International Journal of Forecasting 31 (2015) 188–206

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting economic activity with targeted predictors

Guido Bulligan[a], Massimiliano Marcellino[b,c,*], Fabrizio Venditti[a]

[a] Banca d'Italia, Italy
[b] Bocconi University, IGIER, Italy
[c] CEPR, London, United Kingdom

ARTICLE INFO

Keywords:
Nowcasting
Forecasting
Mixed frequency data
Variable selection
Bridge models
Principal components
Factor models

ABSTRACT

In this paper we explore the forecasting performances of methods based on a pre-selection of monthly indicators from large panels of time series. After a preliminary data reduction step based on different shrinkage techniques, we compare the accuracy of principal components forecasts with that of parsimonious regressions in which further shrinkage is achieved using the General-To-Specific approach. In an empirical application, we show that the two competing models produce accurate current-quarter forecasts of Italian GDP and of its main demand components, outperforming naïve forecasts and comparing favorably with factor models based on all available information. A robustness check conducted on the GDP growth of the euro area and of its major members confirms these results.

1. Introduction

Although a large number of indicators covering all aspects of the economy are usually available at very high frequencies, quarterly national accounts still play a central role in guiding economic decisions and policy analysis. However, the delay with which they are released complicates decision making greatly. Over the past decade, a number of econometric tools have been developed to solve this problem.

One kind of older method, known as bridge models, are single equations in which lower frequency (typically quarterly) target variables are regressed against higher frequency (typically monthly) indicators preliminarily aggregated at the lower frequency, see Baffigi, Golinelli, and Parigi (2004), Barhoumi et al. (2008), Diron (2008) and Hahn and Skudelny (2008). Albeit very simple, these models are still used widely within policy institutions and by private forecasters, for a number of reasons. First, they strike a good compromise between simplicity and accuracy: a small set of indicators appropriately chosen usually guarantees a good forecasting performance. Second, forecasts based on single linear equations are very easy to explain and to communicate to decision makers. Third, dissecting forecast errors is also very easy in a linear context: discrepancies between actual and predicted values of the target variables can be related straightforwardly to those between actual and predicted values in the underlying indicators. However, bridge models present two important drawbacks: (i) they rely on a very parsimonious information set, potentially leaving out informative predictors, (ii) their specification often relies on the judgement and experience of the econometrician.

Their ability to address both of these issues is the reason for the attention that factor models, which have rapidly become the workhorse of short-term forecasting, have attracted in recent years. In these models, the information from a potentially very large dataset is summarized by a small number of linear combinations of the available time series, so that no valuable information is lost. Furthermore, the specification of a factor model requires little judgement: once the number of factors

* Correspondence to: Department of Economics, Bocconi University, Via Roentgen 1, 20136 Milan, Italy.
E-mail addresses: guido.bulligan@bancaditalia.it (G. Bulligan), massimiliano.marcellino@unibocconi.it (M. Marcellino), fabrizio.venditti@bancaditalia.it (F. Venditti).

http://dx.doi.org/10.1016/j.ijforecast.2014.03.004
0169-2070/

has been determined on the basis of some information criterion, the common factors can be estimated using various methods (see Forni, Hallin, Lippi, & Reichlin, 2005; Giannone, Reichlin, & Small, 2008; Kapetanios & Marcellino, 2009; Stock & Watson, 2002, for different estimation techniques), and a forecasting equation can be specified easily. In an empirical application, Angelini, Camba-Mendez, Giannone, Reichlin, and Rünstler (2011) find that, on average, factor models score better than bridge models in forecasting euro area GDP.

The issue of variable selection, which is crucial in the context of bridge models, is usually swept under the carpet in the factor model literature, where it seems that all that is needed is a large number of variables that can be used to average out the influence of idiosyncratic components and to estimate the common factors. A recent branch of the literature has questioned the usefulness of "too much information" for factor forecasts. Boivin and Ng (2006), for example, argue that increasing the N-dimension of large panels can be detrimental, especially if the errors are strongly cross-correlated and the forecasting power is provided by a factor that is dominant in a small panel but dominated in a larger panel. This problem arises because factors are extracted 'blindly', without taking into consideration the properties of the variable that the researcher is really interested in forecasting. To put it roughly, since principal components maximize the signal to noise ratio of the whole panel, they are well suited for forecasting the variables which load the common factors more strongly, but may perform poorly for other variables. Tailoring the predictors to a specific target variable can then provide substantial gains. Bai and Ng (2008) show that factor forecasts can be improved by identifying useful "targeted" predictors and computing principal components on this restricted dataset. In particular, they show that soft thresholding methods like the least absolute shrinkage selection operator (LASSO) can be used successfully to reduce the size of the information set.

An interesting connection between factor model forecasts and thresholding methods has recently been established by De Mol, Giannone, and Reichlin (2008). They find that, as the panel dimension increases, factor forecasts become more highly correlated with those obtained with LASSO, i.e., with a regression on a few selected predictors. They conclude that "… the result that few selected variables are able to capture the space spanned by the common factors, suggests that small models with accurately selected variables may do as well as methods that use information on large panels and are based on regressions on linear combinations of all variables. This point calls for further research…".

This open question constitutes the main motivation of our paper. However, we go beyond simple Lasso regressions, or more generally regressions of target variables on targeted predictors, by intersecting the targeted predictors argument with the General To Specific (GETS) modeling philosophy (see Hoover & Perez, 1999) that underlies the bridge approach (Krolzig & Hendry, 2001). Our analysis also proceeds in two steps. In the first step we follow Bai and Ng (2008) and use a range of hard- and soft-thresholding methods to reduce the dimension of a large dataset to a limited number of potential regressors. In the second step, information extraction is accomplished through an automatic selection algorithm which picks the most informative variables and specifies parsimonious bridge equations, in order to replicate the process usually followed by the econometricians, guided by their judgement and experience, when setting up bridge models.[1] Hence, our first methodological contribution relates to the specification of bridge equations in the presence of a large set of potentially useful indicators, based on sound statistical procedures rather than simply on either the experience and preferences of the bridge model developer or the use of information criteria and/or testing with only a small set of indicators.

Our second contribution is a comparison, in terms of forecasting performances, of this enhanced bridge approach with simple AR models that do not use any external information, with Diffusion Index models estimated on targeted predictors (as in Bai & Ng, 2008), and with general Diffusion Index models based on all of the information available. We can therefore assess: (i) the accuracy gain associated with monthly timely information, (ii) the "harmfulness" of "too much information", and (iii) the relative gains of two alternative ways of extracting information from targeted predictors (selection based on statistical criteria or information extraction by means of factor estimation).

Our empirical analysis focuses on Italian GDP and on the main demand components. The motivation for looking not only at GDP but also at the demand breakdown is twofold. First, factor models have been employed frequently for forecasting GDP, but seldom if ever for forecasting demand components. However, the business cycle behavior of aggregate GDP is very different from that of its components. For example, investment and trade variables are much more volatile than aggregate GDP, while Private Consumption is typically smoother than total activity, see Artis, Marcellino, and Proietti (2004). Checking how models compare for forecasting variables that behave so differently over the business cycle is an interesting exercise on its own. Second, forecasting demand aggregates is extremely important at the turn of the cycle and in turbulent phases. For example, investment tends to trough before GDP, while consumption only achieves momentum when an expansion is well under way, peaking after the cycle. Having models that complement GDP forecasts with a view on the main drivers of economic activity enables business cycle analysts to provide a much more accurate reading of the cyclical phase.

Our application is to one-step-ahead forecasts of Italian GDP and of the main demand breakdown, that is, Private Consumption, Investment in Construction, Other Investment, Exports and Imports. By *one step ahead*, we mean *the next quarterly release*. Given the delay with which quarterly series are published, this actually amounts to performing a nowcast/backcast exercise. We deliberately limit our forecast horizon to the next quarterly release because we are interested in gauging the relative merits

[1] The GETS methodology is implemented using the freeware software GROCER (see http://dubois.ensae.net/grocer.html).

of linear projections of the targets on *different spaces* (one spanned by the factors estimated on the targeted predictors, one by the few indicators included in bridge models, and one by the factors estimated on the whole information set) when some information on the quarter of interest is already available. Once the forecast horizon moves further ahead, the balance of merits in forecasting accuracy is bound to shift from the projection method to the way in which monthly indicators are forecast into the future. The forecasting of monthly business cycle indicators is a very interesting topic in itself, but one that goes beyond the scope of the present work.

Finally, the question of how general our results are, that is, how successful the same method will be in other contexts, arises. Some of the specification choices we make are in fact specific to the Italian dataset, and are dictated mainly by the sample size. Also, we use intervention dummies, which are also dataset-specific. However, we conduct a robustness check by extending the exercise to the euro area, France, Germany and Spain. In this sense, the results are quite encouraging, as the conclusions we draw from the main forecasting exercise carry over to other datasets.

The paper is structured as follows. In Section 2 we discuss some preliminary issues. In Section 3 we briefly present the main features of our dataset. In Section 4 we describe the strategies employed for selecting the targeted predictors and the way in which we implement the general to specific (GETS) procedure. In Section 5 we discuss the results of our empirical analysis. Section 6 concludes. Finally, the Appendix provides additional details on the selection algorithms.

2. Preliminaries: temporal aggregation and ragged-edged data

As we are interested in forecasting a quarterly target variable by means of monthly indicators, we need to clarify the ways in which we address two important issues: temporal aggregation and ragged-edge data.

In the factor model literature, various different approaches have been considered. In recent papers, quarterly and monthly variables have been cast in a factor model with latent variables, with the missing observations being filled using the Kalman filter, see Banbura and Modugno (2010). Other authors fit a factor model to the monthly indicators and then use an auxiliary equation to forecast GDP, either taking quarterly averages of the factors, like Angelini et al. (2011), or using a MIDAS regression, like Marcellino and Schumacher (2010).

In the context of bridge models, the problem of mixed frequencies is usually bypassed by taking quarterly averages of monthly indicators. For example, in a bridge model where quarterly GDP growth is forecast on the basis of industrial production, the predictor is the quarter-on-quarter percentage change in the industrial production index.

As we are not interested in the best way of forecasting monthly indicators into the future, and in order to enable a comparison with the earlier empirical literature on forecasting with bridge models, we work with the lowest frequency (quarterly) by taking a quarterly average of the monthly regressors. The alternative route (monthly interpolation of quarterly GDP) would have required the use of an arbitrary interpolation method, and was therefore discarded.

The second issue is how to deal with the asynchronous release of the indicators. In the absence of a real-time dataset for the large number of indicators that we consider in this study, we use a pseudo real-time exercise in which we replicate the monthly release pattern of the indicators. In particular, we assume that we are in the middle of the month, after industrial production data have been released by Istat but before surveys on the current month become available. Forecasts of the next GDP release are produced until a flash estimate becomes available, roughly 45 days after the end of the quarter.

In Fig. 1, we provide a stylized description of the timing of the information flow and the forecasting cycle relative to GDP in Q2. We start nowcasting Q2 in the middle of May, when a flash estimate for Q1 is made available. In Fig. 1, we report the publication lags of four representative groups of variables: industrial production, surveys, interest rates, stock market indices and trade variables. As can be seen, the industrial production index and trade data for March are released at the same time as the Q1 flash estimate is released. On the other hand, survey data are published at the end of the reference month, so that the April release of the surveys is known by the middle of May. Stock market indices and interest rate data are available at a daily frequency, so their April averages are also known. The next nowcast for Q2 is produced in the middle of June, when one month of hard data and two months of soft data for the quarter of interest are available. Our forecasting cycle terminates with a backcast in the middle of July, when the information on the monthly indicators is almost complete. In August, the flash estimate for Q2 is released and we start forecasting Q3.

