 28%. Findings based on the MAE criterion are quantitatively similar, although less remarkable for the M3 equations. All in all, a substantial gain appears from the information combination strategy. Figure 1, plotting the prediction errors from the different models, provides a visual illustration of this result. Standard errors reported in Table 6 are rather narrow, ranging from 0.02 to 0.04 for both RMSE and MAE evaluation criteria. Although they should be interpreted with care, because in some cases their magnitude could be affected by the short length of both the estimation and the evaluation period, these figures are broadly in line with findings reported elsewhere (see, for instance, Stock and Watson, 2002, for an empirical analysis on US data), and suggest that the predictive gains from the full-information models are statistically relevant.

[Figure 1 about here]

The overall good performance of the full-information model is partly due to a remarkable predictive accuracy during the *Great Recession* episode (mainly 2008Q4 and 2009Q1 for France), when it clearly outperforms its best restricted-information competitors. It is worth noticing from Table 6 and Figure 1 that a peak in the predictive performance of the full-information model, as well as for the BDF model, coincides with the second-month exercise.[16] Unexpectedly, the additional information conveyed by the latest survey data over each quarter seems to deteriorate, although quite slightly, the accuracy of the M3 models, compared to the M2 models. This evidence reinforces the finding pointed out earlier in this Section, that coincident third-month survey data hardly contribute to explain, and also predict, the GDP growth rate in the third-month equations.

5 Information or forecast combinations?

5.1 Preliminary forecast encompassing tests

Chong and Hendry (1986) developed encompassing tests for forecast combinations.[17] These tests allow the researcher to determine whether a specific forecast incorporates all the relevant information contained in the other competing forecasts. An encompassing test on, say, forecast 1 relies on the joint null hypothesis that $\beta_0 = 0$, $\beta_1 = 1$ and $\beta_2 = \cdots = \beta_N = 0$ in:

$$y_{t+h} = \beta_0 + \sum_{i=1}^{N} \beta_i \hat{y}_{t+h,t,i} + e_{t+h,t}. \tag{15}$$

[15] We report evaluation criteria based on prediction errors corrected for a mean-bias, if any, since unbiased forecasts are used in the forecasts combination exercise. Unbiased forecasts are computed as $\hat{\tilde{y}}_{t+h,t} = y_{t+h}^{t+h+1} - (\hat{e}_{t+h} - \mathrm{E}(\hat{e}_{t+h}))$, where $\hat{e}_{t+h}$ is the forecast error from biased forecasts. Uncorrected RMSE and MAE are available upon request from the authors.

[16] As emerges from Tables 2 and 5, the M3 full-information model strongly relies on the M3 BDF model.

[17] See also Diebold and Lopez (1996) and Timmermann (2006) on this point.

If the joint hypothesis is not rejected, then forecast 1 encompasses its competitors and forecast combinations is not appropriate. However, this testing procedure is quite restrictive on the auxiliary specification employed, since it requires unbiasedness ($\beta_0 = 0$) and efficient forecasts ($\sum_{i=1}^{N} \beta_i = 1$). Fair and Schiller (1989) propose a looser version of the previous procedure, involving only a joint null hypothesis that $\beta_2 = \cdots = \beta_N = 0$ in Equation (15), with β_1 supposed to be positive. We report in the top panel of Table 7 the Wald test statistics for the latter joint null hypothesis of each restricted-information model encompassing the others, so as to decide whether these models should be combined or not.[18]

[Table 7 about here]

Test results strongly suggest that the small models are not encompassing each other, since the null is rejected across models and quarterly-equations, and they should be combined together. According to Chong and Hendry (1986), the statistical need for combination probably points to some models misspecification. However, and somewhat expectedly, when the model combining the information sets is included in the forecast encompassing regression (15), the hypothesis of encompassing ($\beta_1 = \cdots = \beta_3 = 0$ and $\beta_4 > 0$) cannot be rejected for this model at any conventional level (see the bottom line of Table 7). These findings are in line with the evidence of superior predictive accuracy for the full-information models discussed in Section 4.3, and suggest an expected superior predictive accuracy even over combined forecasts.

5.2 Are pooled forecasts more accurate than individual forecasts?

As stressed in the previous section, encompassing tests on the restricted-information models clearly suggest that their forecast combinations should outperform individual forecasts. This result is expected to hold almost surely in population, but it might not be the case in small samples. In the latter, the omission of one or more forecasts might still lead to outperforming predictions, even if the coefficients on the omitted forecasts in regression (15) are asymptotically non zero, but there is no clear theoretical guidance on the expected outcome.

Empirical evidence on the superior forecast accuracy of the forecast combination strategy, compared to uncombined individual forecasts, is reported in Table 8. The evaluation of out-of-sample predictions (RMSE and MAE) confirms the recommendation drawn from the encompassing tests, that is combining the individual forecasts systematically yields to some predictive gain, regardless of the weighting scheme implemented. Indeed, based on either the RMSE or the MAE criterion, the maximum average predictive gains from the best combined forecasts, compared to the best restricted-information models, range between 8% and 13%. Table 8 also reveals that, unlike the restricted-information models, the accuracy of the M3 combined forecasts do improve, compared to the M2 combined forecasts. This result suggests that a rationale for taking fruitfully advantage of the information conveyed by the latest survey data can be here provided by the forecast combination strategy. It is also worth noticing that our results are broadly in line with the literature on forecast combinations, because the Naïve approach, which

[18]Test results based on unbiased forecasts are qualitatively similar and available upon request from the authors.

assigns the same fixed weight to each model, seems to outperform quite often the more sophisticated approaches considered here. In addition, standard errors seem to strongly suggest that the predictive accuracy obtained through the Naïve weighting scheme is statistically equivalent to the best predictive accuracy observed across combination methods, models and evaluation criteria.[19] A visual illustration of these findings is provided by Figures 2, 3 and 4, which plot the forecast errors obtained from each combination method.

[Table 8 about here]

Nonetheless, following Diebold (1989), the predictions stemming from the full-information models are expected to outperform the best combined forecasts. Although this theoretical result has been recently recalled into question by Huang and Lee (2010), our empirical results are quite clear-cut: compared to the best combined forecasts challengers, the full-information models display a systematic superior predictive performance, with gains ranging between 4% and 17% according to RMSE. These results are thus consistent with the encompassing tests on the whole set of models (restricted- plus full-information models).[20] Figure 5 compares the forecast errors obtained from the full-information model and the best combined forecasts challenger (here the Naïve weighting scheme, for ease of exposition). From a visual inspection of these plots, we can note that the two series are strongly correlated up to the *Great Recession* episode, but some differences arise thereafter. These intuitive findings are somewhat consistent with the magnitude of the standard errors, which seem to suggest a statistical equivalence between predictions.

