WORKING PAPER SERIES

Big Data Econometrics: Now Casting and Early Estimates

WORKING PAPER
N. 82
JUNE 2018

Dario Buono, George Kapetanios, Massimiliano Marcellino, Gianluigi Mazzi and Fotis Papailias

Università Bocconi
BAFFI CAREFIN
Centre for Applied Research on International Markets, Banking, Finance, and Regulation

Big Data Econometrics: Nowcasting and Early Estimates[1]

Dario Buono[2] George Kapetanios[3] Massimiliano Marcellino[4]
Gian Luigi Mazzi[5] Fotis Papailias[6]

June 5, 2018

Abstract

This paper aims at providing a primer on the use of big data in macroeconomic nowcasting and early estimation. We discuss: (i) a typology of big data characteristics relevant for macroeconomic nowcasting and early estimates, (ii) methods for features extraction from unstructured big data to usable time series, (iii) econometric methods that could be used for nowcasting with big data, (iv) some empirical nowcasting results for key target variables for four EU countries, and (v) ways to evaluate nowcasts and flash estimates. We conclude by providing a set of recommendations to assess the pros and cons of the use of big data in a specific empirical nowcasting context.

Keywords: Big Data, Nowcasting, Early Estimates, Econometric Methods.
JEL Codes: C32, C53, C55.

[1]We are grateful to Eurostat for supporting this research, Jacopo Grazzini, Katerina Petrova and participants at a Banca d'Italia conference for insightful comments, and Mattia Serrano' for research assistance.

[2]Eurostat, European Commission, Data Analytics for Macro and Finance Research Centre. E-mail: dario.buono@ec.europa.eu

[3]King's College London, Data Analytics for Macro and Finance Research Centre. E-mail: george.kapetanios@kcl.ac.uk

[4]Bocconi University, Baffi-Carefin, Bidsa and CEPR. E-mail: massimiliano.marcellino@unibocconi.it

[5]GOPA. E-mail: gianluigi.mazzi@gopa.lu

[6]King's College London, Data Analytics for Macro and Finance Research Centre.
E-mail: fotis.papailias@kcl.ac.uk

1 Introduction

Rapid technological advancements now allow to transform decisions or actions in our every day life into storable data. For example, mobile phones can indicate our position, via GPS, and our activity status, via participation in social media; browsing activity can be monitored - in the form of "cookies" - to provide personalised advertisements; webcams (and satellites) take vast amounts of pictures and monitor activity during day and night. Moreover, connected machines and sensors continuously collect and transmit information. The transformation of such digital traces results in a very large amount of data - henceforth, big data.

While initially big data have been used mainly in the private sector, they also represent an opportunity in other fields, possibly combined with more traditional data sources. In particular, official statistical agencies and other public institutions, such as central banks and governmental bodies, could also benefit from big data, as indicated for example by the High-Level Group for the modernisation of statistical production and services (HLG) or the Eurostat Task Force on big data.[1]

Big data could be particularly useful for nowcasting and the construction of early / flash estimates, namely, for the production of a preliminary estimate for the contemporaneous value of an economic indicator, which has not yet been officially released. Leading examples are Gross Domestic Product (GDP) and its components, deflators, and fiscal variables, which are typically released at least 30-45 days after the end of the reference month or quarter, and later revised. Nowcasts of monthly variables such as the HICP or confidence indicators, sales, trade and labour market indicators could be also of interest.[2] The gains can be expected to be particularly large for countries with less advanced statistical systems, where big data could replace or complement the collection of more standard information.

Therefore, in this paper we focus on the use of big data for macroeconomic nowcasting and the production of early estimates, by surveying, developing and applying proper data handling techniques combined with state of the art econometric methods. Big data have substantial potential in this context, as timely / continuous / large sets of data should provide new or complementary information with respect

[1] See `https://goo.gl/FzQB68` and `https://goo.gl/xQitj6`, respectively

[2] See Buono et al. (2017) and the references therein for a comprehensive discussion and list of examples.

to standard economic indicators. Our aim is to provide a useful starting guide for applied researchers interested in (i) using big data for macroeconomic nowcasting, (ii) improving the quality of early estimates, (iii) increasing the timeliness of the releases, and (iv) complementing standard nowcasts with uncertainty and directional measures.

A related contribution is Bok et al. (2017), who present the factor-based methodology and data underlying the New York Fed Staff nowcasts for the US. With respect to this interesting paper, we focus more on the use of "proper" big data, or big data based indicators, investigating the performance of various econometric methods for several macroeconomic indicators for some of the largest EU economies.

The rest of the paper is organised as follows. Section 2 presents a typology of big data and discusses a set of issues associated with their use. This section draws on Buono et al. (2017), who also present an extensive literature review on the use of big data in economics, with a special focus on nowcasting and forecasting, which therefore is not repeated here. Section 3 studies the conversion from unstructured data to structured time series. Additional details are provided in two working papers underlying this article, Kapetanios et al. (2017a, 2017b), with the latter also discussing filtering techniques and data pre-treatment[3]. Section 4 presents the data for the empirical application. As an illustration of the methods introduced in Section 3, it constructs uncertainty indexes based on Google trends associated with uncertainty related queries over a comparable sample period. Section 5 reviews econometric methods that can handle big data in a nowcasting context, and compares their performance in an empirical exercise on point, directional and interval nowcasting key variables for the four largest European countries. Based on the previous theoretical discussions and empirical results, Section 6 proposes a set of recommendations for the use of big data for economic nowcasting and early estimation. Finally, Section 7 offers brief overall conclusions.

Samples of the code used in this paper can be found on `https://github.com/eurostat/econowcast`.

[3]Additionally, see Buono et al. (2017), Kapetanios et al. (2017c)

2 Typology of big data and related issues

In this section we discuss the most important big data sources and types, as an introduction to the subsequent analysis. Then, we provide an extensive commentary on issues related with the use of such data. This section is based on Buono et al. (2017), who also provide a survey of economic applications based on various big data types.

2.1 Big data sources

Advancements in computer technology during the last decades have allowed the storage, organisation, manipulation and analysis of vast amount of data from different sources and across different disciplines.

A traditional data source, revamped by the IT developments, is represented by **Business Systems** that record and monitor events of interest, such as registering a customer, manufacturing a product, taking an order, etc. The process-mediated data thus collected by either private businesses (commercial transactions, banking/stock records, e-commerce, credit cards, etc.) or public institutions (medical records, social insurance, school records, administrative data, etc.) is highly structured and includes transactions, reference tables and relationships, as well as the metadata that sets its context. Traditional business data is the vast majority of what IT managed and processed, in both operational and BI systems, usually after structuring and storing it in relational database systems.