Formally, our forecasting exercise works as follows. For each year y and for each quarter $q = 1, 2, 3, 4$, we produce three forecasts $i = 1, 2, 3$. We estimate the parameters of a linear equation (either a bridge or a factor model) at the quarterly frequency with available data up to the previous quarter, and then compute a forecast. For example, in the second quarter ($q = 2$) of 2008 ($y = 2008$), the first forecast ($i = 1$) is computed in mid-May by setting up the linear model:

$$\text{target}_{2008,1} = \alpha + \beta x_{2008,1} + \epsilon_{2008,1}, \tag{1}$$

where $\text{target}_{2008,1}$ is the last available observation for our quarterly target (either GDP or other quarterly aggregates), $x_{2008,1}$ is a vector of dimension k which collects the quarterly values of predictors for 2008Q1, as available in mid-May, and $\epsilon_{2008,1}$ is a forecast error. In the case of bridge equations, the vector x is a set of appropriately chosen indicators; in the case of factor models, the vector x collects an autoregressive term and quarterly averages of the monthly factors.

The forecast is then obtained as:

$$\text{target}_{2008,2,1} = \hat{\alpha}^{\text{OLS}} + \hat{\beta}^{\text{OLS}} x_{2008,2,1}, \tag{2}$$

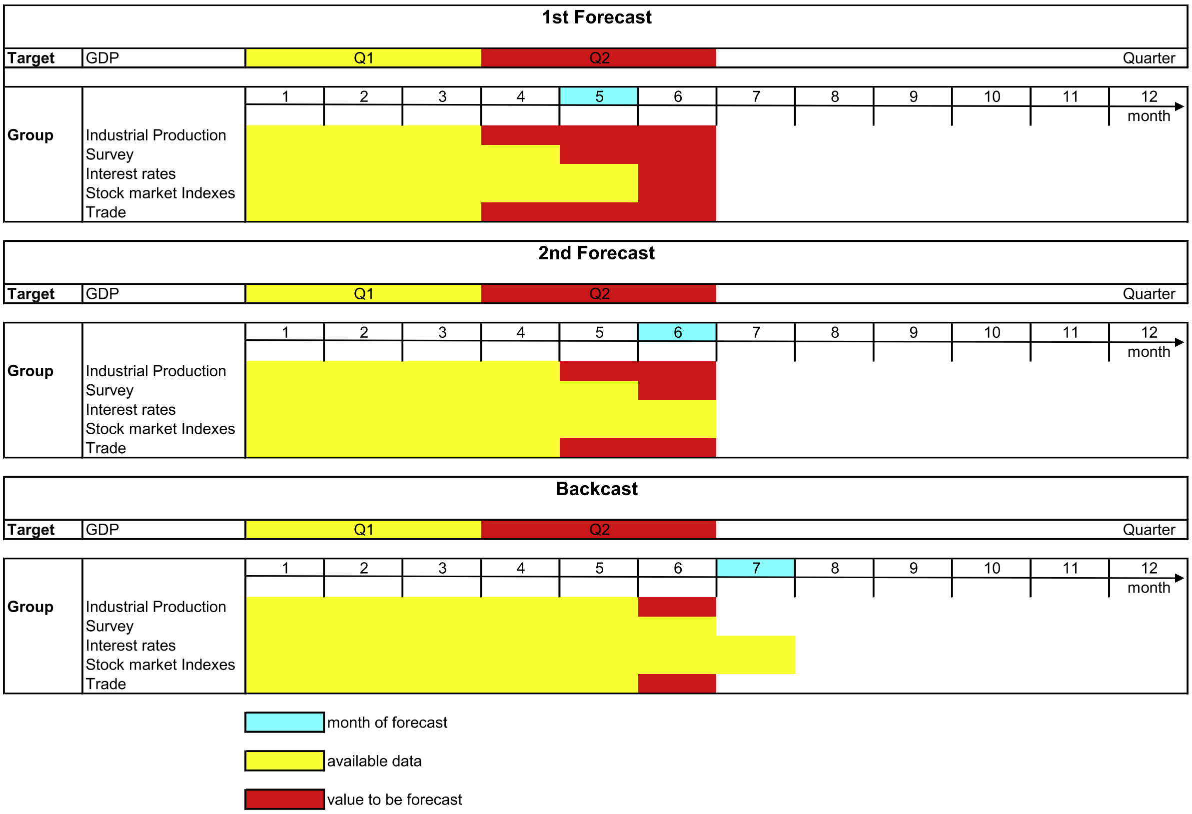

Fig. 1. Timing of the forecasting exercise and the availability of GDP data.

where $\hat{\alpha}^{OLS}$ and $\hat{\beta}^{OLS}$ are consistent OLS estimates of the parameters α and β in Eq. (1), and the subscript $i = 1$ indicates that the quarterly regressors x are averages of monthly indicators that are themselves partly forecast on the basis of the information set available in the first month of the forecast cycle (in this specific example, May).[2] Within the same forecasting cycle (for instance, that for GDP 2008 Q2), as we move from $i = 1$ to $i = 3$, the values of $\hat{\alpha}^{OLS}$ and $\hat{\beta}^{OLS}$ remain unchanged (at the values estimated using data up to 2008 Q1), but the value of the vector x changes as forecast values for missing monthly data are replaced by actual data.

3. Data description

As we want to compare the forecasting performances of different models in a realistic setting, we reconstructed a pseudo real-time dataset for the Italian economy. The dataset is organized into 138 monthly vintages, the first one covering the period 1982m1–1995m1 and the last one 1993m6–2006m6.[3]

Two characteristics of the dataset are worth commenting upon. Firstly, the panel dimension of the dataset increases over time, as it includes indicators that have become available progressively over the past 20 years.[4] Secondly, lacking a real-time database for the Italian economy, we proxied it by using the latest available data and reconstructing the same publication-lag structure for each monthly vintage.[5]

2 Indicators are forecast with autoregressive models as many steps ahead as is required in order to compute quarterly averages. The number of lags is chosen using the Akaike Information Criterion and updated at each forecasting round *i*.

3 The size of the window has been chosen with the aim of striking a balance between the number of observations required for computing correlations among variables and the likely occurrence of structural breaks. As the typical Italian business cycle lasts between 5 and 7 years, using 13 years of data allows us to cover two complete cycles, and therefore have reliable estimates of the true covariances among variables. We do not consider data after 2006m6, as we want to leave the most recent part of our sample for a comparison of the forecasting performances of the various models out-of-sample.

4 We approximate the overall availability of indicators that a forecaster would have faced in the last two decades. In other words, the panel dimension of the dataset N increases over time: for instance, until 1993, the dataset available to the researcher consisted of 68 variables. After 1993 it consisted of 225 variables. According to our reconstruction, the largest increases in the numbers of available variables occur among surveys and interest rates: the number of survey variables increases from a minimum of 8 in January 1995 to a maximum 69 in December 2006, while that of interest rate variables increases from 4 to 24.

5 In other words, we replicated the availability of data that a researcher would observe in the middle of the month, and reproduced it for each monthly vintage. Clearly, in so doing, we did not take into account the effects of statistical revisions on the forecasting performances of the models considered.

While the number and type of indicators can change over time, the information content of the dataset is rather stable, as it is always composed of:

- Hard indicators from the supply side.
- Hard indicators from the demand side.
- Manufacturing, construction, retail and consumer surveys.
- Trade variables.
- Interest rates and stock market indexes.
- Monetary and credit aggregates.
- Exchange rates (nominal, effective, real exchange rates).
- Labour market variables.
- Price variables.

Here, we comment briefly on the main characteristics of our dataset, referring to Table 9 for the full list of indicators and the meta data. Among supply-side hard indicators, the most important group of variables is that of industrial production indexes. Such variables are among the key indicators of the Italian business cycle, see for instance Altissimo, Marchetti, and Oneto (1997). The group includes both the general index and some of its sectoral components, as their performances can differ substantially from that of the aggregate in different phases of the business cycle.

Among demand indicators, the dataset includes retail sales, car registration, and electricity consumption in eight Italian districts, as well as in the railway grid. While such indexes reflect the energy demand for both production and domestic uses, their usefulness for nowcasting industrial production has been documented well by Bulligan, Golinelli, and Parigi (2010) and Marchetti and Parigi (1998).

Among survey variables, business surveys play a central role. In this respect, we include the balances between positive and negative answers for all questions in the Isae manufacturing survey at both the aggregate (total economy) and main industrial grouping levels. We also include in the dataset the aggregate information available in the Reuters-Markit PMI survey. The latter has become one of the most watched indicators of the business cycle in Europe and Italy, and we include here both the manufacturing index and the composite index (which includes information from the service sector). We also include the PMI manufacturing indexes for France, Germany, Spain and the Euro area, which proxy for the international real linkages between Italy and its main euro area trading partners.

Among interest rate variables, we include most reference rates in the term structure and spreads between different maturities, as well as banking rates charged to firms for short and long term borrowing and to households for house purchases. Considering the central role played by banks in the financial intermediation process, the latter, together with some credit aggregate related series (selected according to duration and recipients), might convey additional valuable information on the availability and cost of credit, and therefore on the interaction between monetary and real aggregates.

A further set of variables covers the international linkages of the Italian economy with the rest of the world. Considering the high sensitivity of the Italian economy to external conditions, we include value and volume indexes of exports and imports, as well as nominal and real effective exchange rates based on the consumer price index.

Among labour market indicators, we include the number of extra time hours, as well as the number of hours subsidized through the wage supplementation fund (CIG), a highly countercyclical indicator.

Finally, we include only few aggregate price indexes, while most of the disaggregate indexes are used indirectly to deflate nominal variables.

4. Identifying targeted predictors

Our dataset is characterized by a large degree of collinearity within blocks of variables and non-negligible idiosyncratic errors across variables belonging to different blocks. Picking the right regressors to form parsimonious linear bridge equations and/or selecting the appropriate amount of information for estimating factor models can therefore be quite challenging.

Before proceeding, it will be useful to fix at the outset some terminology that will be used in the rest of the paper. We use the term *target* to indicate a quarterly variable which we want to forecast on the basis of monthly information; the term *indicators* to refer to the monthly variables that are available in our dataset; and the term *targeted predictors* to refer to the *indicators* that have passed our selection tests.

As was mentioned in the introduction, we proceed by ordering and selecting indicators according to the rules suggested by Bai and Ng (2008), to end up with a dataset of a lower dimension. These data reduction methods can be classified as either HARD- or SOFT-thresholding rules. Under HARD-thresholding, an indicator is selected according to the significance of its correlation coefficient with the target. Typically, only indicators whose correlations with the target are above a given threshold are selected as targeted predictors. The obvious shortcoming of this selection criterion is that it only takes into account the bivariate relationship between the target and each indicator, without accounting for the information contained in other indicators. As a result, HARD-thresholding tends to select highly collinear targeted predictors.

On the other hand, SOFT-thresholding rules order and select indicators on the basis of a minimization problem of the following form:

$$\underbrace{\min}_{\beta} \mathrm{RSS} + \lambda\Psi\left(\beta_1, \ldots, \beta_j, \ldots, \beta_N\right), \tag{3}$$

where RSS is the Residual Sum of Squares of a regression of the target on the N indicators, and the Lagrange multiplier λ governs the shrinkage (the higher λ is, the higher the penalty for having extra regressors in the model), while Ψ is a function of RSS and of the N regression coefficients (β_j). Clearly, the cross-correlations among indicators are taken into consideration explicitly when minimizing this loss function.