[Figure 2 to 5 about here]

In light of these results, we can hence conclude that, consistently with Diebold (1989), the information combination strategy outperforms in terms of predictive accuracy the forecast combination strategy. However, in our empirical empirical application, this superior performance is not as strong as expected, and the two strategies appear almost equivalent in statistical terms.

6 Concluding remarks

This paper investigates empirically the accuracy of two alternative forecasting strategies, namely the forecast combinations and the information combinations. Theoretically, there should be no role for forecast combinations in a world where information sets can be instantaneously and costlessly combined. However, following some recent works (Clements and Galvão, 2006; Huang and Lee, 2010), which bring into question this result in small samples, we investigate whether the theoretical predictions hold for the nowcast of the quarterly French GDP. Based on this application, we show that when the model combining the information sets is carefully specified -

[19]For ease of computation, we only report standard errors for the Naïve approach. However, we expect these standard errors to be strongly representative of the order magnitude that should be observed across weighting schemes and windows.

[20]It is worth noticing that, according to MAE criterion, combined forecasts from the M1 models appear somewhat more accurate than those from the full-information model.

here using a general-to-specific strategy and the unconstrained weighting scheme for regressors given by the *blocking* approach - a forecast encompassing model, yielding optimal forecasts, can be found. This result holds even if the restricted-information models are specified as carefully as the full-information, encompassing, models.

This empirical study also reveals that, given a set of predictive models involving coincident survey data, a simple average of individual forecasts outperforms individual forecasts, as long as no individual model encompasses the others. Furthermore, the simple average of individual forecasts outperforms, or it is statistically equivalent to more sophisticated forecast combination schemes. Expectedly, when a predictive encompassing model is obtained by combining information sets, this model outperforms the most accurate forecast combination strategy.

In main economic institutions, such as Central banks, where information is available and the forecaster is also the model builder, it would be hence worth raising the question of combining forecasts *versus* combining the information. In a finite sample world, we believe that this question can be only answered empirically.

References

Aiolfi, M., Timmerman, A., 2006. Persistence of forecasting performance and combination strategies. Journal of Econometrics 135 (1-2), 31–53.

Baffigi, A., Golinelli, R., Parigi, G., 2004. Bridge models to forecast the Euro Area GDP. International Journal of Forecasting 20 (3), 447–460.

Banbura, M., Giannone, D., Modugno, M., Reichlin, L., 2013. Now-casting and the real-time data flow. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting, Volume 2 (forthcoming). Amsterdam (The Netherlands): North-Holland.

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.

Bates, J. M., Granger, C. W. J., 1969. The combination of forecasts. Operational Research Quarterly 20 (4), 451–468.

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

Boivin, J., Ng, S., 2006. Are more data always better for factor analysis? Journal of Econometrics 132 (1), 169–194.

Carriero, A., Clark, T. E., Marcellino, M., 2012. Real-time nowcasting with a Bayesian mixed frequency model with stochastic volatility. Working Paper 12-27, Federal Reserve Bank of Cleveland.

Chen, W., Anderson, B. D. O., Deistler, M., Filler, A., 2012. Properties of blocked linear systems. Automatica 48 (10), 2520–2525.

Chong, Y. Y., Hendry, D. F., 1986. Econometric evaluation of linear macro-economic models. Review of Economic Studies 53 (4), 671–690.

Clark, T. E., McCracken, M. W., 2006. The predictive content of the output gap for inflation: Resolving in-sample and out-of-sample evidence. Journal of Money, Credit, and Banking 38 (5), 1127–1148.

Clemen, R. T., 1989. Combining forecasts: A review and annotated bibliography. International Journal of Forecasting 5 (4), 559–583.

Clements, M. P., Galvão, A. B., 2006. Combining predictors and combining information in modelling: Forecasting us recession probabilities and output growth. In: Milas, C., Rothman, P., van Dijk, D. (Eds.), Nonlinear Time Series Analysis of Business Cycles. Amsterdam (The Netherlands): Elsevier, pp. 55–74.

Clements, M. P., Galvão, A. B., 2008. Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States. Journal of Business and Economic Statistics 26 (4), 546–554.

Clements, M. P., Galvão, A. B., 2009. Forecasting US output growth using leading indicators: An appraisal using MIDAS models. Journal of Applied Econometrics 24 (7), 1187–1206.

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.

Diebold, F. X., 1988. Serial correlation and the combination of forecasts. Journal of Business and Economic Statistics 6 (1), 105–111.

Diebold, F. X., 1989. Forecast combination and encompassing: reconciling two divergent literatures. International Journal of Forecasting 5 (4), 589–592.

Diebold, F. X., Lopez, J. A., 1996. Forecast evaluation and combination. In: Maddala, G. S., Rao, C. R. (Eds.), Statistical Methods in Finance (Handbook of Statistics, Volume 14). Amsterdam (The Netherlands): North-Holland, pp. 241–268.

Diron, M., 2008. Short-term forecasts of Euro Area real GDP growth: an assessment of real-time performance based on vintage data. Journal of Forecasting 27 (5), 371–390.

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.

Eklund, J., Karlsson, S., 2007. Forecast combination and model averaging using predictive measures. Econometric Reviews 26 (2-4), 329–363.

Fair, R. C., Schiller, R. J., 1989. The informational content of ex ante forecasts. The Review of Economics and Statistics 71 (2), 325–331.

Giannone, D., Reichlin, L., Small, D., 2008. Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics 55 (4), 665–676.

Granger, C. W. J., Ramanathan, R., 1984. Improved methods of combining forecasts. Journal of Forecasting 3 (2), 197–204.

Hendry, D. F., Nielsen, B., 2007. Econometric Modeling: A Likelihood Approach. Princeton (USA): Princeton University Press.

Hendry, D. F., Richard, J. F., 1989. Recent developments in the theory of encompassing. In: Cornet, B., Tulkens, H. (Eds.), Contributions to Operations Research and Economics. The XXth Anniversary of CORE. Cambridge (USA): MIT Press, pp. 393–440.