A novel data source is represented by **Social Networks** (human-sourced information). This information is the record of human experiences, by now almost entirely digitally stored in personal computers or social networks. Data, typically, loosely structured and often ungoverned, include those saved in proper Social Networks (such as Facebook, Twitter, Tumblr etc.), in blogs and comments, in specialized websites for pictures (Instagram, Flickr, Picasa etc.) or videos (Youtube, etc.) or internet searches (Google, Bing, etc.), but also text messages, user-generated maps, e-mails, etc.

Yet another source of big data, and perhaps the fastest expanding one, is the so-called **Internet of Things**. Machine-generated data are derived from sensors and

machines used to measure and record the events and situations in the physical world. It is becoming an increasingly important component of the information stored and processed by many businesses. Its well-structured nature is suitable for computer processing, but its size and speed is beyond traditional approaches. Examples include data from sensors, such as fixed sensors (home automation, weather/pollution sensors, traffic sensors/webcam, etc.) or mobile sensors (mobile phones, connected cars, satellite images, etc.) but also data from computer systems (logs, web logs, etc.).

The resulting many types of big data have been already exploited in many scientific fields, such as climatology, oceanography, biology, medicine, and applied physics. Specific areas of economics have also seen a major interest in big data and business analytics, in particular marketing and finance. Instead, in conventional macroeconomics there have so far been limited applications, mostly concentrated in the areas of nowcasting/forecasting.

2.2 Types of Big Data by Dominant Dimension

Following, e.g., Doornik and Hendry (2015), we can also distinguish three main types of big data according to their dominant dimension: Fat (big cross-sectional dimension, N, small temporal dimension, T), Tall (small N, big T), or Huge (big N, big T).

Huge numerical datasets, possibly coming from the conversion of even larger but unstructured big data, pose substantial challenges for proper econometric analysis, but also offer potentially the largest informational gains for nowcasting. Fat or Tall datasets can be also relevant in specific applications, for example for microeconomic or marketing studies in the case of Fat data or for financial studies in the case of Tall data. Tall data resulting from targeted textual analysis could be also relevant for macroeconomic nowcasting.

This classification is relevant to associate the proper econometric method to each data type. In this paper we are mainly concerned with Tall and Huge datasets, where Huge datasets will be transformed to Tall following the procedures discussed in Section 3. Basically, the information is condensed in a few indicators which are added to large sets of standard economic and financial indicators.

2.3 Data Issues

The previous discussion has already emphasized several advantages of big data in a nowcasting context, starting with the fact that they provide potentially relevant *complementary* information with respect to standard data, being based on rather different information sets. Moreover, big data are timely available and, generally, they are not subject to subsequent revisions, all relevant features for potential coincident and leading indicators of economic activity. Furthermore, big data could be helpful to provide a more granular perspective on the indicator of interest, both in the temporal and in the cross-sectional dimensions. In the temporal dimension, they can used to update nowcasts at a given frequency, such as weekly or even daily, so that the policy and decision makers can promptly update their actions according to the new and more precise estimates. In the cross-sectional dimension, big data could provide relevant information on units, such as regions or sectors, not fully covered by traditional coincident and leading indicators.

Besides all these actual and potential benefits of big data, it is however important to also consider some possible drawbacks, in particular related to internet data. The use of internet data for the production of official statistics has been evaluated in details, e.g., in the report "Analysis of methodologies for using the Internet for the collection of information society and other statistics"[4]. We are instead mainly interested in those aspects related to the use of big data for nowcasting economic indicators.

A first issue concerns data availability. As it is clear from the data categorization presented above, most data pass through private providers and are related to personal aspects. Hence, continuity of data provision could not be guaranteed. For example, Google could stop providing Google Trends, or at least no longer make them available for free or even totally shut down the service as they did for their Maps Engine. Indeed, they have recently decided to substantially restrict the availability of weekly Trends. Or online retail stores could forbid access to their websites to crawlers for automatic price collection. Or individuals could extend the use of software that prevent tracking their internet activities, or tracking could be more tightly regulated by law for privacy reasons. Similarly, there are Twitter data limitations through

[4]See `https://goo.gl/skaXxJ` for more information

Firehose.

Continuity of data availability is more an issue for the use of internet data in official statistics than for a pure nowcasting purpose, as it often happens in nowcasting that indicators become unavailable or no longer useful and must be replaced by alternative variables. That said, continuity and reliability of provision are important elements for the selection of a big data source.

Another concern related to data availability is the start date, which is often quite recent for big data, or the overall number of temporal observations in low frequency (months/quarters), which is generally low, even if in high frequency or cross-sectionally there can be thousands of observations. A short temporal sample is problematic as the big data based indicators need to be related to the target low frequency macroeconomic indicators and, without a long enough sample, the parameter estimators can be noisy and the ex-post evaluation sample for the nowcasting performance too short. On the other hand, several informative indicators, such as surveys and financial condition indexes, are also only available over short samples, starting after 2000, and this feature does not prevent their use.

One more issue for internet based big data is related to the "digital divide", the fact that a sizable fraction of the population still has no or limited internet access. This implies that the available data are subject to a sample selection bias, and this can matter for their use. Suppose, for example, that we want to nowcast unemployment at a disaggregate level, either by age or by regions. Internet data relative to older people or people resident in poorer regions could lead to underestimation of their unemployment level, as they have relatively little access to internet based search tools, e.g. referred to as sample selectivity and representativeness issue.

For nowcasting, the suggestion is to carefully evaluate the presence of a possible selection bias, but this is likely less relevant when internet based data are combined with more traditional indicators and are therefore used to provide additional marginal rather than basic information. There can also be other less standard big data sources for which the digital divide can be less relevant. For example, use of mobile phones is quite widespread and mobility of their users, as emerging from calls and text messages, could be used to measure the extent of commuting, which is in turn typically related to the employment condition.

Another issue is that both the size and the quality of internet data keeps changing

over time, in general much faster than for standard data collection. For example, applications such as Twitter or WhatsApp were not available just a few years ago, and the number of their users increased exponentially, in particular in the first period after their introduction. Similarly, other applications can be gradually dismissed or used for different uses. For example, the fraction of goods sold by EBay through proper auctions is progressively declining over time, being replaced by other price formation mechanisms.

This point suggests that the relationship between the target variable and the big data (as well as that among the elements of the big data) could be varying over time, and this is a feature that should be properly checked and, in case, taken into consideration at the modelling stage.