Depending on the functional form of Ψ, various different SOFT-thresholding rules can be obtained. In our empirical application, we will focus on the following four:

- Least angle regressions (LARS)
- Least absolute shrinkage selection operator (LASSO)
- Elastic net estimator (NET)
- Forward selection regressions (FWD).

In the Appendix, we provide details on how they each arise from the general penalized regression in Eq. (3).

The next step consists of designing a selection algorithm that combines the five screening rules above so as to extract the targeted predictors from our large dataset. Within each estimation sample, we use the five methods above to rank the available indicators. We then associate with each indicator and selection method a binary variable which takes the value 1 if that indicator was ranked among the top 15 by that given selection method, and 0 otherwise. We run this exercise on all 138 of the estimation samples.

At the end of the exercise, we obtain, for each indicator, five binary variables (one for each selection method) with 138 observations. The sample mean of these binary variables, which is in the range from 0 (if the indicator was *never* selected by the specific selection method) to 1 (if the indicator was *always* selected by the specific selection method), can be interpreted as the probability of being selected, conditional on a given thresholding method. We include the indicator in the pool of targeted predictors if the probability of being included exceeds 0.6, conditional on at least one of the thresholding rules.

A specific example can make the algorithm clearer. Take for example industrial production (IP). For each rolling window, we check whether IP was ranked among the top 15 predictors by HARD, LARS, LASSO, NET and FWD. We then obtain five binary variables, which we will call IP^{HARD}, IP^{LARS}, IP^{LASSO}, IP^{NET} and IP^{FWD}, with 138 zero/one realizations. If the mean of at least one of these binary variables is above 0.6, IP is included in the set of targeted predictors. Since an indicator can be selected on the basis of more than one thresholding method, we take the *union* of the indicators selected by each thresholding method as our final set of targeted predictors.

Two issues need further clarification: (1) the choice to work with a rolling window instead of running the exercise once and for all on the full sample, and (2) the choice of the 0.6 probability threshold. We choose to work with a rolling window in order to have some control over possible structural breaks. Suppose that we run the selection exercise on the full sample, and an indicator which was very informative at the beginning of the sample is poorly correlated with the target in the last part of the sample. Under some conditions (for example, if the measurement error of the target variable has also varied over time), it can happen that the indicator is selected by some thresholding rule in spite of the fact that its predictive content at the end of the sample is low. The same would happen if the correlation between an indicator and the target depended on an outlier located somewhere in the sample. We therefore require that an indicator be picked as consistently as possible over the 138 rolling windows in order to be included in the pool of targeted predictors.

The 0.6 probability threshold is worked out backwards as the cut-off that ensures that we do not end up with more than 30 targeted predictors, which is roughly the maximum number of variables that the GETS routine can handle, given the number of available observations for the targets. In a way, this threshold can therefore be seen as an additional shrinkage parameter that we calibrate as part of making our selection algorithm operational.

4.1. Forecasting with targeted predictors

Having identified a subset of targeted predictors, two alternative strategies are available for extracting their predictive content for the target variable. The one analyzed by Bai and Ng involves setting up diffusion index models (usually abbreviated as "DI models"), where the estimated factors (principal components) condense the information dispersed in the dataset efficiently.

The alternative approach is to specify linear (bridge) models which relate the target variable to a few indicators and their lags. While the specification of bridge models is usually accomplished by a mixture of judgement, econometricians' experience and formal discipline, here we try to minimize the contributions of the first two by applying the so-called general-to-specific (GETS) model selection procedure, advocated by Hoover and Perez (1999) and Krolzig and Hendry (2001). This method starts from a general statistical model (here defined by the set of targeted predictors), then employs standard testing procedures to reduce its complexity by eliminating statistically insignificant variables and checking that the resulting model satisfies some predetermined criterion.

In our application, following the recommendations of Krolzig and Hendry (2001), we include the following tests in the battery of diagnostic tests:

- Chow predictive failure test with a break at 50% of the sample for parameter constancy;
- Chow predictive failure test with a break at 90% of the sample for parameter constancy;
- Doornik and Hansen's test for normality of the residuals;
- LM autocorrelation test up to fourth order autocorrelation in the residuals;
- Heterosckedasticity test for the residuals.

The significance levels for the selection t-tests are set to 0.05, while the significance levels for the five diagnostic tests are set to 0.01.

5. Results

In this section we present the results of the empirical analysis. In the first subsection we discuss the selected indicators. In the second subsection we present the main forecasting results. In the third subsection we focus on the great recession. In the final subsection we present the results for the euro area, France, Germany and Spain.

Table 1
Bridge models: indicators selected by GETS.

GDP	Exports
IP: Other manuf. repair (lag)	Electricity consumption—Milan (lag)
PMI manufacturing euro area	Export volume
Electricity consumption—Florence	Hourly wage rate (deflated)
IP: Intermediate goods	Real effective exchange rate—CPI-based
IP: Investment goods	
D9601	
Imports	**I-constr.**
IP: Means of transport (lag)	Stock price index—electricity
Import volume	IP: Construction
IP: Rubber and materials	IP: Investment goods—new orders (motor vehicles)
	D974
	D981
I-other	**Consumption**
Commercial vehicle registration	PMI manufacturing France
IP: Investment goods	IP: Metal and metal products
D031	New passenger car registration
	Electricity consumption—Palermo (lag)
	Retail survey: future business situation
	Retail sales volume index
	D9601

5.1. Selected indicators

Before turning to the forecast accuracy performances of the various models, we first show in Table 1 the indicators that enter the bridge models for our six quarterly targets: GDP, exports, imports, investment in construction, other investment and consumption. It is worth noting at the outset that the GETS procedure delivers very parsimonious specifications, as the number of regressors (excluding the constant) is four in most cases, and seven in only one case (consumption).[6] Also notice that the lag of the dependent variable is never selected. Since the lack of residual autocorrelation is one of the prerequisites for an acceptable specification of the GETS procedure, this means that the contemporaneous and lagged values of the indicators are generally sufficient to catch the dynamics of the dependent variables.

For each equation, it is possible to single out a few regressors that are usually considered as monthly proxies for the quarterly variable of interest, and that receive a specific scrutiny from most economic analysts and commentators. This is a comforting outcome, as it indicates that the procedure delivers interpretable results, and interpretability is one of the strengths of bridge models.[7]

Starting with GDP, the GETS procedure selects three industrial production indexes: the index for the intermediate goods sector, that for the investment goods sector, and the first lag of the index of the repair and installation of machinery. Although the total index, a standard regressor in most bridge equations, is not picked up by the algorithm, the industrial production subcomponents which are included are known to play a leading role in driving cyclical fluctuations in economic activity.

Among the regressors of the consumption equation, the number of car registrations, the index of retail sales volume, and the assessment of the future business situation from the Isae survey among retailers are selected.

The main driver in the equation for construction investment is the index of production in construction.[8] Turning to investment in machinery equipment, transport and patents (other investment), the regressors are the number of registered commercial vehicles and the IP index.

In the import and export equations, the drivers are the volumes of imported and exported goods, respectively.

5.2. Forecasting exercise

To gauge the forecasting abilities of our models, we run a pseudo real-time out-of-sample forecasting exercise over the period January 2007 to January 2010.[9] As was explained in Section 3, we produce two forecasts and one backcast for each quarter.

For each macro aggregate, the two main models to be compared are the bridge (BRIDGE) model and the diffusion index forecasting model, which we rename the Targeted Predictors factor model (TP-F), while the benchmark models are in turn a simple AR process (AR) and a diffusion index forecasting model, where the factors are extracted from the full dataset (ALL-F). While comparison with the AR model is a standard exercise in most forecasting applications, as it allows us to quantify the accuracy gains associated with models that incorporate additional exogenous information, comparison with the ALL-F model allows us to quantify the importance of pre-selecting indicators, and therefore to reduce the information set in a meaningful way.

For completeness, in the following out-of-sample forecasting exercise, we also report the results obtained by selecting a specific diffusion index forecasting model for each of the five selection methods (LARS-F, LASSO-F, NET-F, FWD-F and HARD-F). When interpreting the results obtained with these models, it must be kept in mind that a researcher could not exploit this information in a real

6 In some equations, intervention dummies capturing specific episodes were added to the specification after the GETS selection. In no case did adding the dummy variables lead to any loss of significance of the remaining coefficients. In order to preserve the pseudo-real time nature of the exercise, dummy variables were included based on available information up until December 2006.

7 Note that we use very intuitive transformations of the indicators. In some studies, goodness-of-fit is achieved through complicated nonlinear transformations of the indicators, at the cost of reducing the interpretability of the equation.

8 As Istat only releases a quarterly version of its underlying monthly index, our index is obtained by aggregating the real turnover of industries which provide inputs for the construction industry.

9 As was explained in Section 3, in the absence of a real-time database for the Italian economy, the exercise falls into the category of pseudo real-time analyses, as we use the latest vintage available for both the target variables and the monthly indicators, and therefore do not take into account the revisions that occur to data after their first release. The exercise is out of sample because data up to 2006 are used for the in-sample specification and diagnostics and data from 2007 to 2010 are used for the forecast comparison.

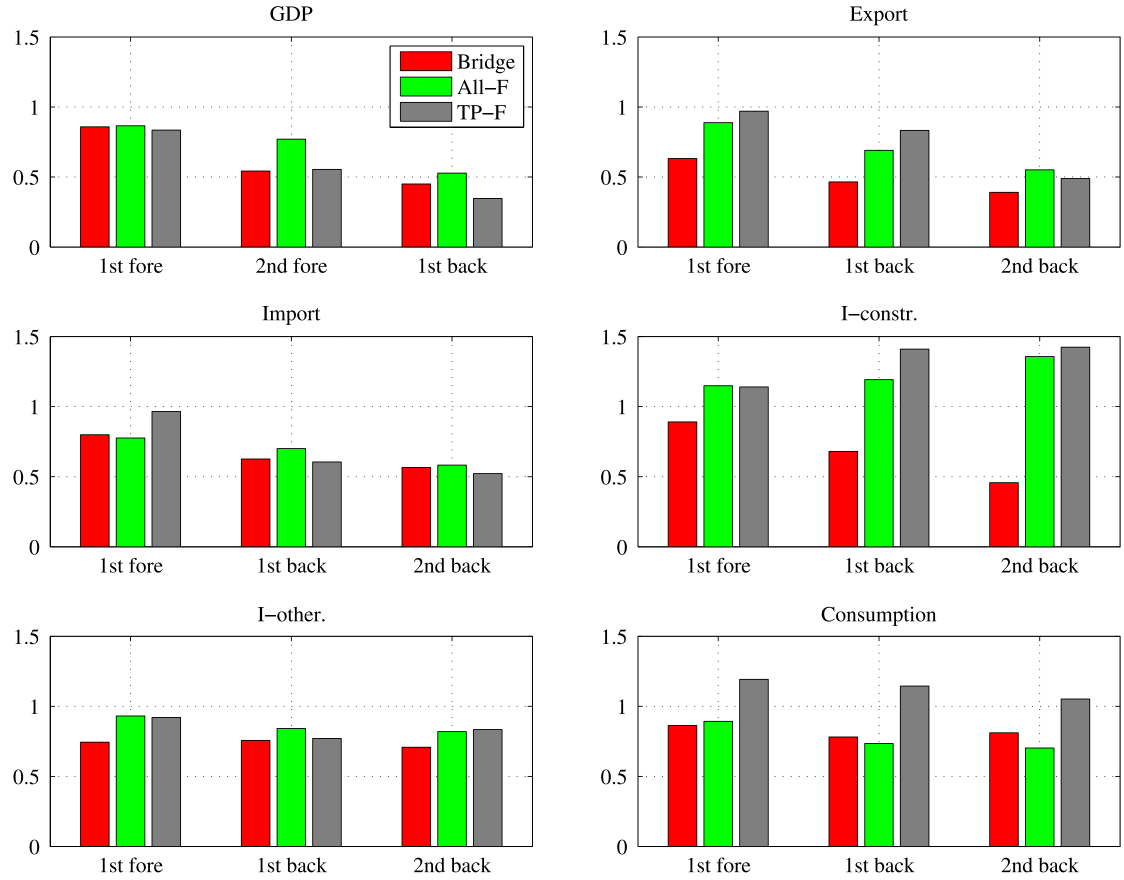

Fig. 2. RMSFEs: Bridge, TP-F and All-F models relative to AR. Notes: the figure reports the Root Mean Squared Forecast Errors (RMSFEs) of the Bridge, TP-f and ALL-F models relative to that of an AR model. The number of lags of the AR model is selected optimally based on the Akaike criterion at each step. The three bars refer to the three different forecast horizons, which are the second and third month of the current quarter (1st forecast and 2nd forecast) and the first month of the next quarter (backcast).

time context, as she would not know a priori which selection procedure would deliver the most accurate results. Nonetheless, it is important to check whether any one of these criteria emerges as uniformly superior in an ex post forecast evaluation.