Huang, H., Lee, T. H., 2010. To combine forecasts or to combine information. Econometric Reviews 29 (5-6), 534–570.

Jackson, T., Karlsson, S., 2004. Finding good predictors for inflation : A Bayesian model averaging approach. Journal of Forecasting 23 (7), 479–496.

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.

Kuzin, V., Marcellino, M., Schumacher, C., 2011. MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the Euro Area. International Journal of Forecasting 27 (2), 529–542.

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.

Mizon, G., Richard, J. F., 1986. The encompassing principle and its application to testing non-nested hypotheses. Econometrica 54 (3), 657–678.

Mizon, G. E., 1984. The encompassing approach in econometrics. In: Hendry, D. F., Wallis, K. F. (Eds.), Econometrics and Quantitative Economics. Oxford (UK): Blackwell, pp. 135–172.

Newbold, P., Granger, C. W. J., 1974. Experience with forecasting univariate time series and the combination of forecasts. Journal of the Royal Statistical Society Series A 137 (2), 131–146.

Rünstler, G., Sédillot, F., 2003. Short-term estimates of Euro Area real GDP by means of monthly data. Tech. rep., European Central Bank Working Paper Series N. 276.

Santos, C., Hendry, D. F., Johansen, S., 2008. Automatic selection of indicators in a fully saturated regression. Computational Statistics 23 (2), 317–335.

Smith, J., Wallis, K. F., 2009. A simple explanation of the forecast combination puzzle. Oxford Bulletin of Economics and Statistics 71 (3), 331–355.

Stock, J. H., Watson, M. W., 1999. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In: Engle, R. F., White, H. (Eds.), Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger. Oxford (UK): Oxford University Press, pp. 1–44.

Stock, J. H., Watson, M. W., 2002. Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics 20 (2), 147–162.

Stock, J. H., Watson, M. W., 2007. Why has U.S. inflation become harder to forecast? Journal of Money, Credit and Banking 39 (S1), 3–33.

Timmermann, A., 2006. Forecast combinations. In: Elliott, G., Granger, C. W. J., Timmermann, A. (Eds.), Handbook of Economic Forecasting, Volume 1. Amsterdam (The Netherlands): North-Holland, pp. 135–196.

White, H., 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126.

Appendix

The MSE-based class of combination weights

Let us consider the simple example presented in Bates and Granger (1969), where two different individual forecasts of y are available. These forecasts generate the unbiased forecast errors $e_1 = y - \hat{y}_1$ and $e_2 = y - \hat{y}_2$, whose variances are respectively σ_1^2 and σ_2^2 and where $\sigma_{12} = \rho_{12}\sigma_1\sigma_2$ denotes the covariance between e_1 and e_2, with ρ_{12} their correlation. Assuming that the vector of weights is $\boldsymbol{\omega} = (\omega, 1-\omega)$, the forecast error from the corresponding combination of $\hat{y}_1$ and $\hat{y}_2$ is:

$$e^c = \omega e_1 + (1-\omega)e_2, \tag{A-1}$$

which has zero mean and the following variance:

$$\sigma_c^2(\omega) = \omega^2\sigma_1^2 + (1-\omega)^2\sigma_2^2 + 2\omega(1-\omega)\sigma_{12}. \tag{A-2}$$

The optimal combination weights $\boldsymbol{\omega}^*$ are then found as the ones minimizing the objective function (A-2), which gives:

$$\omega^* = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}, \tag{A-3a}$$

$$1 - \omega^* = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}. \tag{A-3b}$$

By substituting ω^* in (A-2), we obtain the following optimal combination forecast error variance:

$$\sigma_c^2(\omega^*) = \frac{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)}{\sigma_1^2 + \sigma_2^2 - 2\rho_{12}\sigma_1\sigma_2}. \tag{A-4}$$

It is worth noticing that $\sigma_c^2(\omega^*) \leq min(\sigma_1^2, \sigma_2^2)$: this result grounds the diversification argument usually put forward in this literature (See the Appendix in Bates and Granger (1969) for a proof). Here, the strict equality holds if and only if *i)* $\sigma_1 = 0$ or $\sigma_2 = 0$, *i.e.*, one individual forecast clearly outperforms the other one, *ii)* $\sigma_1 = \sigma_2$ and $\rho_{12} = 1$, *i.e.*, the individual forecasts are equally good and highly (perfectly) correlated, or *iii)* $\rho_{12} = \sigma_1/\sigma_2$. It can also be seen from (A-3) that by setting ρ_{12} to zero, we obtain a combination scheme that weighs the forecasts inversely to their relative MSE:

$$\omega_{\sigma_2} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}, \tag{A-5a}$$

$$1 - \omega_{\sigma_1} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}. \tag{A-5b}$$

If $\rho_{12} = 0$, then ω_{σ_2} and $(1 - \omega_{\sigma_1})$ are optimal weights. Yet, due to difficulties in precisely estimating σ_{12} from short samples of time series, many authors such as Bates and Granger

(1969), Newbold and Granger (1974), Stock and Watson (1999), Aiolfi and Timmerman (2006) or Smith and Wallis (2009), simply recommend to ignore correlations across forecast errors and to express the weights in the following alternative form:

$$\omega_{\text{IMSE}} = \frac{\sigma_1^{-2}}{\sigma_1^{-2} + \sigma_2^{-2}}, \tag{A-6a}$$

$$1 - \omega_{\text{IMSE}} = \frac{\sigma_2^{-2}}{\sigma_1^{-2} + \sigma_2^{-2}}. \tag{A-6b}$$

Note that the latter is also optimal if $\sigma_1 = \sigma_2$, in which case $\omega_{\text{IMSE}} = \omega^* = 1/2$. This implies that the Naïve forecast combination scheme is optimal when the two forecasts are equally good, no matter how correlated they are.