Yet another issue, again more relevant for digital than standard data collection, is that individuals or businesses could not report truthfully their experiences, assessments and opinions. For example, some newspapers and other sites conduct online surveys about the feelings of their readers (happy, tired, angry, etc.) and one could think of using them, for example, to predict election outcomes, as a large fraction of happy people should be good for the ruling political party. But, if respondents are biased, the prediction could be also biased, and a large fraction of non-respondents could lead to substantial uncertainty[5].

As for the case of the digital divide, this is less of a problem when the internet data are complementary to more traditional information, such as phone or direct interviews, or indicators of economic performance.

Data could also not be available in a numerical format, or not in a directly usable numerical format. A similar issue emerges with standard surveys, for example on economic conditions, where discrete answers from a large number of respondents have to be somewhat summarized and transformed into a continuous index. However, the problem is more common and relevant with internet data.

A related problem is that the way in which the big data measure a given phenomenon is not necessarily the same as in official statistics, given that the data are typically the by-product of different activities. A similar issue arises with the use of proxy variables in econometric studies, e.g. measures of potential output or in-

[5]See `https://goo.gl/PAkU2x` for more information about types of "response bias" in standard surveys

flation expectations. The associated measurement error can bias the estimators of structural parameters but is less of a problem in a forecasting context, unless the difference with the relevant official indicator is substantial.

Clearly, the collection and preparation of big data based indicators is far more complex than that for standard coincident and leading indicators, which are often directly downloadable in ready to use format from the web through statistical agencies or data providers. The question is whether the additional costs also lead to additional gains, and to what extent, and this is mainly an empirical issue. The literature review we have presented suggests that there seem to be cases where the effort is worthwhile.

A final issue, again common also with standard data but more pervasive in internet data due to their high sampling frequency and broad collection set, relates to data irregularities (outliers, working days effects, missing observations, etc.) and presence of seasonal / periodic patterns, which require properly de-noising and smoothing the data.

As for the previous point, proper techniques can be developed and the main issue is to assess their cost and effectiveness, which is again an empirical issue (which will be shortly considered later on).

A few other, more methodological, potential problems associated with big data are discussed in the next subsection.

2.4 Methodological Issues

As Hartford (2014) put it: " 'Big data' has arrived, but big insights have not. The challenge now is to solve new problems and gain new answers – without making the same old statistical mistakes on a grander scale than ever."

The statistical mistakes he refers to are well summarized by Doornik and Hendry (2015): " ... an excess of false positives, mistaking correlations for causes, ignoring sampling biases and selecting by inappropriate methods."

An additional critic is the "big data hubris", formulated by Lazer et al. (2014): ""Big data hubris" is the often implicit assumption that big data are a substitute for, rather than a supplement to, traditional data collection and analysis. They also identify "Algorithm Dynamics" as an additional potential problem, where algorithm

dynamics are the changes made by engineers to improve the commercial service and by consumers in using that service. Specifically, they write: "All empirical research stands on a foundation of measurement. Is the instrumentation actually capturing the theoretical construct of interest? Is measurement stable and comparable across cases and over time? Are measurement errors systematic?".

Yet another caveat comes from one of the biggest fans of big data. Hal Varian, Google's chief economist, in a 2014 survey wrote: "In this period of big data it seems strange to focus on sampling uncertainty, which tends to be small with large datasets, while completely ignoring model uncertainty which may be quite large. One way to address this is to be explicit about examining how parameter estimates vary with respect to choices of control variables and instruments."

In our nowcasting context, we can therefore summarize the potential methodological issues about the use of big data as follows.

First, do we get any relevant insights? In other words, can we improve nowcast precision by using big data? As we mentioned in the previous subsection, this is mainly an empirical issue and, from the studies reviewed in the previous sections, it seems that for some big data and target variables this is indeed the case.

Second, do we get a big data hubris? Again as anticipated, we think of big data based indicators as complements to existing soft and hard data-based indicators, and therefore we do not get a big data hubris (though this is indeed the case sometimes even in some of the nowcasting studies, for example those trying to anticipate unemployment using only Google Trends).

Third, do we risk false positives? Namely, can we get some big data based indicators that nowcast well just due to data snooping? This risk is always present in empirical analysis and is magnified in our case by the size of the dataset since this requires the consideration of many indicators with the attendant risk that, by pure chance, some of them will perform well in sample. Only a careful and honest statistical analysis can attenuate this risk. In particular, as mentioned, we suggest to compare alternative indicators and methods over a training sample, select the preferred approach or combine a few of them, and then test if they remain valid in a genuine (not previously used) sample.

Fourth, do we mistake correlations for causes? Again, this is a common problem in empirical analysis and we will not be immune for it. For example, a large number

of internet searches for "filing for unemployment" can predict future unemployment without, naturally, causing it. This is less of a problem in our nowcasting context, except perhaps at the level of economic interpretation of the results.

Fifth, do we use the proper econometric methods? Here things are more complex because when the number of variables N is large we can no longer use standard methods and we have to resort to more complex procedures. Some of these were developed in the statistical or machine learning literatures, often under the assumption of i.i.d. observations. As this assumption is likely violated when nowcasting macroeconomic variables, we have to be careful in properly comparing and selecting methods that can also handle correlated and possibly heteroskedastic data. This is especially the case since these methods are designed to provide a good control of false positives, but this control depends crucially on data being i.i.d. To give an example, it is well known that exponential probability inequalities, that form the basis of most methods that control for false positives, have very different and weaker bounds for serially correlated data leading to the need for different choices for matters like tuning parameters used in the design of the methods[6]. Overall, as we will see, a variety of methods are available, and they can be expected to perform differently in different situations, so that also the selection of the most promising approach is mainly application dependent.

Sixth, do we have instability due to Algorithm Dynamics or other causes (e.g., the financial crisis, more general institutional changes, the increasing use of internet, discontinuity in data provision, etc.)? Instability is indeed often ignored in the current big data literature, while it is potentially relevant, as we know well from the economic forecasting literature. Unfortunately, detecting and curing instability is complex, even more so in a big data context. However, some fixes can be tried mostly borrowing from the recent econometric literature on handling structural breaks.

Finally, do we allow for variable and model uncertainty? As we will see, it is indeed important to allow for both variable uncertainty, by considering various big data based indicators rather than a single one, and for model uncertainty, by comparing alternative procedures and then either selecting or combining the best performing ones. Again, all issues associated with model selection and uncertainty are likely magnified due to the fact that large data also allow for larger classes of

[6]See Roussas (1996) and the references therein for more information

models to be considered, and model selection methods, such as information criteria, may need modifications in many respects.

3 A General Data Conversion Framework for Unstructured Numerical Big Data

A rather common feature of big data is the lack of a structure that makes them directly suitable for econometric or statistical analysis. This section aims to offer a general unstructured data conversion framework.