Before turning to the results, we stress that, throughout the forecasting exercise, while the number and type of indicators entering each model are those which are selected on the basis of information up until December 2006, the parameters of the models (both the regression coefficients and the parameters that determine the number of static factors, as well as the numbers of their lags) are re-estimated at each monthly iteration.

To give an initial indication of how well our models forecast, we present in Fig. 2 the ratio of the Root Mean Square Forecast Errors (RMSFE) of the bridge model (BM), the factor model based on the whole information set (ALL-F) and that of the diffusion index models based on targeted predictors (TP-F) to that obtained with the benchmark autoregressive model (AR).[10] Furthermore, to evaluate the uncertainty regarding these results, we also run standard pairwise Diebold–Mariano tests on the forecast errors and report the results in Table 2.

Table 2
Equal forecast accuracy tests.

	1st fore	2nd fore	1st back	1st fore	2nd fore	1st back
	GDP			Exports		
TP-F	0.76	0.22	**0.02**	0.18	0.30	0.55
Bridge	0.90	0.39	0.76	0.41	**0.08**	**0.11**
	Imports			I-constr.		
TP-F	0.32	**0.12**	0.62	0.99	0.60	0.84
Bridge	0.45	0.76	0.94	0.68	0.14	**0.06**
	I-other			Consumption		
TP-F	0.93	0.46	0.94	0.41	**0.05**	0.28
Bridge	0.66	0.42	0.41	0.79	0.93	0.98

Notes: the table reports the *p*-values of pairwise equal forecast accuracy (Diebold–Mariano) tests. The tests compare the forecast accuracies of, respectively, Bridge and TP-F models with those of the All-F models. The null hypothesis is that of equal forecast accuracy. We highlight values lower than 0.15 in bold.

In the case of GDP, pre-selecting indicators (through either BM models or the TP-F model) delivers more accurate forecasts than either an autoregressive or a factor model

[10] The number of lags to be included in the autoregressive model is re-optimized at each forecasting step.

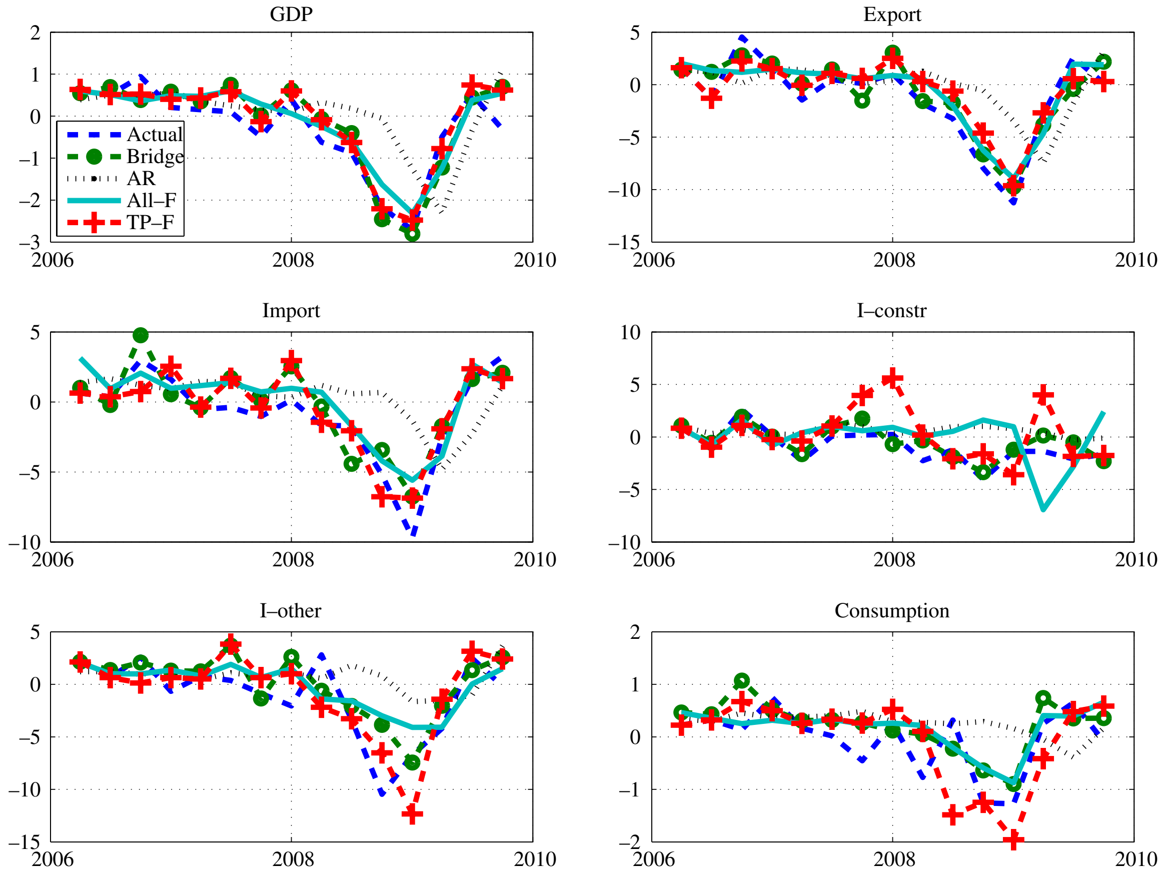

Fig. 3. Model forecasts. Notes: the figure shows the forecasts of the Bridge, TP-f, ALL-F, and AR models, together with the actual values. We only report results for the third month of the forecast cycle (backcast).

without pre-selection (ALL-F). In fact, the RMSFE ratios of the BM and TP-F models are always below one (meaning that these models outperform the AR benchmark), and never higher than those of the ALL-F model.

However, the Diebold–Mariano tests indicate that, while the AR model is outperformed by the remaining models in the first two forecasting steps, no significant differences in accuracy arise among the BM, TP-F and ALL-F models. This may be due to the fact that the gains from pre-selecting indicators are counteracted by the inaccuracy of projecting the monthly indicator to close the quarter. In the backcasting round (that is, when two months of hard data have already been released), the forecast accuracy of the TP-F model is significantly higher than those of the competing setups.

When turning to demand components, bridge models always outperform both the AR benchmark and the TP-F models, and often the ALL-F model as well. They also generally show a uniformly declining RMSFE over the forecasting exercise, indicating that they make good use of the new monthly releases. On the other hand, the TP-F models are outperformed by AR models in a couple of cases (consumption and investment in construction). Furthermore, they do not always make an efficient use of incoming information, as their RMSFEs increase rather than decaying over the forecast cycle in the cases of investment in construction and, to a lesser extent, other investment.

A comparison of the forecasts with the actual data, plotted in Fig. 3, gives further insights into the forecast accuracy results.[11] Firstly, in the case of GDP, the differences across models are negligible, with the exception of the AR model, which over-predicts GDP growth throughout the crisis. This result further emphasizes the value of monthly information for tracking the business cycle in real time.

Secondly, turning to the components, the forecasts produced by the TP-F model tend to be more volatile than those of the competing ones. This is especially true in the cases of investment in construction and consumption, which explains the poor performances highlighted above in terms of RMSEs. Bridge model forecasts, on the other hand, are generally smoother and follow the turning points quite closely.

The full set of results from the forecasting exercise are shown in Tables 3–8. Before commenting on the numbers, an explanation of the way in which these tables are organized is in order. Each table is divided into three panels. The upper panel shows the RMSFEs, while the remaining two report the relative RMSFEs with respect to the benchmarks (autoregressive model (AR) in the central panel, DI

[11] To avoid complicating the graphs too much, we only plot the backcasts.

Table 3
RMSFEs: GDP.

	AR	Bridge	ALL-F	TP-F	LARS-F	LASSO-F	NET-F	FORW-F	HARD-F	Best
Panel A: RMSFE										
1st fore	0.93	0.80	0.80	0.77	0.82	0.88	0.84	0.84	0.86	TP-F
2nd fore	0.93	0.50	0.71	0.51	0.58	0.57	0.53	0.77	0.59	Bridge
Backcast	0.90	0.41	0.48	0.31	0.31	0.40	0.41	0.75	0.36	Lars-F
Panel B: relative RMSFE (benchmark AR)										
1st fore	1.00	**0.86**	**0.87**	**0.83**	**0.88**	**0.94**	**0.91**	**0.91**	**0.93**	
2nd fore	1.00	**0.54**	**0.77**	**0.55**	**0.62**	**0.61**	**0.57**	**0.83**	**0.64**	
Backcast	1.00	**0.45**	**0.53**	**0.35**	**0.34**	**0.45**	**0.46**	**0.83**	**0.40**	
Panel C: relative RMSFE (benchmark ALL-F)										
1st fore	1.16	**0.99**	1.00	**0.96**	1.02	1.09	1.05	1.05	1.07	
2nd fore	1.30	**0.70**	1.00	**0.72**	**0.81**	**0.80**	**0.74**	1.07	**0.83**	
Backcast	1.89	**0.85**	1.00	**0.66**	**0.64**	**0.84**	**0.86**	1.58	**0.75**	

Notes: the table reports the Root Mean Squared Forecast Errors (Panel A), and the relative RMSFEs with respect to a AR model (Panel B) and the DI model with no pre-selection of indicators (ALL-F, Panel C). In panels B and C, a value below one indicates that the model in the column outperforms the benchmark. For ease of visualization, entries below one are highlighted in bold. The rows refer to the months of the forecasts (the 1st forecast is performed in the second month of the quarter, the 2nd forecast is performed in the third month of the quarter, and the backcast is performed in the first month of the following quarter). The last column (Best) reports the model with the smallest RMSFE for each forecast period. Each model is evaluated in a rolling (the estimation sample is 13 years) pseudo out-of-sample real time forecasting exercise over the period 2006Q4–2009Q4.

Table 4
RMSFEs: exports.