Generalizing Equations (A-6a) and (A-6b) to N forecasts, we obtain the following expression for the combination weights:

$$\omega_{i,\text{IMSE}} = \frac{\sigma_i^{-2}}{\sum_j^N \sigma_j^{-2}} = \frac{\text{MSE}_i^{-1}}{\sum_j^N \text{MSE}_j^{-1}}. \tag{A-7}$$

Table 1: Survey Data

Name	Label	Balance of Opinion	Time	Source
Production	EVPRO	Change in output	m_t/m_{t-1}	BDF
Total orders	EVCOM	Change in overall level of new orders	m_t/m_{t-1}	BDF
Total foreign orders	EVCOME	Change in overall level of new orders from abroad	m_t/m_{t-1}	BDF
Past deliveries	EVLIV	Evolution of past deliveries of finished products	m_t/m_{t-1}	BDF
Order books	PREVPRO	Expected production	m_t/m_{t+1}	BDF
Stocks	EVSTPF	Evolution of stocks of finished products	m_t/m_{t-1}	BDF
Recent changes in output	PASTPROD	Change in production	m_t/m_{t-3}	INSEE
Demand and total order levels	CARNET	Level of the current total order intake	$m_t/E[m]$	INSEE
Finished-goods inventory level	STOCKS	Level of the current total finished products inventories	$m_t/E[m]$	INSEE
Personal production expectations	PERSEXP	Expectations about the evolution of production volume	m_t/m_{t+3}	INSEE
General production expectations	GENEXP	Expectations about the evolution of production volume for French manufacturing industries	m_t/m_{t+3}	INSEE
Output	OUTPUT	Volume of units produced	m_t/m_{t-1}	PMI
New Orders	NEWORDERS	Level of new orders received (units)	m_t/m_{t-1}	PMI
Stocks of Finished Goods	INVENT	Level of finished products come off the production line and awaiting shipment/sales (units)	m_t/m_{t-1}	PMI
Suppliers Delivery	SUPPLIERS	Average length of time agreed to deliver the goods	m_t/m_{t-1}	PMI

Notes: $m_t/m_{t\pm j}$ denotes a balance of opinions over the current month compared with the situation in the previous (next) j month(s). $m_t/E[m]$ denotes a balance of opinions over the current month compared with a norm.

Table 2: BDF models (1992Q1-2012Q4)

M1		M2		M3	
Var	Coeff	Var	Coeff	Var	Coeff
β_0	0.165*** (0.044)	β_0	0.082* (0.042)	β_0	0.091** (0.044)
y_{t-1}	-0.414*** (0.083)	y_{t-1}	-0.371*** (0.076)	y_{t-1}	-0.392*** (0.081)
$\text{evliv}_t^{(m1)}$	0.019*** (0.005)	$\text{evliv}_t^{(m2)}$	0.021*** (0.004)	$\text{evliv}_t^{(m3)}$	0.009** (0.004)
$\text{prevpro}_t^{(m1)}$	0.041*** (0.007)	$\text{evliv}_t^{(m1)}$	0.019*** (0.005)	$\text{evliv}_t^{(m2)}$	0.022*** (0.004)
		$\text{prevpro}_t^{(m2)}$	0.022*** (0.007)	$\text{evliv}_t^{(m1)}$	0.025*** (0.004)
d_{09Q1}	-1.032*** (0.284)	d_{09Q1}	-0.827*** (0.268)	d_{09Q1}	-0.964*** (0.274)
Adj-R^2	0.67	Adj-R^2	0.71	Adj-R^2	0.69
σ_ϵ	0.26	σ_ϵ	0.24	σ_ϵ	0.25
SIC	0.33	SIC	0.22	SIC	0.28
Normality	3.24 [0.19]	Normality	0.99 [0.61]	Normality	0.61 [0.73]
AR(4)	0.59 [0.67]	AR(4)	3.23 [0.02]	AR(4)	1.56 [0.19]
Hetero	0.42 [0.79]	Hetero	0.25 [0.94]	Hetero	0.29 [0.91]

Notes: Standard errors in parentheses. ***, ** and * denote statistical significance at the 1, 5 and 10% levels, respectively. Normality is the Bera-Jarque test for residual normal distribution. AR(p) is the Breusch-Godfrey test for residual serial correlation up to order $p = 4$. Hetero is the Breusch-Pagan-Godfrey test for heteroskedasticity. p-values in brackets.

Table 3: INSEE models (1992Q1-2012Q4)

M1		M2		M3	
Var	Coeff	Var	Coeff	Var	Coeff
β_0	0.505^{***} (0.053)	β_0	0.491^{***} (0.055)	β_0	0.488^{***} (0.049)
y_{t-1}	-0.369^{***} (0.084)	y_{t-1}	-0.345^{***} (0.084)	y_{t-1}	-0.362^{***} (0.077)
				$\Delta_{\frac{2}{3}}\text{carnet}_t^{(m3)}$	0.011^{***} (0.004)
$\Delta\text{prodpass}_t^{(m1)}$	0.011^{***} (0.003)	$\Delta\text{prodpass}_t^{(m2)}$	0.009^{***} (0.003)	$\Delta\text{prodpass}_t^{(m2)}$	0.009^{***} (0.003)
$\text{prodprev}_t^{(m1)}$	0.012^{***} (0.004)	$\text{prodprev}_t^{(m2)}$	0.015^{***} (0.004)	$\text{prodprev}_t^{(m2)}$	0.014^{***} (0.003)
$\text{persgen}_t^{(m1)}$	0.007^{***} (0.002)	$\text{persgen}_t^{(m2)}$	0.007^{***} (0.002)	$\text{persgen}_t^{(m1)}$	0.006^{***} (0.002)
d_{96Q1}	1.028^{***} (0.246)	d_{96Q1}	0.689^{***} (0.237)	d_{96Q1}	0.727^{***} (0.228)
d_{08Q2}	-0.721^{***} (0.247)	d_{08Q2}	-0.569^{**} (0.233)	d_{08Q2}	-0.517^{**} (0.224)
d_{08Q4}	-0.707^{***} (0.255)	d_{08Q4}	-0.511^{**} (0.253)		
d_{09Q1}	-0.873^{***} (0.277)	d_{09Q1}	-0.902^{***} (0.267)	d_{09Q1}	-0.797^{***} (0.253)
Adj-R^2	0.71	Adj-R^2	0.73	Adj-R^2	0.75
σ_ϵ	0.24	σ_ϵ	0.23	σ_ϵ	0.22
SIC	0.33	SIC	0.26	SIC	0.18
Normality	0.55 [0.76]	Normality	0.20 [0.90]	Normality	0.15 [0.93]
AR(4)	0.77 [0.55]	AR(4)	0.84 [0.50]	AR(4)	1.53 [0.20]
Hetero	1.21 [0.30]	Hetero	0.69 [0.69]	Hetero	0.72 [0.67]

Notes: see Table 2. Δ_z denotes a difference operator, where $z = 1$ means a quarter-on-quarter first-difference (same month, across quarters), and $z < 1$ is a month-on-month difference, with order given by the value of z ($z = \frac{1}{3}$ means one-month difference and $z = \frac{2}{3}$ means two-months difference).