Data structuring can be made in different ways, producing different results according to the targeted phenomenon, the granularity of the phenomenon and the characteristics of the unstructured data. Also, the dimensions of the structured dataset can have relevant impact on the modelling process. In particular, if the output of the structuring process is just few to many time-series, then common econometric methods for small or large datasets can be used, while specific techniques are needed for really big, and possibly sparse, datasets. As mentioned, additional details are provided in two working papers underlying this article, Kapetanios et al. (2017a, 2017b), with the latter also discussing filtering techniques and data pre-treatment.

3.1 Conceptual Setting

We start by considering an unstructured dataset which consists of N events, each described by a vector of the form $y_i = (y_{1,i}, ..., y_{m,i}, t_i)' = (\tilde{y}_i, t_i)'$, $i = 1, ..., N$, where N is potentially very large. The vector $\tilde{y}_i'$ contains information on the event, while t_i denotes the time the event has occurred, and is referred to as a time stamp. This definition of an event in terms of a vector of characteristics and time stamp is very general. It is difficult to envisage any big dataset with a temporal dimension that cannot be accommodated by this definition. Time is defined continuously in an interval $(0, T]$. Our aim is to describe a mapping between these events and potential time series defined in discrete time, $t = 1, ..., T$, which can then be used for nowcasting or other econometric analyses.

We first need to define a mapping between the set $\mathcal{T} = \{t_i\}_{i=1}^N$ and time series observations. That is, a set function given by $\mathcal{T}_t = f(\mathcal{T}, t)$ where $\mathcal{T}_t$ is a subset of $\mathcal{T}$

containing the event time stamps that relate to time period t. The obvious mapping is for $\mathcal{T}_t$ to contain the events whose time stamp satisfies $t-1 < t_i \leq t$ but more complex setups, e.g. with lags, are possible. For example, it may be that $\mathcal{T}_t$ contains the events whose time stamp satisfies $t-p < t_i \leq t-p+1$.

Then, we can define a time series as

$$x_t = \sum_{t_i \in \mathcal{T}_t} g_t\left(y_i, \gamma\right), \tag{1}$$

where $g_t\left(y_i, \gamma\right)$ is a function possibly depending on t (to reduce the notational burden we sometimes suppress the t subscript), and parameterized by γ, mapping information on the event (contained in $\tilde{y}_i$) into a single number. This is a very general formulation and our discussion below provides ways in which we can better interpret its applicability. Before proceeding, it is important to note that the only restriction imposed here is a form of linear separability across events. The alternative would be a formulation of the type $x_t = g_t\left(y_1, ..., y_{\mathcal{T}_t}; \gamma\right)$. This would provide the ability to aggregate events in a nonlinear fashion, but we feel this might be unwieldy in practice. We also note that there is the potential for different time scales to be used in defining the mapping from events to time series. In particular, we can use the above mapping framework to construct, for example, both quarterly and monthly time series which can then be used by considering mixed frequency methods. Finally, we note that missing fields or elements in y_i can simply be treated by removing the events or, if the incidence of missing values is great, by setting indicator functions, associated with selecting particular features, to zero.

3.2 Aggregation

We suggest to consider g functions of the type:

$$g_t\left(y_i, \gamma\right) = \sum_{j=1}^{m} w_{tj} g_{tj}\left(y_{j,i}, \gamma_j\right), \tag{2}$$

where the w_{tj} are weights generally depending on t. For example, the function could be $g_{tj}\left(y_{j,i}, \gamma_j\right) = 1$, or $g_{tj}\left(y_{j,i}, \gamma_j\right) = y_{j,i}^{\gamma_j}$ or $g_{tj}\left(y_{j,i}, \gamma_j\right) = y_{j,i} I(|y_{j,i}| > \gamma_j)$, with $(w_{t1}, w_{t2}, ..., w_{tm})$ equal to, e.g., $(1, 1, ..., 1)$ or $(1/m, 1/m, ..., 1/m)$ or even $w_{tj} = \frac{\tilde{w}_{tj}}{\mathcal{T}_t}$,

thereby providing a useful normalisation for the summation.

The function in (2) is a linear transformation of (possibly non-linear) g_{tj} functions. As an alternative, a non-linear transformation could be considered. Moreover, the w_{tj} weights could be also optimally computed given a certain error loss function.

An interesting possibility is to cast in this set-up the construction of internet search-based indicators, such as Google trends. In this context, $(y_{1,i}, ..., y_{m,i}) = (y_1, ..., y_m)$ denotes the numerical representation of alphanumeric sequences that describe generic internet searches, for example y_1 corresponds to "recession" and y_2 to "income" (so $m = 2$ in this example).[7] Then, we can define

$$g(y_i, \gamma) = I(y_i \in \mathcal{G}_{t_i}(\gamma)), \tag{3}$$

where $\mathcal{G}_{t_i}(\gamma)$ denotes (possibly a subset of) all Google searches during period t_i (transformed into numerical representation), and may depend on the parameter vector γ (for example, γ can select searches for only a specific country or in a specific language), so that the function $g(y_i, \gamma)$ counts how many times y_i (a search for both "recession" and "income") happened in period t_i. Computing $g(y_i, \gamma)$ for $i = 1, ..., N$ and plugging the results into (1) returns a "Google trend" time series for period $t = 1, 2, ..., T$. Naturally, if $m = 1$ only single keyword searches are considered, so we would get two separate Google trends for "recession" and "income".

Once parameterised by, e.g., setting a subset of w_j to zero and choosing the functional form for g_j, the parameters , w_j, γ_j, $j = 1, ..., m$ can be either fixed a priori or calibrated. The most obvious approach is to consider many different events $\tilde{y}_i$ and/or functions g, by, e.g., defining a grid for γ and w, and so obtain a large number of time series indicators, which can then be handled by appropriate econometric methods.

An interesting alternative is to pin down the function g by considering which

[7]It is important to note that the list of key words that should be entered into Google Trends is a matter for consideration. However, in our view the wider the list the better, since our recommendation is that at a later stage all these variables constructed via Google Trends (or other Google facilities such as Google Correlate) can be analysed in a regression setting using various data rich methodologies. As a result one can design a wide variety of relevant keywords (the use of a dictionary or a thesaurus may be useful here, together with an automatic translation service such as Google translate) and these can then be used on their own or combined through the use of logical relationships based on "or" or "and" to provide the final list.

choice of g has the best forecasting ability for the target variable when used on its own or, potentially, in combination with a lag of the target variable.

3.3 Features Extraction

Apart from the above function, we can extract features from the distribution of the events which occur at time t. Depending on the nature of the underlying data, we could extract the variance, various percentiles, the skewness and kurtosis, or even higher moments, and then use the resulting time series of big data features.