	AR	Bridge	ALL-F	TP-F	LARS-F	LASSO-F	NET-F	FORW-F	HARD-F	Best
Panel A: RMSFE										
1st fore	3.66	2.31	3.25	3.55	2.81	3.07	2.86	3.16	3.15	Bridge
2nd fore	3.50	1.63	2.41	2.91	2.71	2.53	2.60	3.16	1.90	Bridge
Backcast	3.50	1.37	1.93	1.71	1.48	1.07	1.38	2.52	1.18	Lasso-F
Panel B: relative RMSFE (benchmark AR)										
1st fore	1.00	**0.63**	**0.89**	**0.97**	**0.77**	**0.84**	**0.78**	**0.86**	**0.86**	
2nd fore	1.00	**0.47**	**0.69**	**0.83**	**0.77**	**0.72**	**0.74**	**0.90**	**0.54**	
Backcast	1.00	**0.39**	**0.55**	**0.49**	**0.42**	**0.30**	**0.39**	**0.72**	**0.34**	
Panel C: relative RMSFE (benchmark ALL-F)										
1st fore	1.13	**0.71**	1.00	1.09	**0.87**	**0.94**	**0.88**	**0.97**	**0.97**	
2nd fore	1.45	**0.67**	1.00	1.21	1.12	1.05	1.08	1.31	**0.79**	
Backcast	1.81	**0.71**	1.00	**0.89**	**0.76**	**0.55**	**0.71**	1.30	**0.61**	

Note: see notes to Table 3.

Table 5
RMSFEs: imports.

	AR	Bridge	ALL-F	TP-F	LARS-F	LASSO-F	NET-F	FORW-F	HARD-F	Best
Panel A: RMSFE										
1st fore	3.12	2.50	2.43	3.01	2.97	3.69	3.41	2.39	3.10	Forw-F
2nd fore	3.00	1.88	2.10	1.81	2.08	2.32	2.37	2.59	1.99	TP-F
Backcast	3.00	1.70	1.75	1.57	1.56	1.97	2.00	2.36	1.80	Lars-F
Panel B: relative RMSFE (benchmark AR)										
1st fore	1.00	**0.80**	**0.78**	**0.96**	**0.95**	1.18	1.09	**0.76**	**0.99**	
2nd fore	1.00	**0.63**	**0.70**	**0.61**	**0.69**	**0.77**	**0.79**	**0.86**	**0.66**	
Backcast	1.00	**0.57**	**0.58**	**0.52**	**0.52**	**0.66**	**0.67**	**0.79**	**0.60**	
Panel C: relative RMSFE (benchmark ALL-F)										
1st fore	1.29	1.03	1.00	1.24	1.23	1.52	1.41	**0.98**	1.28	
2nd fore	1.43	**0.89**	1.00	**0.86**	**0.99**	1.10	1.13	1.23	**0.95**	
Backcast	1.72	**0.97**	1.00	**0.90**	**0.89**	1.12	1.14	1.35	1.03	

Note: see notes to Table 3.

forecasting model without pre-selection (ALL-F) in the bottom panel). In the central and bottom panels, entries higher (lower) than one mean that the corresponding models perform worse (better) than the benchmark. In each panel, the three rows refer to the three monthly forecasts with increasing amounts of information. The first four columns re-

Table 6
RMSFEs: investment in construction.

	AR	Bridge	ALL-F	TP-F	LARS-F	LASSO-F	NET-F	FORW-F	HARD-F	Best
Panel A: RMSFE										
1st fore	2.03	1.81	2.33	2.31	2.42	1.98	2.17	1.74	2.47	Forw-F
2nd fore	1.99	1.35	2.37	2.81	2.31	2.18	2.18	2.08	2.69	Bridge
Backcast	1.99	0.91	2.70	2.83	3.00	2.12	2.71	1.97	2.46	Bridge
Panel B: relative RMSFE (benchmark AR)										
1st fore	1.00	**0.89**	1.15	1.14	1.19	**0.98**	1.07	**0.86**	1.22	
2nd fore	1.00	**0.68**	1.19	1.41	1.16	1.09	1.10	1.05	1.35	
Backcast	1.00	**0.46**	1.36	1.42	1.51	1.07	1.36	**0.99**	1.24	
Panel C: relative RMSFE (benchmark ALL-F)										
1st fore	**0.87**	**0.78**	1.00	**0.99**	1.04	**0.85**	**0.93**	**0.75**	1.06	
2nd fore	**0.84**	**0.57**	1.00	1.18	**0.97**	**0.92**	**0.92**	**0.88**	1.13	
Backcast	**0.74**	**0.34**	1.00	1.05	1.11	**0.79**	1.00	**0.73**	**0.91**	

Note: see notes to Table 3.

Table 7
RMSFEs: other investment.

	AR	Bridge	ALL-F	TP-F	LARS-F	LASSO-F	NET-F	FORW-F	HARD-F	Best
Panel A: RMSFE										
1st fore	3.76	2.80	3.50	3.46	3.63	3.28	3.53	3.22	3.65	Bridge
2nd fore	3.65	2.76	3.08	2.81	3.33	3.23	2.84	2.59	3.05	Forw-F
Backcast	3.65	2.59	2.99	3.05	3.39	2.70	3.06	3.06	2.94	Bridge
Panel B: relative RMSFE (benchmark AR)										
1st fore	1.00	**0.74**	**0.93**	**0.92**	**0.96**	**0.87**	**0.94**	**0.86**	**0.97**	
2nd fore	1.00	**0.76**	**0.84**	**0.77**	**0.91**	**0.89**	**0.78**	**0.71**	**0.83**	
Backcast	1.00	**0.71**	**0.82**	**0.83**	**0.93**	**0.74**	**0.84**	**0.84**	**0.80**	
Panel C: relative RMSFE (benchmark ALL-F)										
1st fore	1.07	**0.80**	1.00	**0.99**	1.04	**0.94**	1.01	**0.92**	1.04	
2nd fore	1.19	**0.90**	1.00	**0.91**	1.08	1.05	**0.92**	**0.84**	**0.99**	
Backcast	1.22	**0.86**	1.00	1.02	1.13	**0.90**	1.02	1.02	**0.98**	

Note: see notes to Table 3.

Table 8
RMSFEs: consumption.

	AR	Bridge	ALL-F	TP-F	LARS-F	LASSO-F	NET-F	FORW-F	HARD-F	Best
Panel A: RMSFE										
1st fore	0.72	0.62	0.64	0.86	0.91	0.82	0.74	0.80	0.44	Hard-F
2nd fore	0.67	0.52	0.49	0.77	0.61	0.62	0.72	0.61	0.55	All-F
Backcast	0.67	0.54	0.47	0.71	0.49	0.74	0.73	0.64	0.47	All-F
Panel B: relative RMSFE (benchmark AR)										
1st fore	1.00	**0.86**	**0.89**	1.19	1.27	1.14	1.04	1.12	**0.61**	
2nd fore	1.00	**0.78**	**0.73**	1.14	**0.91**	**0.92**	1.08	**0.91**	**0.81**	
Backcast	1.00	**0.81**	**0.70**	1.05	**0.73**	1.10	1.08	**0.96**	**0.71**	
Panel C: relative RMSFE (benchmark ALL-F)										
1st fore	1.12	**0.97**	1.00	1.34	1.42	1.28	1.16	1.25	**0.69**	
2nd fore	1.36	1.07	1.00	1.56	1.23	1.25	1.47	1.24	1.11	
Backcast	1.42	1.15	1.00	1.50	1.04	1.56	1.54	1.36	1.01	

Note: see notes to Table 3.

fer to the AR model, the Bridge model, the ALL-F model and the TP-F model, respectively. For completeness, the next five columns then report the DI forecasting models based on each of the five selection criteria. Finally, the last column reports the best performing model for each month.

Some interesting results emerge from an overview of the tables. Firstly, one can see that, for most quarterly targets, exploiting the additional information embedded in monthly indicators generally leads to more accurate forecasts than using simple AR models (entries lower than one are highlighted in bold in the central panels). The only exception is investment in construction (and to a lesser extent consumption).

Secondly, focusing on the relative benefits of pre-selecting information (benchmark ALL-F model in Panel C), the accuracy gains from forecasting using targeted predictors (highlighted in bold) are more scattered across targets, so that one cannot state that filtering information always

Table 9
Data description.

Block	Description	Start date	Lag	Treatment
DEM	New passenger car registrations	Jan-90	1	deltap
DEM	Car registrations—new light commercial vehicles up to 3.5T	Jan-91	2	deltap
DEM	Exports: motor vehicles—trailers/semi-trailers	Jan-91	3	deltap
DEM	New orders: manufacturing—motor vehicles	Jan-90	3	deltap
DEM	New orders: mfg.—motor vehicles/trailer bodies (coachwork)	Jan-90	3	deltap
DEM	New orders: mfg.— motor vehicles, trailers/semi-trailers	Jan-90	3	deltap
DEM	Retail sales volume—food	Jan-96	3	deltap
DEM	Retail sales volume—non food	Jan-96	3	deltap
DEM	Retail sales volume	Jan-96	3	deltap
DEM	Commercial vehicle registration	Jan-80	1	deltap
ENEL	Electricity consumption total	Jan-80	1	deltap
ENEL	Electricity consumption railway	Jan-80	1	deltap
ENEL	Electricity consumption Turin	Jan-80	1	deltap
ENEL	Electricity consumption Milan	Jan-80	1	deltap
ENEL	Electricity consumption Venice	Jan-80	1	deltap
ENEL	Electricity consumption Florence	Jan-80	1	deltap
ENEL	Electricity consumption Rome	Jan-80	1	deltap
ENEL	Electricity consumption Neaples	Jan-80	1	deltap
ENEL	Electricity consumption Palermo	Jan-80	1	deltap
ENEL	Electricity consumption Cagliari	Jan-80	1	deltap
EXCH	Real effective exchange rate—CPI based	Jan-80	1	deltap
EXCH	Italian lire to Euro (ECU)	Jan-80	1	deltap
EXCH	Italian lire to US $ (mth.avg.)	Jan-80	1	deltap
INT	Interest rate on 3-month Italian bond	Feb-88	0	delta
INT	Interest rate on 6-month Italian bond	Feb-88	0	delta
INT	Interest rate on 12-month Italian bond	Apr-87	0	delta
INT	Interest rate on 10-year Italian bond	Apr-91	0	delta
INT	3-month interbank rate on deposits	Jan-80	1	delta
INT	Interest rate on 3-month German bond	Jan-80	0	delta
INT	Interest rate 10-year German bond	Jan-80	0	delta
INT	3 years gov bond yield	Oct-92	0	delta
INT	5 years gov bond yield	Nov-88	0	delta
INT	10 years gov bond yield	Mar-91	0	delta
INT	30 years gov bond yield	Nov-93	0	delta
INT	Lending rate to firms–short term (less than 1y)	Apr-99	2	delta
INT	Lending rate to firms–long term	Jan-95	2	delta
INT	Mortgage rate	Jan-95	2	delta
INT	Spread 12-month 3-month	Feb-88	0	none
INT	Spread 3-year 3-month	Oct-92	0	none
INT	Spread 10-year 3-month	Mar-91	0	none
INT	Spread 30-year 3-month	Nov-93	0	none
INT	Spread 10-year 3-year	Oct-92	0	none
INT	Spread 30-year 3-year	Nov-93	0	none
INT	Spread 10-year 5-year	Mar-91	0	none
INT	Spread 30-year 5-year	Nov-93	0	none
INT	Spread it-3-month de-3-month	Jan-80	1	none
INT	Spread it-10-year de-10-year	Mar-91	0	none
INT	Spread on lending rate short	Apr-89	2	none
INT	Spread on lending rate long	Jan-95	2	none
INT	Spread on mortgage rate	Jan-95	2	none
LAV	Cig in manufacturing	Jan-80	1	deltap
LAV	Cig ordinary	Jan-80	1	deltap
LAV	Cig construction	Jan-84	1	deltap
MON	Money supply: m1—it contribution to the Euro area	Jan-80	2	deltap
MON	Money supply: m2—it contribution to the Euro area	Jan-80	2	deltap
MON	Money supply: m3—it contribution to the Euro area	Jan-80	2	deltap
MON	Itm1-defl	Jan-80	2	deltap
MON	Itm2-defl	Jan-80	2	deltap
MON	Itm3-defl	Jan-80	2	deltap
MON	Loans to household—consumer credit	Mar-80	2	deltap
MON	Loans to household—for house purchase	Mar-80	2	deltap
MON	Loans to household—other credit	Mar-80	2	deltap
MON	Loans to non financial corporation—total	Mar-80	2	deltap
MON	Loans to non financial corporation—above 1 year	Mar-80	2	deltap
MON	Loans to non financial corporation—below 1 year	Mar-80	2	deltap
MON	Credit to private sector—total	Jan-83	2	deltap
MON	Non performing loans (percentage)	Jan-90	2	delta
MON	Consumer loans–defl	Mar-80	2	deltap