Table 4: PMI models (1998Q2-2012Q4)

M1		M2		M3	
Var	Coeff	Var	Coeff	Var	Coeff
β_0	-1.617*** (0.365)	β_0	-2.774*** (0.303)	β_0	-2.774*** (0.303)
y_{t-1}	-0.270** (0.107)	y_{t-1}	-0.415*** (0.101)	y_{t-1}	-0.415*** (0.101)
$\Delta_{\frac{1}{3}}\text{suppliers}_t^{(m_1)}$	-0.038** (0.015)	$\Delta_{\frac{1}{3}}\text{output}_t^{(m_2)}$	0.075*** (0.012)	$\Delta_{\frac{1}{3}}\text{output}_t^{(m_2)}$	0.075*** (0.012)
$\text{neworders}_{t-1}^{(m_3)}$	0.058** (0.013)	$\text{neworders}_{t-1}^{(m_3)}$	0.063*** (0.006)	$\text{neworders}_{t-1}^{(m_3)}$	0.063*** (0.006)
$\text{output}_{t-1}^{(m_2)}$	-0.059*** (0.017)				
$\text{neworders}_{t-1}^{(m_1)}$	0.043*** (0.014)				
d_{08Q4}	-0.824*** (0.243)				
d_{09Q1}	-1.189*** (0.268)				
		d_{10Q4}	-0.786*** (0.243)	d_{10Q4}	-0.786*** (0.243)
Adj-R^2	0.74	Adj-R^2	0.73	Adj-R^2	0.73
σ_ϵ	0.22	σ_ϵ	0.23	σ_ϵ	0.23
SIC	0.24	SIC	0.15	SIC	0.15
Normality	1.31 [0.52]	Normality	1.16 [0.56]	Normality	1.16 [0.56]
AR(4)	1.02 [0.41]	AR(4)	0.67 [0.62]	AR(4)	0.67 [0.62]
Hetero	2.21 [0.05]	Hetero	1.51 [0.21]	Hetero	1.51 [0.21]

Notes: see Table 2 and 3.

Table 5: Full-information models (1998Q2-2012Q4)

M1		M2		M3	
Var	Coeff	Var	Coeff	Var	Coeff
β_0	-0.141 (0.250)	β_0	-0.396** (0.197)	β_0	-0.315 (0.201)
y_{t-1}	-0.521*** (0.080)	y_{t-1}	-0.391*** (0.072)	y_{t-1}	-0.459*** (0.074)
$\text{prodpass}_t^{(m1)}$	0.009** (0.004)	$\text{prodpass}_t^{(m2)}$	0.022*** (0.003)	$\text{evliv}_t^{(m3)}$	0.010*** (0.003)
$\text{prevpro}_t^{(m1)}$	0.029*** (0.006)	$\text{prevpro}_t^{(m2)}$	0.019*** (0.005)	$\text{evliv}_t^{(m2)}$	0.021*** (0.003)
$\text{persgen}_t^{(m1)}$	0.007** (0.002)	$\text{evliv}_t^{(m2)}$	0.022*** (0.003)	$\text{prodpass}_t^{(m2)}$	0.025*** (0.003)
$\text{evliv}_t^{(m1)}$	0.015*** (0.005)	$\text{evliv}_t^{(m1)}$	0.009** (0.004)	$\text{evliv}_t^{(m1)}$	0.014*** (0.004)
$\text{suppliers}_{t-1}^{(m3)}$	0.011* (0.006)	$\text{suppliers}_{t-1}^{(m3)}$	0.011** (0.004)	$\text{suppliers}_{t-1}^{(m3)}$	0.009** (0.004)
$\text{prodpass}_{t-1}^{(m1)}$	-0.007*** (0.003)	$\text{prodpass}_{t-1}^{(m2)}$	-0.010*** (0.002)	$\text{prodpass}_{t-1}^{(m2)}$	-0.010*** (0.002)
d_{06Q2}	0.532*** (0.175)	d_{06Q2}	0.455*** (0.151)	d_{06Q2}	0.481*** (0.156)
d_{07Q3}	0.475*** (0.178)				
d_{09Q1}	-0.866*** (0.209)				
Adj-R^2	0.86	Adj-R^2	0.89	Adj-R^2	0.88
σ_ϵ	0.16	σ_ϵ	0.15	σ_ϵ	0.15
SIC	-0.21	SIC	-0.55	SIC	-0.51
Normality	1.25 [0.53]	Normality	1.86 [0.39]	Normality	2.25 [0.32]
AR(4)	0.80 [0.53]	AR(4)	0.69 [0.60]	AR(4)	0.34 [0.85]
Hetero	0.51 [0.87]	Hetero	0.61 [0.76]	Hetero	1.26 [0.28]

Notes: see Table 2 and 3.

Table 6: Individual forecasts: one-step-ahead evaluation (2005Q1-2012Q4)

	RMSE				MAE			
	BDF	INSEE	PMI	Full	BDF	INSEE	PMI	Full
M1	0.302	0.285	0.325	**0.253**	0.211	0.216	0.226	**0.186**
	(0.036)	(0.037)	(0.040)	(0.029)	(0.027)	(0.026)	(0.022)	(0.019)
M2	0.242	0.255	0.261	**0.175**	0.170	0.185	0.212	**0.145**
	(0.032)	(0.035)	(0.029)	(0.019)	(0.025)	(0.025)	(0.022)	(0.015)
M3	0.247	0.242	0.261	**0.182**	0.156	0.173	0.212	**0.147**
	(0.034)	(0.030)	(0.029)	(0.020)	(0.026)	(0.023)	(0.022)	(0.016)

Notes: Criteria computed over mean-bias corrected forecast errors. Standard errors (in parentheses) are computed by non-parametric bootstrap with 10,000 draws. Bold entries denote the smallest RMSE and MAE, for each indicated monthly equation.