Formally, and assuming that N events occur at a particular time period, t, we can extract the features using:

$$\text{Variance: } g_t(y_i) = \frac{1}{N}\sum_{i=1}^{N}(y_{it} - \overline{y}_t)^2, \tag{4}$$

$$\text{P-Percentile: } g_t(y_i) = y_{\lceil \frac{P}{100}N \rceil t}, \text{ where } \lceil \cdot \rceil \text{ denotes ordinal ranking,} \tag{5}$$

$$\text{Skewness: } g_t(y_i) = \frac{\frac{1}{N}\sum_{i=1}^{N}(y_{it} - \overline{y}_t)^3}{\left[\frac{1}{N-1}\sum_{i=1}^{N}(y_{it} - \overline{y}_t)^2\right]^{3/2}}, \tag{6}$$

$$\text{Kurtosis: } g_t(y_i) = \frac{\frac{1}{N}\sum_{i=1}^{N}(y_{it} - \overline{y}_t)^4}{\left[\frac{1}{N}\sum_{j=1}^{N}(y_{it} - \overline{y}_t)^2\right]^2} - 3, \tag{7}$$

where $\overline{y}_t$ denotes the sample mean of N events at time t.

These measures are rather intuitive, with the specific choice depending on the nature of the big data and the purpose of the exercise.

3.4 Data Mining

Another approach to reduce dimensionality is data mining and, in particular, clustering. To keep things simple and intuitive, we illustrate this approach by means of a basic clustering method, which is the $k - means$ clustering.

$k - means$ clustering aims to partition N observations into k clusters, with generally $k << N$, in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. Given a set of observations, $(y_1, ..., y_i)$,

$k-means$ clustering aims to split the data into k sets, e.g. a partition $S = (S_1, ..., S_k)$, so as to minimise the within-cluster sum of squares which is the sum of distance functions of each point in the cluster to the K center. The relevant objective function is:

$$\arg_s \min \sum_{j=1}^{k} \sum_{y_i \in S_j} \|y_i - \mu_j\|^2 , \tag{8}$$

where μ_j is the mean of points in S_j. This will return k centers. In this way, the researcher succeeds in the reduction of the number of events to be considered. This method is particularly useful when similar observations repeatedly appear in the sample. In this case simple aggregation could be suboptimal, as it could hide substantial heterogeneity. See the seminal work of Lloyd (1982) for a starting point.

3.4.1 Random Subsampling

The above time series conversion techniques are expected to work well in applications where the researcher is interested in summaries and/or features of the data and where the size of the dataset allows the required computations. However, how should we approach cases where data is huge? For example, consider datasets which are measured in terabytes, petabytes, exabytes, zettabytes or even yottabytes. Can these conversion techniques still be applied? In cases where even database managers struggle to work well with software to construct structured time series, we can adopt the random subsampling approach; see Ng (2016) and references therein.

In this approach, random subsamples of the data are selected to create a new big dataset which is significantly smaller. This can be related to Equation (1) using indicator functions. For example, we could select observations of random seconds in an hour, or random hours in a day or even random days in a month. Then, the conversion techniques as described above can be applied and continue in the standard way. In this cost-efficient way, the researcher can randomly extract a large number of subsamples, which consequently leads to a large number of structured time series. At this point, an average of the resulting time series can be used or principal components can be extracted. The issue here is mostly one of data management and less of econometric difficulty, since a number of methods stay, in principle, valid for any dataset size while the informational loss associated with subsampling has an

unknown cost.

4 Data

In this Section we discuss, in turn, the target variables for the subsequent nowcasting exercise; the macroeconomic and financial predictors; two big data based uncertainty indexes, which also provides an illustration of some of the techniques for structuring and summarizing unstructured big data presented in the previous Section; the variable transformations; and issues related with timing, mixed frequency and unbalancedness.

4.1 Targets

We consider the four largest EU economies: Germany (DE), France (FR), Italy (IT) and the UK. For each of them, we have collected data on three key monthly economic indicators: the harmonised index of consumer prices (HICP), the industrial production (IP), and the unemployment rate (UR). The datasets considered are made available by Eurostat and have been downloaded from their online dissemination database.[8]

We have also considered quarterly GDP as a target variable. However, as discussed below, the evaluation sample is even shorter than for the monthly indicators. Hence, we prefer to make the results for GDP available upon request.

4.2 Predictors

Our set of monthly macroeconomic predictors includes various coincident and leading indicators plus additional key economic variables for each country. Specifically, we consider: Bank Lending Rate, Bankruptcies, Capital Flows, Car Registrations, Construction Output, Consumer Credit, Core Consumer Prices, various CPI components, Crude Oil Production, Export Prices, Exports, Gasoline Prices, House Price Index, Import Prices, Imports, Job Vacancies, Manufacturing Production, Mining

[8]The Eurostat codes we use are: target.ei_isin_m.XX, target.prc_hicp_midx.XX and target.ei_lmhr_m.XX for IP, HICP and UR respectively. XX denotes the corresponding country code. Data can be downloaded from `https://goo.gl/pThhqB`.

Production, Money Supply, M1, M2 and M3, Producer Prices, Retail Sales, Steel Production, Youth Unemployment Rate, Consumer Confidence Indicators and various surveys.

Our set of weekly variables includes mainly financial indicators: interest rates at various maturities and spreads (yield curve), equity indexes and volatility indexes.

Finally, we have constructed a big data uncertainty indicator based on Google searches, which is discussed in more details below. There is a growing literature suggesting that uncertainty can have substantial negative effects on the economy, starting with the pioneering work of Baker et al. (2016); see also Jurado et al. (2015) and Carriero et al. (2018). Hence, it is interesting to assess whether big data based uncertainty indicators are also helpful for nowcasting.

Table 1 lists the variables we consider, along with their frequency and transformation type. For each country we use the corresponding variables.

4.3 The Google Uncertainty Index

The Google Uncertainty Index is based on the use of Google Trends. In particular, we obtain Google Trends for the following keywords:

- Germany: "unsicherheit" and "risiko" across web and news searches in the region of Germany.
- France: "incertitude" and "risque" across web and news searches in the region of France.
- Italy: "incertezza" and "rischio" across web and news searches in the region of Italy.
- UK: "uncertainty" and "risk" across web and news searches in the region of the UK.

To construct the general (non-country specific) Google based uncertainty index we consider the keywords "uncertainty", "economic uncertainty", "financial uncertainty" and "policy uncertainty".