(*continued on next page*)

Table 9 (*continued*)

Block	Description	Start date	Lag	Treatment
MON	Mortgage loans–defl	Mar-80	2	deltap
MON	Loans to non-fin corp.: less 1y–defl	Mar-80	2	deltap
MON	Loans to non-fin corp.: over 1y–defl	Mar-80	2	deltap
MON	Other loans to househld–defl	Mar-80	2	deltap
MON	Credit to private sector–defl	Jan-83	2	deltap
MON	Total loans to non-fin. corp–defl	Mar-80	2	deltap
ORD-TURN	New orders	Jan-90	3	deltap
ORD-TURN	New orders: domestic	Jan-90	3	deltap
ORD-TURN	New orders: foreign	Jan-90	3	deltap
ORD-TURN	Industrial turnover	Jan-90	3	deltap
ORD-TURN	Sales: domestic	Jan-90	3	deltap
ORD-TURN	Sales: foreign	Jan-90	3	deltap
ORD-TURN	New orders–defl	Jan-90	3	deltap
ORD-TURN	Domestic orders–defl	Jan-90	3	deltap
ORD-TURN	Foreign orders–defl	Jan-90	3	deltap
ORD-TURN	Turnover–defl	Jan-90	3	deltap
PREZZI	PPI—linked/rebased	Jan-81	2	deltap
PREZZI	CPI including tobacco (NIC)	Jan-80	1	deltap
PREZZI	Baltic dry index	May-85	0	deltap
PREZZI	Composite price index—food commodities	Jan-80	2	deltap
PREZZI	Market price index—primary commodities	Jan-83	2	deltap
PREZZI	Export price—non fuel primary commodities index	Feb-80	2	deltap
PREZZI	Price of oil brent	Feb-82	0	deltap
PROD-IND	IP	Jan-80	2	deltap
PROD-IND	IP: manufacturing	Jan-80	2	deltap
PROD-IND	IP: consumer goods	Jan-90	2	deltap
PROD-IND	IP: consumer goods—durable	Jan-90	2	deltap
PROD-IND	IP: consumer goods—non-durable	Jan-90	2	deltap
PROD-IND	IP: investment goods	Jan-90	2	deltap
PROD-IND	IP: intermediate goods	Jan-90	2	deltap
PROD-IND	IP: energy	Jan-90	2	deltap
PROD-IND	IP: chemical products/synthetic fibres	Jan-90	2	deltap
PROD-IND	IP: coke manufacture/petroleum refining	Jan-90	2	deltap
PROD-IND	IP: extraction of minerals	Jan-90	2	deltap
PROD-IND	IP: food, drink/tobacco	Jan-90	2	deltap
PROD-IND	IP: machines/mechanical apparatus	Jan-90	2	deltap
PROD-IND	IP: means of transport	Jan-90	2	deltap
PROD-IND	IP: metal/metal products	Jan-90	2	deltap
PROD-IND	IP: rubber items/plastic materials	Jan-90	2	deltap
PROD-IND	IP: textile/clothing	Jan-90	2	deltap
PROD-IND	IP: wood/wood products	Jan-90	2	deltap
PROD-IND	IP: manufacture of computer, electronic and optical products	Jan-90	2	deltap
PROD-IND	IP: manufacture of electrical equipment	Jan-90	2	deltap
PROD-IND	IP: manufacture of basic pharmaceutical products	Jan-90	2	deltap
PROD-IND	IP: electricity, gas, steam and air conditioning	Jan-90	2	deltap
PROD-IND	IP: other manufacturing, and repair and installation of machinery an	Jan-90	2	deltap
PROD-IND	IP in construction	Jan-91	3	deltap
SP	MIB price index	Jan-80	0	deltap
SP	MIB price index—banking sector	Jan-80	0	deltap
SP	MIB price index—pharm sector	Feb-86	0	deltap
SP	MIB price index—telecom sector	Jan-80	0	deltap
SP	MIB price index—industry sector	Jan-80	0	deltap
SP	MIB price index—insurance sector	Jan-80	0	deltap
SP	MIB price index—info sector	Feb-86	0	deltap
SP	MIB price index—oil/gas sector	Feb-86	0	deltap
SP	MIB price index—electronic sector	Feb-86	0	deltap
SP	MIB price index—media sector	Feb-86	0	deltap
SP	MIB price index—banking sector	Dec-88	0	deltap
SP	MIB price index—auto sector	Jan-80	0	deltap
SP	ITALY-DS market–price earning ratio	Feb-86	0	none
SP	ITALY-DS market—dividend yield	Jan-80	0	none
SUR-BUS	Isae business confidence indicator	Jan-91	1	none
SUR-BUS	Isae business svy.: order books—domestic, net	Jan-91	1	none
SUR-BUS	Isae business svy.: order books—export, net	Jan-91	1	none
SUR-BUS	Isae business svy.: order books, net	Jan-91	1	none
SUR-BUS	Isae business svy.: stocks of finished goods, net	Jan-91	1	none
SUR-BUS	Isae business svy.: production level, net	Jan-91	1	none
SUR-BUS	Isae business svy.: order books in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae business svy.: production in next 3mos., net	Jan-91	1	none

Table 9 (*continued*)

Block	Description	Start date	Lag	Treatment
SUR-BUS	Isae business svy.: selling price in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae business svy.: economy in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed. gds.—order books domestic, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed. gds.—order books export, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed. gds.—order books, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed. gds.—stocks of fin.gds., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed. gds.—production level, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: investment goods—order books domestic, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: investment goods—order books export, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: investment goods—order books, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: investment goods—stocks of fin.gds., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: investment goods—production level, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: consumer goods—order books domestic, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: consumer goods—order books export, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: consumer goods—order books, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: consumer goods—stocks of fin.gds., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: consumer goods—production level, net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed.gds.—order books in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed.gds.—production in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed.gds.—sell.price in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: intermed.gds.—economy in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: inv.gds.—order books in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: inv.gds.—production in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: inv.gds.—selling price in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: inv.gds.—economy in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: cons.gds.—order books in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: cons.gds.—production in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: cons.gds.—selling price in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae bus.svy.: cons.gds.—economy in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae business svy.: order books, net	Jan-91	1	none
SUR-BUS	Isae business svy.: order books—domestic, net	Jan-91	1	none
SUR-BUS	Isae business svy.: order books—export, net	Jan-91	1	none
SUR-BUS	Isae business svy.: order books in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae business svy.: stocks of finished goods, net	Jan-91	1	none
SUR-BUS	Isae business svy.: production level, net	Jan-91	1	none
SUR-BUS	Isae business svy.: production in next 3mos., net	Jan-91	1	none
SUR-BUS	Isae business svy.: selling price in next 3mos., net	Jan-91	1	none
PMI	PMI manufacturing reconstr Euro area	Apr-87	1	none
PMI	PMI manufacturing reconstr Germany	Apr-87	1	none
PMI	PMI manufacturing reconstr France	Apr-87	1	none
PMI	PMI manufacturing reconstr Italy	Apr-87	1	none
PMI	PMI manufacturing reconstr Spain	Apr-87	1	none
PMI	PMI composite reconstructed Italy	Jan-87	1	none
SUR-CONS	Isae household confidence index	Jan-80	1	none
SUR-CONS	Isae consumer survey: general economic situation (balance)	Jan-82	1	none
SUR-CONS	Isae consumer survey: general economic expectations (balance)	Jan-82	1	none
SUR-CONS	Isae consumer survey: unemployment expectations (balance)	Jan-82	1	none
SUR-CONS	Isae consumer survey: economic situation of households (balance)	Jan-82	1	none
SUR-CONS	Isae consumer survey: economic expectations of households (balance)	Jan-82	1	none
SUR-CONS	Isae consumer survey: households budget (balance)	Jan-82	1	none
SUR-CONS	Isae cons.svy.: present intentions to purchase durables (balance)	Jan-82	1	none
SUR-CONS	Consumer survey: prices next 12 months—Italy	Jan-85	1	none
SUR-CONS	Consumer survey: savings at present—Italy	Jan-85	1	none
SUR-CONS	Consumer survey: savings over next 12 months—Italy	Jan-85	1	none
SUR-CONS	Consumer survey: prices last 12 months—Italy	Jan-85	1	none
SUR-OTH	Retail confidence indicator—Italy	Oct-85	1	none
SUR-OTH	Retail survey: current business situation—Italy	Oct-85	1	none
SUR-OTH	Retail survey: stocks—Italy	Oct-85	1	none
SUR-OTH	Retail survey: future business situation—Italy	Oct-85	1	none
SUR-OTH	Retail survey: orders placed with suppliers—Italy	Jan-86	1	none
SUR-OTH	Retail survey: employment—Italy	Nov-03	1	none
SUR-OTH	Construction confidence indicator—Italy	Jan-85	1	none
SUR-OTH	Construction survey: order book position—Italy	Jan-85	1	none
SUR-OTH	Construction survey: employment expectations—Italy	Jan-85	1	none
SUR-OTH	Construction survey: act. compared to last month—Italy	Jan-85	1	none
SUR-OTH	Construction survey: price expectations—Italy	Jan-85	1	none
SUR-OTH	Construction survey: limits to activity—demand, Italy	Jan-85	1	none
SUR-OTH	Services survey: evolution of demand in recent months—Italy	Apr-96	1	none
SUR-OTH	Services survey: evolution of demand expected in mth.ahead—Italy	Apr-96	1	none

(*continued on next page*)

Table 9 (*continued*)