Table 7: Forecast encompassing tests

Model	M1		M2		M3	
$N = 3$, i = BDF, INSEE, PMI in Eq.(15)						
BDF	21.1	[0.00]	12.8	[0.00]	15.7	[0.00]
INSEE	12.9	[0.00]	14.2	[0.00]	14.8	[0.00]
PMI	29.4	[0.00]	23.4	[0.00]	12.1	[0.00]
$N = 4$, i = BDF, INSEE, PMI, Full in Eq.(15)						
Full	7.79	[0.05]	1.35	[0.72]	3.33	[0.34]

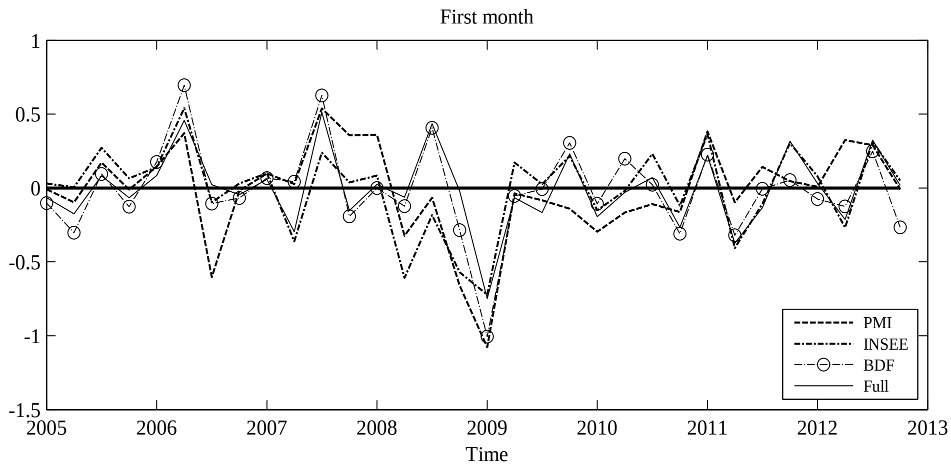
Notes: The table reports Wald test statistics and their associated p-values (in brackets).

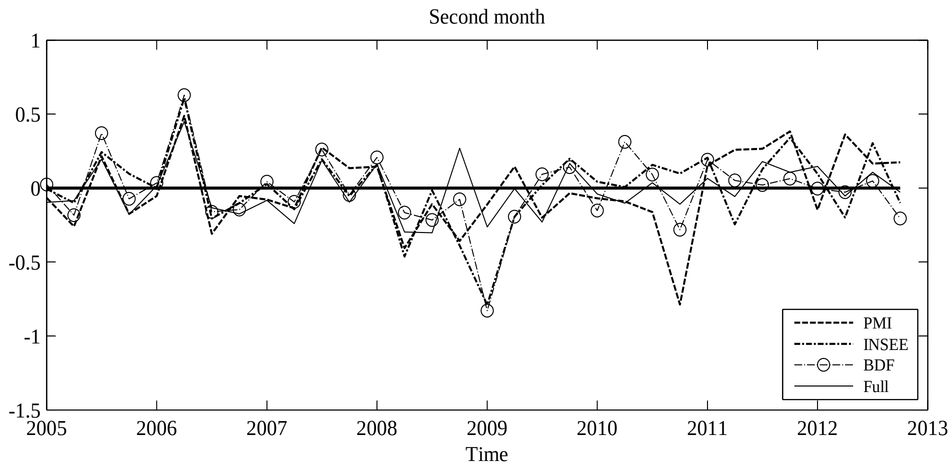
Table 8: Combined forecasts: one-step-ahead evaluation (2005Q1-2012Q4)

		RMSE			MAE		
		M1	M2	M3	M1	M2	M3
Naïve		0.270 (0.025)	0.212 (0.021)	0.209 (0.020)	0.181 (0.017)	0.152 (0.016)	0.150 (0.016)
IMSE	*FS*	0.269	0.212	0.209	0.182	**0.151**	0.149
	Rol	0.265	0.231	0.227	**0.178**	0.159	0.160
	Rec	0.269	0.216	0.213	0.182	0.152	0.148
IR	*FS*	0.269	0.217	0.212	0.187	0.152	0.151
	Rol	0.263	0.226	0.226	0.180	0.156	0.158
	Rec	0.273	0.226	0.219	0.189	0.151	**0.148**
OLS	*FS*	0.269	**0.211**	**0.209**	0.187	0.154	0.151
	Rol	**0.263**	0.261	0.255	0.181	0.183	0.176
	Rec	0.284	0.240	0.234	0.198	0.162	0.152
PLIK	*FS*	0.297	0.236	0.237	0.211	0.167	0.154
	Rol	0.318	0.247	0.251	0.231	0.186	0.170
	Rec	0.315	0.240	0.241	0.226	0.168	0.161

Notes: see Table 6.

Figure 1: Individual forecast errors (2005Q1-2012Q4)