The final national and global Google uncertainty indexes are the equally weighted average of the respective Google Trends. We also tries to use data-dependent weights

based on the spectral entropy of each series[9]. However, the resulting uncertainty indexes do not change significantly and we omit the results from presentation.

Comparing the Google based uncertainty indexes to the Economic Policy Uncertainty (EPU) index of Baker et al. (2016), we see that the two indexes are positively correlated with a coefficient of 0.69 across the full sample. In particular, before 2012 the correlation coefficient of the two indexes is 0.45 and increases to 0.66 in the period after 2012; see Figure 1.

[9]Lower spectral entropy values indicate that the series is more predictable

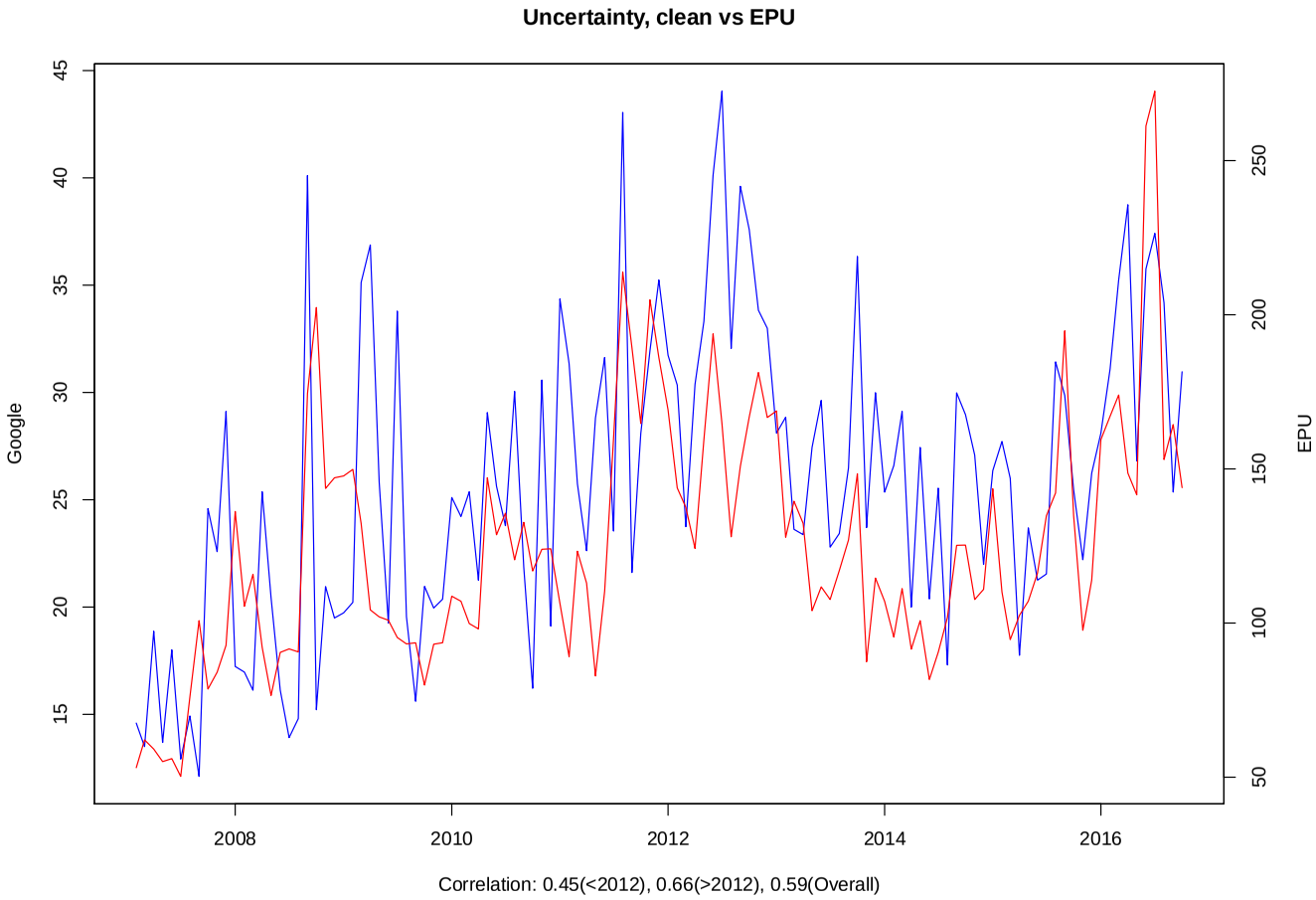

Figure 1: Comparing the General Uncertainty Index of Google (left axis, blue colour) to the corresponding EPU index (right axis, red colour).

4.4 Time Span

Our monthly variables (including monthly targets) span from 2007-01-31 to 2016-10-31 (118 months). The nowcasting exercise starts in 2014-01-31, to ensure that there is sufficient data to be used in the estimation also for the complex econometric models. A lag of the target variable is included in most specifications, i.e., an autoregressive term. However, lags of the predictors are not included, to avoid overfitting due to the short sample span.

This results in 34 evaluation periods (2014-01-31 to 2016-10-31) for our nowcasting exercise. The short evaluation sample should be kept in mind when assessing the empirical results. As mentioned, this is a common problem when using big data as the collection starts only recently.

4.5 Transformations

To ensure that the variables under analysis are stationary, a pre-requisite for several of the econometric methods we will implement, we use a set of transformations, which include: (i) log, (ii) first difference, (iii) percentage change, (iv) log difference, (v) second log difference. The specific transformation for each variable follows standard practice in the literature, see, e.g., McCracken and Ng (2015) as well as Stock and Watson (2002a and 2002b) among others. Table 1 provides a detailed list of the transformation used for each variable.

We nowcast the period-to-period percentage change of IP, HICP and GDP and the period-to-period first difference of UR. The nowcasts for most models are first produced using growth rates, and then translated to levels, as Eurostat and other official statistical agencies typically publish their nowcasts in levels.

4.6 Timing, Mixed-Frequency & Unbalancedness

For each variable, we mark the publication delay and the day of the month that the release for this variable is due. For example, suppose that industrial production for month T is released on the 25th day of the next month, month T+1. In that case, we note one time period publication lag (i.e., -1) and the 25th day as publication release. Repeating this procedure for each variable allows us to construct a

nowcasting exercise that is accurate (on average) in terms of the information which is available at each point in time. For example, when we are about to construct nowcasting estimates five weeks prior to the official release, we use only the available information up to that point. The use of real time data would be even preferable, but unfortunately real time data are not available for most of the predictors used in the analysis.