Block	Description	Start date	Lag	Treatment
SUR-OTH	Services survey: evolution of emp expected in mth. ahead—Italy	Apr-96	1	none
SUR-OTH	Services survey: price expectation in months ahead—Italy	Jan-97	1	none
TRADE	Exports of goods FOB	Jan-91	2	deltap
TRADE	Imports of goods CIF	Jan-91	2	deltap
TRADE	Export unit value index	Jan-80	2	deltap
TRADE	Import unit value index	Jan-88	2	deltap
TRADE	Export volume index	Jan-80	2	deltap
TRADE	Import volume index	Jan-80	2	deltap
TRADE	Exports: consumer goods	Jan-91	2	deltap
TRADE	Exports: consumer goods—non-durable	Jan-91	2	deltap
TRADE	Exports: consumer goods—durable	Jan-91	2	deltap
TRADE	Exports: intermediate goods	Jan-91	2	deltap
TRADE	Exports: investment goods	Jan-91	2	deltap
TRADE	Imports from eu27: energy	Jan-93	2	deltap
TRADE	Imports from eu27: consumer goods	Jan-93	2	deltap
TRADE	Imports from eu27: intermediate goods	Jan-93	2	deltap
TRADE	Imports from eu27: investment goods	Jan-93	2	deltap
TRADE	Imports from non-eu countries (eu27): intermediate goods	Jan-93	2	deltap
TRADE	Imports from non-eu countries (eu27): investment goods	Jan-93	2	deltap
TRADE	Imports from non-eu countries (eu27): consumer goods—durable	Jan-93	2	deltap
TRADE	Imports from non-eu countries (eu27): consumer goods—non-durable	Jan-93	2	deltap
TRADE	Export value index	Jan-80	2	deltap
TRADE	Import value index	Jan-80	2	deltap
TRADE	Export unit value index	Jan-80	2	deltap
TRADE	Import unit value index	Jan-80	2	deltap
TRADE	Export volume index	Jan-80	2	deltap
TRADE	Import volume index	Jan-80	2	deltap
WAGE	Hourly wage rate index: all industry—manual/clerical workers	Nov-84	2	deltap
WAGE	Wage per employee	Jan-96	2	deltap
WAGE	It wage per employee–defl	Jan-96	2	deltap
WAGE	It hourly wage rate index: all ind-manual/clerical workers–defl	Nov-84	2	deltap

leads to better forecasts. Indeed, the differences in accuracy between models based on pre-selecting indicators (either BM models or TP-F models) and models that do not pre-select information (ALL-F) are seldom significant (Table 2). In this respect, however, using targeted predictors in combination with the GETS procedure generally gives better results than using the TP-F model. In fact, bridge models are outperformed by the ALL-F model in only three cases out of eighteen (1st forecast of imports, 2nd forecast and backcast of consumption).

Thirdly, the selection method that delivers the best results is not consistent either across tables or within a single table across monthly forecasts, suggesting that it would be difficult to choose a specific selection method a priori, and therefore "averaging across methods" (TP-F or Bridge) is a good idea. Focusing on GDP (Table 3), the accuracy gain with respect to the AR model ranges from 6% (LASSO-F) to 17% (TP-F) in the first forecast, from 17% (FWD-F) to 46% (Bridge) in the second forecast, and from 17% (FWD-F) to 66% (LARS-F) in the third forecast (backcast).

In this rich landscape of forecasting models, picking bridge models seems to be a safe choice, as they are selected as the best performing model in seven cases out of 18, and track the performance of the best models closely in the other cases.

5.3. Forecasting during the Great Recession

As the forecasting period is not particularly long and includes the so-called Great Recession, the question arises as to whether the previous results are driven by a few observations. In particular, the accuracy gains from models based on targeted predictors with respect to the benchmark AR model could be a reflection of the poor performance of the latter during a unique event with no historical precedent. Indeed, several authors, for instance D'Agostino and Giannone (2012), have highlighted the fact that even sophisticated models failed to outperform simple AR models during the late 1990s and early 2000s, a period often referred to as the "Great Moderation".

Dissecting the performances of the ALL-F, TP-F and Bridge models over appropriate subsamples is also an interesting exercise for evaluating the reliability of the models which are regularly used by policy makers during crisis periods.

In what follows, we therefore distinguish the performances of models of GDP during the period 2008Q2–2009Q2 (crisis) from those during the rest of sample (no crisis).[12] Considering the rather small numbers of observations available, formal statistical inference cannot be performed, so that the exercise should be taken as an event study providing qualitative results.

Despite the increase in the size of the forecast errors for all models during the crisis subsample, the results that

12 This period was classified by ISTAT as a recession for the Italian economy. Relative performances change slightly with the subsample definition, although the superior performances of models based on targeted predictors are confirmed.

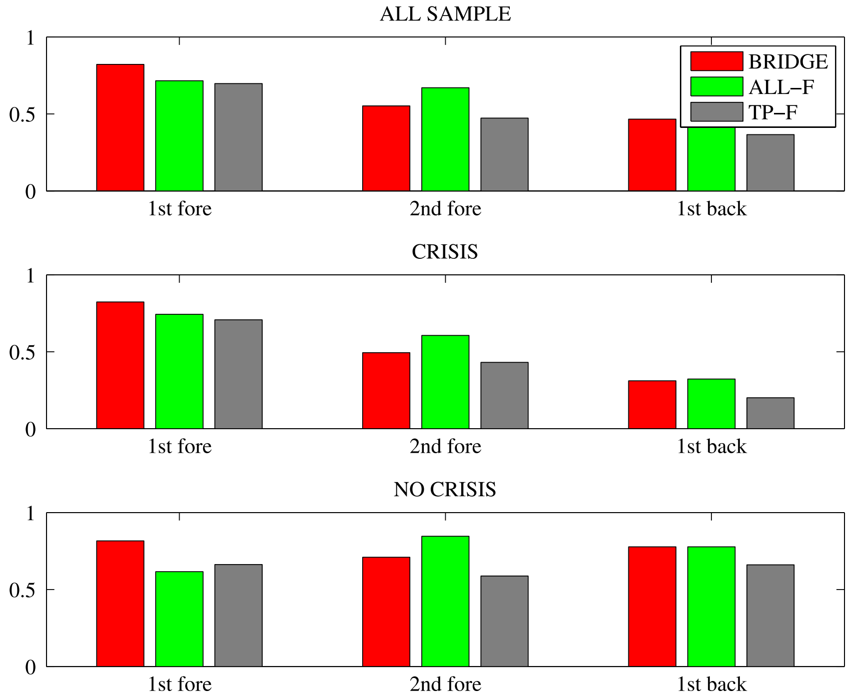

Fig. 4. RMSFEs of the bridge, TP-F and ALL-F models relative to that of AR—GDP subsample decomposition. Notes: the figure shows the ratios of the Root Mean Squared Forecast Errors (RMSFEs) of the bridge model, the targeted predictor factor model (TP-F) and the factor model with no pre-selection (ALL-F) to the those of the benchmark AR model over the whole sample and two subsamples (crisis, 2008Q2–2009Q2, and no crisis, 2006Q4–2008Q1 and 2009Q3–2009Q4). The three bars refer to the three different forecast horizons, which are the second and third month of the current quarter (first forecast and second forecast) and the first month of the next quarter (backcast).

emerge on the whole sample hold in terms of relative performances even when the Great Recession is separated from less turbulent quarters. Fig. 4 reports the RMSFEs of the ALL-F, TP-F and Bridge models relative to those of the benchmark AR computed over the whole sample, the crisis sample and the no-crisis period. A comparison of the accuracy gains between the two subsamples suggests that either the bridge model or the TP-F model outperforms the alternative models in almost all forecasting rounds and in both subsamples.[13] This finding confirms that selecting indicators is generally a useful step towards reducing the forecasting errors associated not only with naïve AR models but also with factor models that do not pre-select indicators on the basis of their information content.

This exercise also sheds light on the difficulties of forecasting during "the Great Recession". Focusing on the crisis subsample, one can see how the accuracy gains relative to the benchmark AR model increase as more information enters the forecaster information set. While this result had already been noted in the full-sample analysis, the limits of the AR models are clearly amplified during the crisis period. Furthermore, even within the class of models that exploit conjunctural indicators, it is interesting to note the sizable reduction in RMSFEs across forecasting rounds, which suggests the inadequacy of replacing missing data within the quarter using simple AR models during the crisis period. Overall, the sub-sample analysis, although of a qualitative nature, suggests that the gains associated with models based on targeted predictors are not driven by few exceptional observations.

5.4. *Results for the euro area, Germany, France and Spain*

Our main empirical exercise focuses on Italian GDP growth and its main demand components. However, the methodology that we use to first shrink the information set and then specify the bridge equations is general and can be applied to any large dataset. In order to check the robustness of our results and validate our methodology, we repeat the evaluation of GDP growth forecasts for the euro area, France, Germany and Spain.

The analysis follows the exact same steps as those described above. Firstly, for each country we gather a monthly dataset comprising the most commonly watched indicators of economic activity. Despite being unable to replicate the wide coverage of the Italian dataset, the dimensions of these information sets are quite large, varying from 60 indicators for Spain to 140 for Germany. More importantly, they all include indicators that are considered to be reliable indicators of the business cycle, according to both the available empirical country-specific literature and indications from the results for the Italian GDP. After removing seasonal factors and outliers, all data are transformed in order to guarantee stationarity.[14]

13 The only exception is the first forecast in the no-crisis period, for which the ALL-F model shows the best performance.

14 The large datasets used for this robustness check are drawn from Cristadoro, Saporito, and Venditti (2013). Details on the variables included for each country can be found therein.

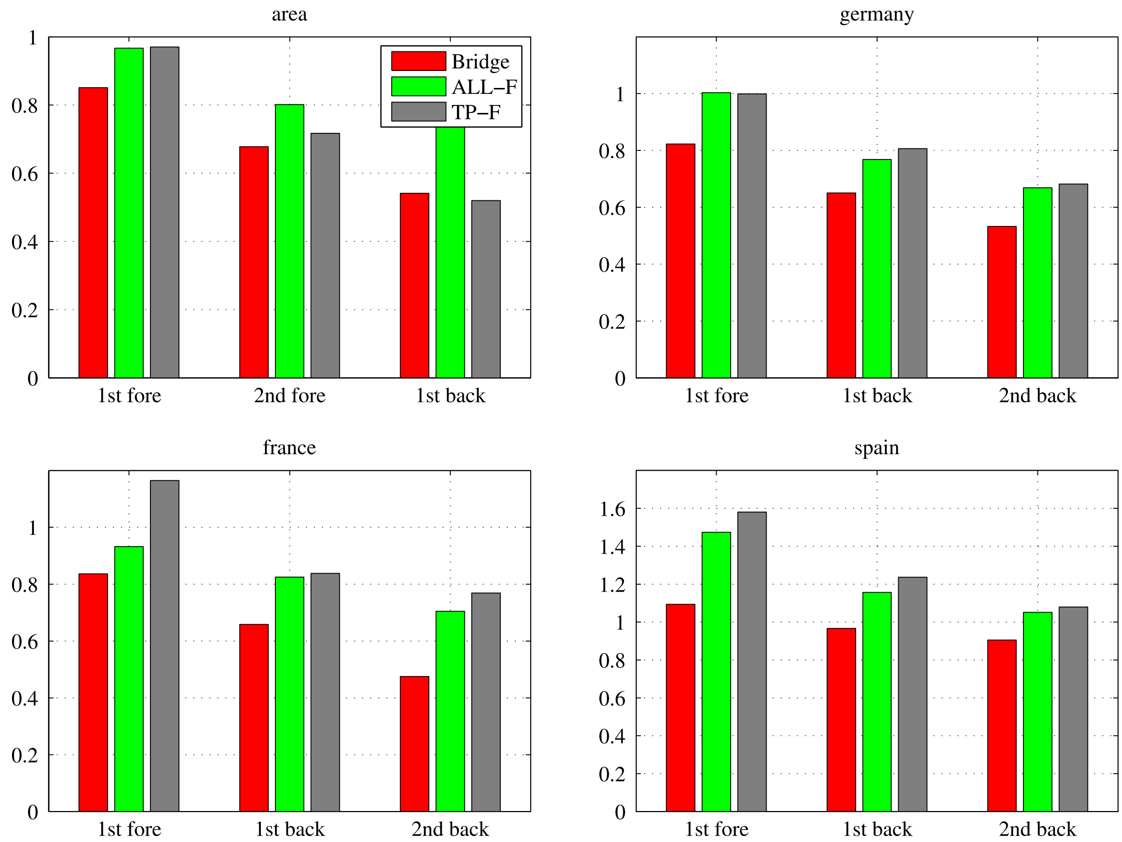

Fig. 5. RMSFEs for other euro area countries: Bridge, TP-F and ALL-F models relative to AR—GDP.