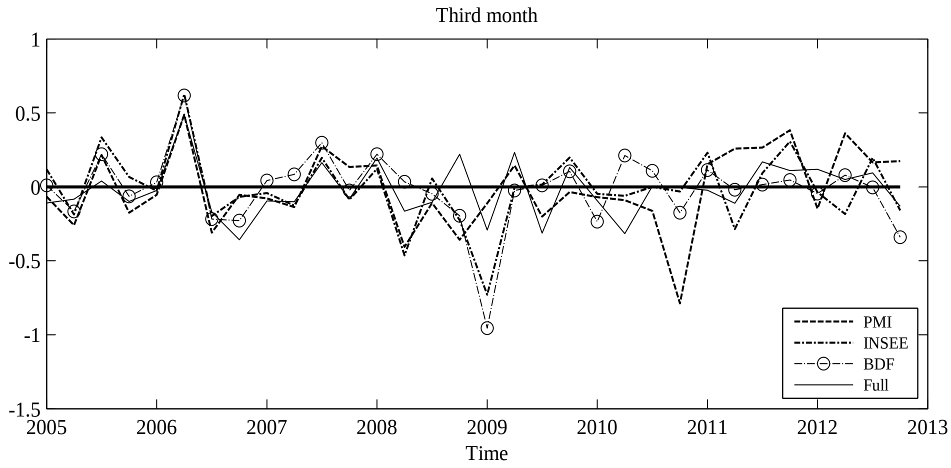
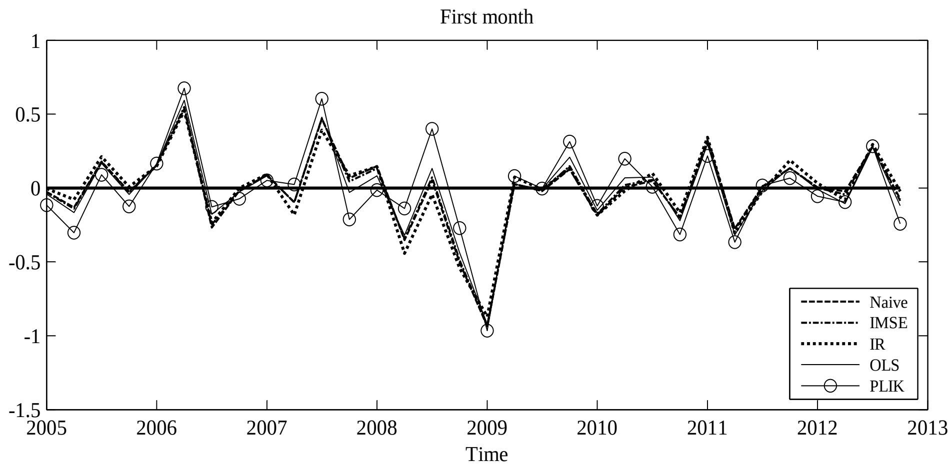
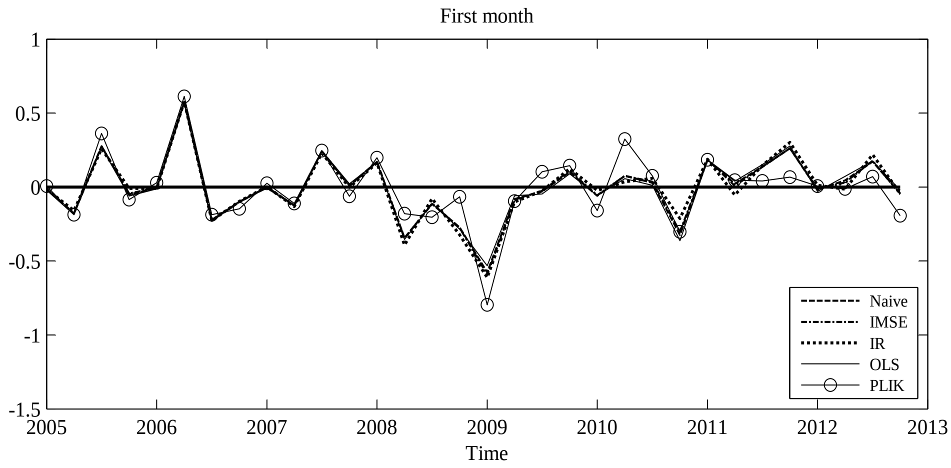


Figure 2: Combined forecast errors - full sample weights (2005Q1-2012Q4)





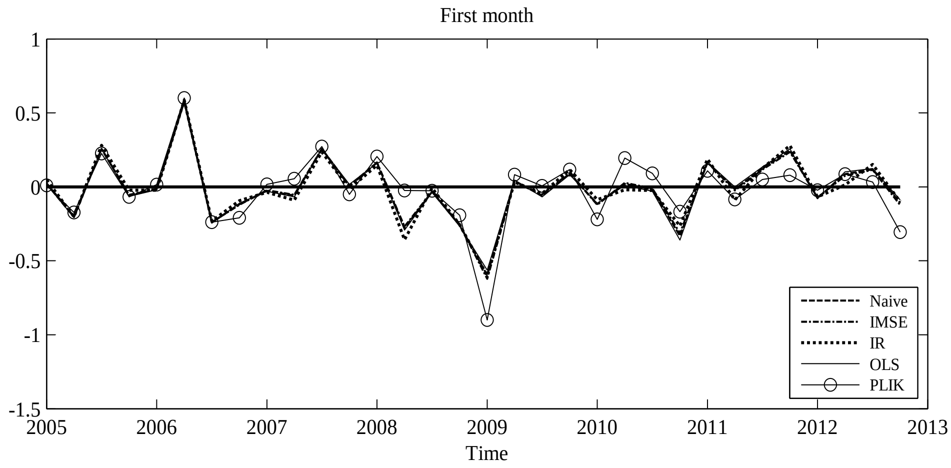


Figure 3: Combined forecast errors - recursive weights (2005Q1-2012Q4)