In our nowcasting exercise we transform weekly observations to monthly (or quarterly) by averaging the available information in each month (or quarter), as routinely done when constructing bridge models. This creates an unbalanced panel of variables in which weekly-to-monthly variables have zero publication lag, i.e., information is up-to-date, but macroeconomic variables might have 1, 2, or more periods of publication lags. In such cases of unavailable information, we assign missing values.

The U-MIDAS approach could also be an attractive alternative, in general, to deal with mixed frequency data. However, in our specific context the available sample is quite short, the difference in sampling frequency is rather high compared to the sample size, and the number of weeks in a month changes over time. The MIDAS approach could reduce the number of parameters to be estimated but it would introduce non-linearity, which would make an estimation with a large number of explanatory variables practically unfeasible. Hence, our preference for simple temporal aggregation. See Ghysels et al. (2004) and Foroni et al. (2015), among others, for MIDAS and U-MIDAS models. Limited experimentation with the UMIDAS approach did not yield major improvements with respect to the averaging approach.

To deal with the missing values at the end of the sample due to publication delays, we adjust each series by moving it forward so that the last observed data point matches the observed value in the dependent variable. Other solutions could be to replace the missing values with the median or mean, or to use extrapolation employing a simple autoregressive (AR) or other model. However, the median or mean could be different from the recent trend of the variables, and extrapolation is complex when dealing with a very large dataset. Hepenstrick and Marcellino (2018) discuss these various options. Limited experimentation in our context did not yield major improvements with respect to the vertical re-alignment approach.

5 Econometric Methods, Evaluation Criteria & Empirical Results

In this section we provide a short overview of approaches that can be used for macroeconomic nowcasting using big data based indicators, rather than directly big data. This may seem restrictive but, as we have discussed, it is typically convenient to structure and reduce the dimensionality of big data prior to using them for macroeconomic forecasting. Next, we discuss the construction and evaluation of point, directional and density nowcasts. Finally, we present and comment the empirical results.

5.1 Econometric Models

In the nowcasting exercise we employ several methodologies to assess the relative gains from more complex methods that can handle large datasets with respect to simpler procedures. Specifically, we consider (with the number of models in each method reported in parentheses):

- (7) Naive and AR models. The first group of models consists of some simple models which assume that the best nowcast is given by: the average of the last four periods ($Avg(4)$), the average of the last twelve periods ($Avg(12)$) and the average of the last twenty-four periods ($Avg(24)$). The $Naive$ model uses the last observed value as the best nowcast (which coincides with that from a pure random walk model). Then, we include an $AR(1)$, $AR(4)$ and an $AR(p_{AIC})$ model where the p_{AIC} is determined via the Akaike's Information Criterion.

- (6) Simple Linear Regressions (LinReg). These are simple specifications using the Google (G) Uncertainty Indexes with zero, one and three lags of the target variable.

- (8) Various other univariate models. The following models have been shown to work well in various forecasting competitions by the International Journal of Forecasting. More details can be found in Hyndman and Khandakar (2008).

- AutoArima: chooses the best ARIMA(p,d,q) model using AIC. Transformation of the univariate series not necessary as the model handles integrated series.
- ETS and BaggedETS: Exponential smoothing methods as in Hyndman et al. (2002) and Bergmeir et al. (2016). The methodology is fully automatic and performed extremely well on the M3-competition data. The bootstrapped series are obtained using the Box-Cox and Loess-based decomposition (BLD) bootstrap; see Bergmeir et al. (2016).
- BATS and TBATS: This class of models are exponential smoothing state space models with Box-Cox transformation, ARMA errors, Trend and Seasonal components. It is part of automatic procedure forecasting. See De Livera et al. (2011).
- Neural Networks (NN): This is a simple feed-forward neural network with a single hidden layer and lagged inputs.
- Spline forecasts: this model produces local linear nowcasts using cubic smoothing splines.
- Theta method: The theta decomposition method of Assimakopoulos and Nikolopoulos (2000).

- (16) Dynamic Factor Analysis (DFA). Models in this class have been the standard choice for nowcasting purposes with large datasets, see, e.g., Bok et al. (2017). We use the default setup of Giannone et al. (2008) with $(q, r, p) = (2, 2, 1)$ where q is the dynamic rank, r is the static rank and p is the AR order of the state vector. We also try three more settings with $(q, r, p) = (3, 3, 1)$, $(q, r, p) = (4, 4, 1)$ and $(q, r, p) = (5, 5, 1)$. For each setting we use: (i) the set of macroeconomic and financial indicators only (MF), and (ii) the set of macroeconomic and financial indicators including the Google Uncertainty Indexes (MF-G).

- (20) Partial Least Squares (PLS). We use PLS to extract one, two, three, four and five factors; the resulting models and nowcasts are labeled, respectively, $PLS(1)$, $PLS(2)$, $PLS(3)$, $PLS(4)$ and $PLS(5)$. As above, for each model

we use MF and MF-G. PLS has the advantage with respect to DFA that the estimated factors are targeted towards the specific variables to be forecasted.

- (20) Sparse Principal Components (SPC). We use SPC to extract one, two, three, four and five factors; the resulting models and nowcasts are labeled, respectively, $SPC(1)$, $SPC(2)$, $SPC(3)$, $SPC(4)$ and $SPC(5)$. As above, for each model we use MF and MF-G. SPC can improve upon standard Principal Components (PC) in the presence of "sparse" variance covariance matrixes of the explanatory variables, which are common with various types of big data (though perhaps not so much with economic variables that tend to be correlated).

- (8) LASSO and Elastic Net (EN). We use the standard 10-fold cross-validation to determine the value for λ in LASSO. The chosen value is the one which minimises the in-sample MSE. As above, we use MF and MF-G. Both methods can be given a Bayesian interpretation and, as all shrinkage estimators, they trade-off bias versus efficiency, which can pay off particularly in a forecasting context.

- (4) Spike and Slab (SS) regressions using MF and MF-G. These specific priors have been often used in empirical studies with big data based indicators, see, e.g., Scott and Varian (2013).

- (4) Data-Driven Automated Forecasting Strategies. On top of the above methodologies we introduce some automated data-driven *"forecasting strategies"*. Our idea is simple and intuitive: we suggest the use of a *"model rotation"* strategy which chooses the model with the smallest cumulative nowcast error. Furthermore, we use an equally-weighted average of the top three, five and ten models. Forecast combination has a long tradition of good empirical performance, noted already by Bates and Granger (1969).

In addition to the above procedures, we also consider the inclusion of a lag in the set of predictors and let each method choose the necessary variables unconditionally (even though in some cases a lag might not be chosen as significant in terms of nowcasting/forecasting).

5.2 Nowcasting Exercise

The nowcasting exercise is based on the algorithm described in the following steps.