Secondly, we apply hard and soft thresholding techniques to reduce the country specific datasets to around 20 indicators.[15]

Thirdly, we turn to the GETS procedure, and, starting from the reduced information set, we specify four parsimonious country-specific bridge equations for forecasting GDP. Finally, we set up four TP-F models for GDP by regressing the quarterly GDP growth on quarterly averages of the first f principal components extracted from each national dataset and its p lags.

We run a forecast exercise akin to that performed for Italy and compare the forecast accuracies of the bridge, TP-F and ALL-F models. The main findings from this exercise are reported in Fig. 5, where the RMSEs of the different models are compared to that of a benchmark AR model.

By looking at the histograms, one can easily see the superior performance of bridge models compared to both the benchmark AR model and the two factor models (TP-F and All-F). Relative to the AR, the RMSEs from bridge models are always below one (with the exception of the first forecast for Spain), and decrease monotonically, as conjunctural information accumulates. Turning to factor models, selecting indicators (TP-F models) leads to an improvement over the ALL-F model for the euro area, but to a slight deterioration of accuracy in the other cases (more markedly for France). Also, notice that, in the case of Spain, factor forecasts (both ALL-F and TP-F models) perform worse than the AR model.

Taken together, these results confirm those obtained in the case of Italy: bridge models specified by combining targeted predictors with the GETS specification procedure perform at least as well as factor models either based on all the available information or subject to a pre-screening of the variables.

6. Conclusions

Forecasting quarterly variables relies on the availability of timely monthly indicators. In this paper we have compared two approaches to information extraction from large panels. The two approaches both rely on a pre-selection of the available indicators, but differ in the way in which the information is extracted.

While factor models based on targeted predictors exploit the covariance structure of the data in order to estimate the driving common factors to be used in the forecasting equation, bridge models rely on a general-to-specific approach for finding the most accurate approximation of the true unknown data generating process within a set of admissible models.

15 We kept the value of the probability threshold at 0.6, as for the Italian dataset.

The resulting models generally show good forecasting performances, clearly outperforming a benchmark AR model and comparing well with factor models that use all of the available information.

The lesson that we learn from the exercise is that the forecasting gains obtained by exploiting the timely information provided by monthly indicators can be increased further by carefully screening the available information.

Acknowledgments

The views expressed in this paper are those of the authors and do not necessarily reflect those of the Banca d'Italia. We would like to thank an anonymous referee for helpful comments on a previous draft.

Appendix. Selection algorithms in detail

In this appendix we summarize the algorithms used to pre-select the monthly indicators used in the analysis. The description of the algorithms broadly follows Bai and Ng (2008) and Hastie, Tibshirani, and Friedman (2009). In what follows, we neglect time subscripts without loss of generality and consider the problem of selecting a subset of the M covariates X for predicting the output variable Y, where X is a $T \times M$ matrix and Y is a T-dimensional vector.

A.1. Hard thresholding

The hard thresholding algorithm consists of running a regression of the target variable Y on each indicator X_i at the time. Under this rule, only the variables that show regression coefficients which are significant at the 5% level are kept as targeted predictors. When deciding whether or not to include each variable in the set of predictors, this method ignores the information contained in all of the other variables. It can therefore end up selecting variables that are too strongly correlated with each other.

A.2. Soft thresholding

Soft thresholding methods are based on special penalized regressions. To understand how these methods work, let us start from the basic penalized regression, namely the ridge regression. The ridge regression performs a form of *shrinkage*; that is, it penalizes coefficients that are too large by adding a constraint on the size of the coefficients to the usual sum of squares minimization problem, on which ordinary least squares is based. Specifically, the ridge coefficients are obtained by solving the following problem:

$$\underbrace{\min}_{\beta} \|Y - X\beta\| + \lambda \sum_{i=1}^{M} \beta_i^2, \tag{4}$$

where β is a M dimensional vector and $\|Y - X\beta\|$ indicates the L^2 norm, that is $\|Y - X\beta\| = (Y - X\beta)'(Y - X\beta)$. By tilting the estimated parameters towards zero, the ridge regression induces a bias in the coefficients, but reduces the estimation variance. This could result in more accurate forecasts when the covariates in X are strongly correlated with each other. The parameter λ governs the degree of shrinkage; that is, the higher λ the closer to zero are the β_i.

A.2.1. LASSO

The problem with the ridge regression is that the estimated parameters are smaller than under OLS, but they are never exactly zero. There is therefore scope for reducing the estimation variance further by setting some of the parameters to zero. This is what the LASSO (Least Absolute Shrinkage and Selection Operator) regression does. LASSO imposes a constraint on the sum of the absolute (rather than square) values of the coefficients:

$$\underbrace{\min}_{\beta} \|Y - X\beta\| + \lambda \sum_{i=1}^{M} |\beta_i|. \tag{5}$$

Like the ridge regression, the LASSO departs from OLS estimation and induces a bias, but unlike the ridge it sets some of the coefficients to exactly zero. Thus, in technical terms, it achieves *sparsity*; that is, it reduces a large regression problem to a simpler model in which only a few predictors have nonzero coefficients. This can be seen geometrically by noticing that the absolute value function has a kink at zero, while the square function, which is used in the ridge regression, is smooth at zero, see Hastie et al. (2009) for a graphical example in the case $M = 2$.

A.2.2. Elastic net

By augmenting the OLS problem with a weighted average of the ridge and LASSO penalties, one obtains the elastic net estimation. In this problem, the parameter estimates are the solution to the following minimization problem:

$$\underbrace{\min}_{\beta} \|Y - X\beta\| + \lambda_1 \sum_{j=1}^{M} \left|\beta_j\right| + \lambda_2 \sum_{j=1}^{M} \beta_j^2. \tag{6}$$

In the elastic net estimation, shrinkage depends on two tuning parameters, λ_1 and λ_2. Zou and Hastie (2005) show that the elastic net can be more efficient than LASSO when the number of predictors M is higher than the number of observations T.

A.2.3. LARS as a general algorithm for soft thresholding

Unlike the ridge regression, neither the LASSO nor the elastic net has a closed form solution, so that quadratic programming techniques are needed to solve the respective minimization problems. However, Efron, Hastie, Johnstone, and Tibshirani (2004) show that these two minimization problems are closely related to a more general selection algorithm, the LARS (least angle regression). LARS starts by identifying the covariate X_i that has the highest correlation with the target Y. To understand the LARS algorithm, we will start from the popular forward selection method. This method starts from an empty model with no explanatory variables, and adds variables one by one until the model cannot be improved significantly (according to a statistical criterion) by adding another variable. This means enriching the model progressively at each step k with the covariate that shows the maximum correlation with the residuals from the orthogonal projection of Y on the $k_{\text{th}} - 1$ covariate, where the only regressor in the model for $k = 1$ is the zero vector. Formally, if $\hat{\mu}_k$is the current estimate of Y with k predictors and $\hat{c} = X'(Y - \hat{\mu}_k)$ is the set of current correlations (where we assume that the data series are standardized), then forward selection uses the following updating rule:

$$\hat{\mu}_{k+1} = \hat{\mu}_k + |\hat{c}_j|\text{sign}(\hat{c}_j)X_j, \tag{7}$$

where j is such that $|\hat{c}_j| = \max(\hat{c})$.

The LARS algorithm starts as forward selection, but the two methods separate at $k = 2$, when a different updating rule is used. The exact details of the LARS updating rule are provided by Bai and Ng (2008, p. 308). The intuition is that while forward selection proceeds by orthogonal projections, the LARS rule uses skewed projections. This implies that covariates which are mildly correlated with other variables that are already in the model set have a chance to enter the set. In this respect, LARS strikes a balance between hard thresholding, which selects too many correlated variables, and forward selection, which turns out to be too parsimonious.

In the LARS regression, the desired shrinkage can be seen as a stopping rule for k. Efron et al. (2004) show that the LARS algorithm encompasses other popular shrinkage methods, including forward selection itself, LASSO and the elastic net.

LASSO can be obtained via the LARS algorithm by imposing a restriction at each step of the algorithm on the sign of the correlation between the new candidate regressor and the residual obtained by the skewed LARS projection in the previous step. To understand this intuitively, let us start again from step 1, where the variable which is most correlated with the target enters the active set. Suppose that this correlation is positive. In selecting the second variable for the active set, LARS is agnostic on the sign of the correlation between the target and the new variable. If one imposes the restriction that the sign of this correlation must not switch, the LASSO regression is obtained. Bai and Ng (2008) show that, in order to reformulate the EN as a LASSO problem, it is sufficient to apply a variable transformation, which can therefore be obtained through the LARS algorithm.

References

Altissimo, F., Marchetti, D., Jr., & Oneto, P. (1997). The Italian business cycle: coincident and leading indicators and some stylised facts. *Giornale degli Economisti e Annali di Economia*, *60*, 147–220.

Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., & Rünstler, G. (2011). Short-term forecasts of euro area GDP growth. *Econometrics Journal*, *14*, C25–C44.

Artis, M., Marcellino, M., & Proietti, T. (2004). Dating business cycles: a methodological contribution with an application to the euro area. *Oxford Bulletin of Economics and Statistics*, *66*, 537–565.

Baffigi, A., Golinelli, R., & Parigi, G. (2004). Bridge models to forecast the euro area GDP. *International Journal of Forecasting*, *20*, 447–460.

Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, *146*, 304–317.

Banbura, M., & Modugno, M. (2010). *Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data.* Working paper series 1189. European Central Bank.

Barhoumi, K., Brunhes-Lesage, V., Darné, O., Ferrara, L., Pluyaud, B., & Rouvreau, B. (2008) *Monthly forecasting of French GDP: a revised version of the OPTIM model*. Working papers 222, Banque de France.

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

Bulligan, G., Golinelli, R., & Parigi, G. (2010). Forecasting monthly industrial production in real-time: from single equations to factor-based models. *Empirical Economics*, *39*, 303–336.

Cristadoro, R., Saporito, G., & Venditti, F. (2013). Forecasting inflation and tracking monetary policy in the euro area: does national information help? *Empirical Economics*, *44*, 1065–1086.

D'Agostino, A., & Giannone, D. (2012). Comparing alternative predictors based on large-panel factor models. *Oxford Bulletin of Economics and Statistics*, *74*, 306–326.

De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, *146*, 318–328.

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*, 371–390.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *Annals of Statistics*, *32*, 407–499.

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2005). The generalized dynamic factor model: one-sided estimation and forecasting. *Journal of the American Statistical Association*, *100*, 830–840.

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

Hahn, E., & Skudelny, F. (2008). *Early estimates of euro area real GDP growth—a bottom up approach from the production side.* Working paper series 975. European Central Bank.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction*. New York: Springer.

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

Kapetanios, G., & Marcellino, M. (2009). A parametric estimation method for dynamic factor models of large dimensions. *Journal of Time Series Analysis*, *30*, 208–238.

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

Marcellino, M., & Schumacher, C. (2010). Factor-MIDAS for now- and forecasting with ragged-edge data: a model comparison for German GDP. *Oxford Bulletin of Economics and Statistics*, *72*, 518–550.

Marchetti, D. Jr., & Parigi, G. (1998). *Energy consumption, survey data and the prediction of industrial production in Italy*. Working paper series 342. Bank of Italy.

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

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, *67*, 301–320.