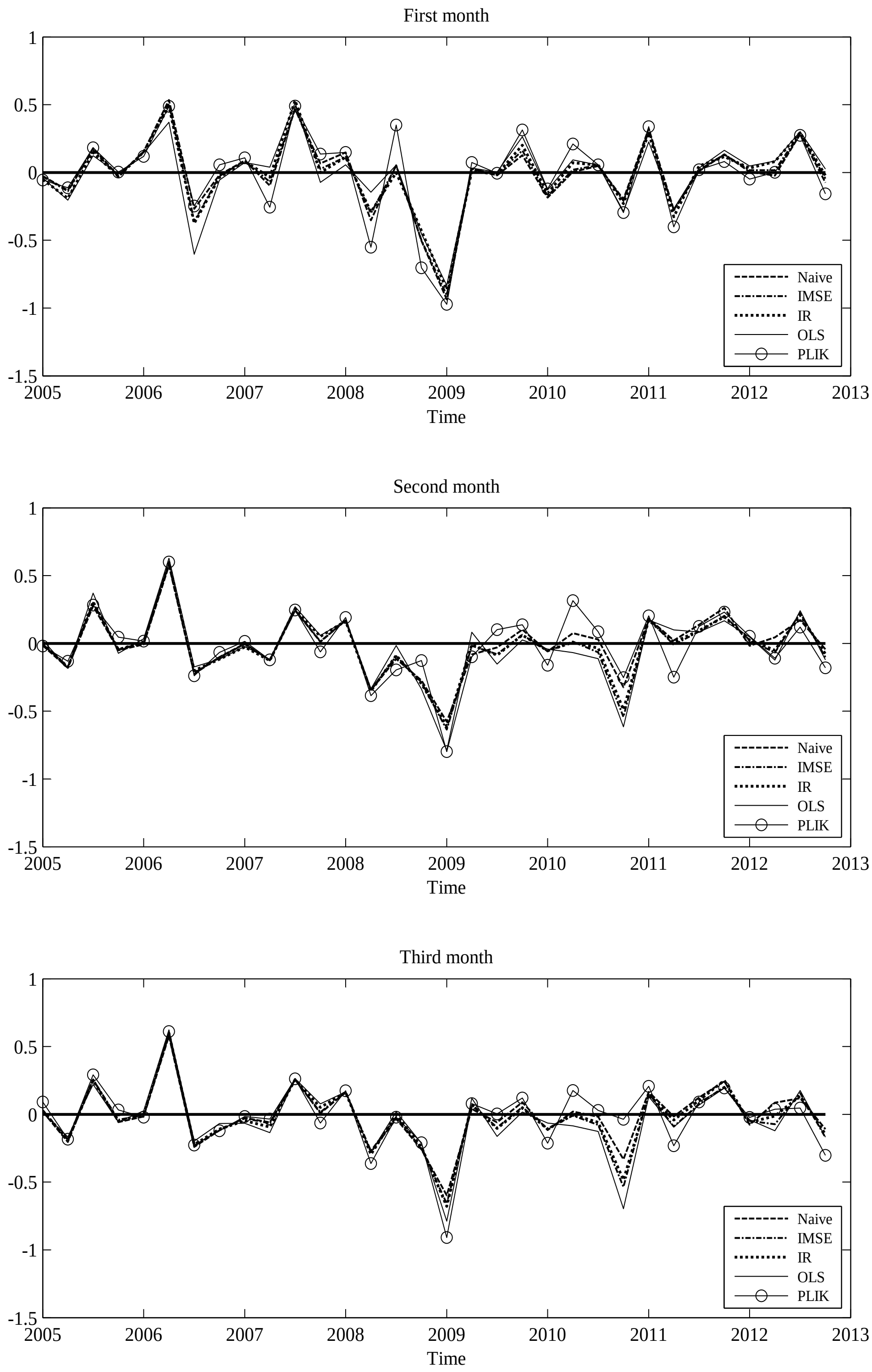
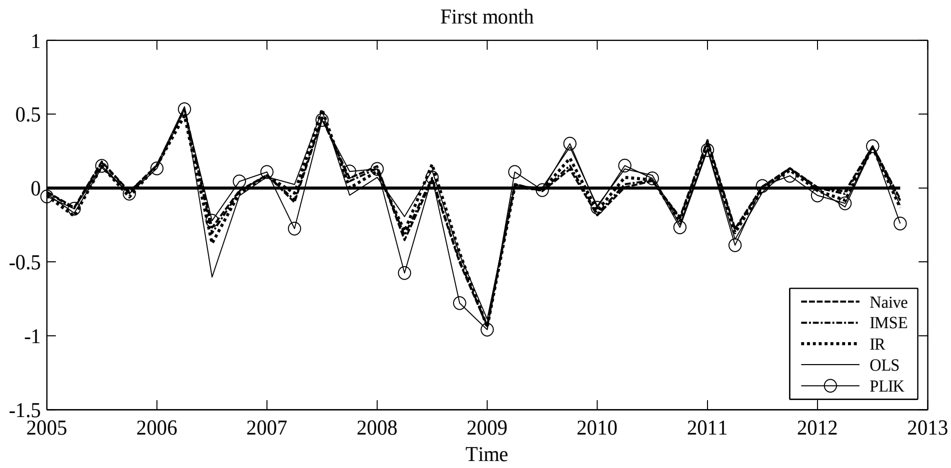
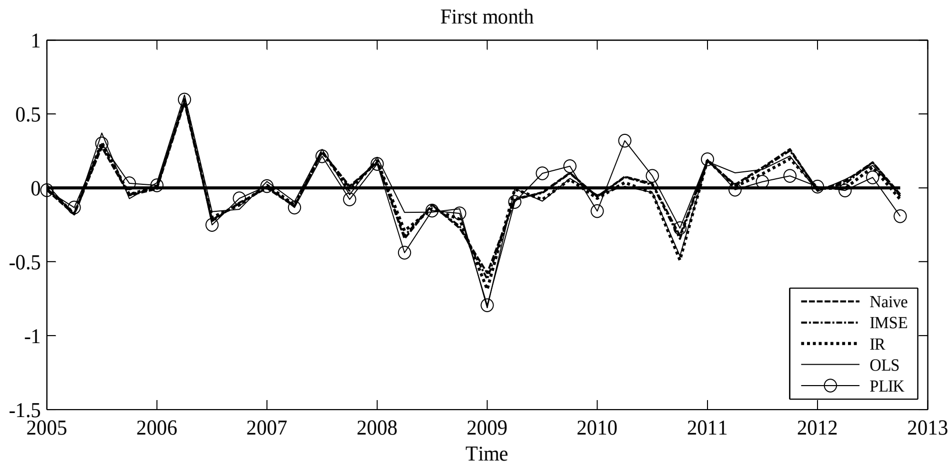


Figure 4: Combined forecast errors - rolling weights (2005Q1-2012Q4)





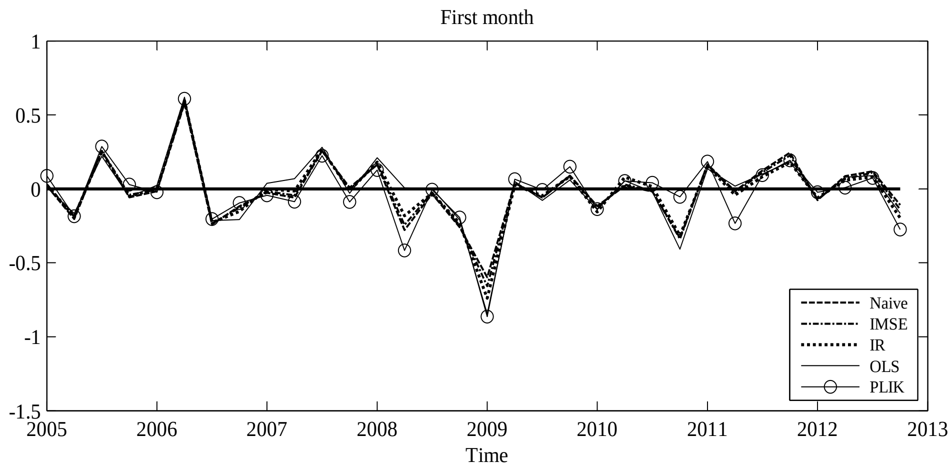


Figure 5: Full-information forecast errors vs Naïve combined forecast errors (2005Q1-2012Q4)