1. First, we leave a number of observations, T^{OUT}, out-of-sample, in order to use them in the evaluation of the nowcasting performance of different models. In our experiments, $T^{OUT} = 34$ (with monthly data).

2. The initial sample we use in the first round of estimation and nowcasting is $T_1^{IN} = \{1, .., \big(T - T^{OUT} + 1\big)\}$. Then, we estimate the parameters and produce the nowcasts from the various models. We construct nowcasts for $h = \{-5, -4, -3, -2, -1\}$ weeks prior to the target date when the official release is due. For each h, we keep the same target date, however we re-estimate and produce different nowcasts (updates) using all available information up to that time point. This produces five estimates for each corresponding target date.

3. We repeat Step 2 in a recursive manner, i.e. $T_2^{IN} = \{1, .., \big(T - T^{OUT} + 2\big)\}$ and generally $T_j^{IN} = \{1, .., \big(T - T^{OUT} + j\big)\}$. We stop when $T_j^{IN} = \{1, .., (T - 1)\}$, as we always need the true value of the next period to evaluate the nowcasts.

At the end of the above recursive procedure we end up with T^{OUT} nowcasts for each model under consideration.

5.3 Nowcasts Evaluation

Once we have computed T^{OUT} nowcasts for 5 to 1 weeks prior to the release with each of the many alternative econometric models, and transformed them in levels, we evaluate their performance using the root mean squared forecast error (RMSFE), defined as:

$$RMSFE_{i,h} = \left(\frac{1}{T^{OUT}} \sum_{t=1}^{T^{OUT}} e_{i,h,t}^2 \right)^{\frac{1}{2}},$$

where $e_{i,h}$ is the out-of-sample forecast error (in levels) for model i and weekly nowcast h weeks prior to the release.

In our results we present the RMSFE across different horizons to illustrate how nowcast errors change as we approach the release date. We report the RMSFE of each model relative to $AR(1)$, a simple and common benchmark to allow an intuitive cross-comparison of the results. We have also calculated the relevant p-value for the two-sided Diebold and Mariano (1995) test statistic. Often the differences are not statistically significant, due to the short evaluation sample, hence we do not present the results but make them available upon request.

We also evaluate the directional forecasting precision, using the Sign Success Ratio (SSR) defined as:

$$SSR_{i,h} = \frac{\sum_{t=1}^{T^{OUT}} I\left(\text{Sign}\left(\widehat{y}^{f}_{i,h,t}\right)\right)}{T^{OUT}}, \tag{9}$$

where $\widehat{y}^{f}_{i,h,t}$ denotes the forecast for model i made h weeks prior to the release, and $I\left(\text{Sign}\left(\widehat{y}^{f}_{i,h,t}\right)\right)$ is an indicator function that receives the value 1 if $\text{Sign}\left(\widehat{y}^{f}_{i,h,t}\right)$=Sign$(y_t)$ and 0 otherwise. Hence, $0 \leq SSR_{i,h} \leq 1$, with values close to 1 indicating that model i correctly predicts the "direction" of the target variable during the evaluation period.

Of course, the drawback of SSR is that it does not consider the size of the forecast error but only its sign. For example, we might have a model that correctly predicts the direction 95% or even 100% of the times but the forecasts could be very different the actual values. Therefore, SSR should be used as a complement to the standard $RMSFE$.

Point forecasts are informative but do not provide a measure of uncertainty around them, which can be relevant in particular when using the forecasts for decision making. Density forecasts can be better suited in this context. We can construct density forecasts f_t making a Normality assumption so that, over the evaluation sample, it is

$$f_t \sim N(\widehat{y}^{f}_{i,h,t}, V(e_{i,h,t})),$$

where $t = 1, ..., T^{OUT}$.

As our evaluation sample is short, it is preferable to focus on interval forecasts, which can be easily constructed starting from the density forecasts f_t. Their evaluation is rather simple. In particular, let us assume that we construct forecast intervals at the 60% confidence level, which correspond, respectively, to the (0.2, 0.6) quan-

tiles of the density forecast f_t.[10] We can then compute the empirical coverage rate, measuring the number of times the true values lie inside the specified intervals, and compare it with the theoretical confidence level.

5.4 Discussion of the Results

The main goal of the empirical exercise is to assess whether the use of big data, as proxied by Google based uncertainty indexes, can, either in isolation or in combination with high frequency economic and financial indicators, improve the precision of nowcasts and flash estimates. We are interested in gains in terms of RMSFE for point forecasts, sign success ratio for directional forecasts, and coverage rate for interval forecasts.[11]

5.4.1 Presentation of the results

We have a massive amount of results, due to the many models, forecast horizons, variables, countries, and evaluation criteria we consider. To provide a focused summary of them, we group the models in five categories and in the main text we only report results for the best model across each category and overall.[12] Tables 2 to 13 present, for each of the four countries, the results using RMSFE, SSR and coverage rate. Each table consists of three panels for, respectively, HICP, IP and UR.

The five model categories are:

- **AveNaive**: simple averages and naïve forecasts
- **TS**: univariate time series models
- **G**: any model which includes Google-based indicators
- **MF**: models which only include standard Macro and Financial indicators
- **Best**: best model selection strategies.

[10] More generally, the researcher could construct $(1-\alpha)\%$ interval nowcasts based on the $(\alpha/2, 1-\alpha/2)$ quantiles.

[11] We do not consider density forecasts, even though they could be easily computed, as the forecast sample is too short for a reliable evaluation.

[12] Additional detailed results are available in a separate Online Appendix.

In each table, after the columns with the best models within each category and overall Top-5 models, we provide the corresponding statistics for each model. The nowcast horizons are $h = \{-5, -4, -3, -2, -1\}$ weeks prior to the official release of each target variable.

5.4.2 RMSFE

Tables 2 to 5 cover the (point) nowcasting results for the four countries under consideration and the three monthly target variables. The reported statistic is the RMSFE relative to an AR(1) benchmark, to facilitate cross-model comparison.

Starting with Table 2, we see that in the Top 5 models there are very often specifications based also on the big data uncertainty indicators, particularly so for HICP and UR, less so for IP. Looking in details at the three variables, for HICP, AR(AIC) seems to be the best performing method overall. However, for early estimates (e.g., -5w and -4w), sparse regressions are better. At all horizons, models with the big data uncertainty indicator are ranked second, and the differences with respect to the top models are small. For IP, univariate time series methods are best. Only for -5w and -4w Spike and Slab regressions are included in the top models, with minor differences in terms of RMSFE. UR is instead a clear case where big data indicators yield nowcasting gains. From -5w to -1w the Top 2 models always include big data indicators. Univariate models also perform adequately, with an average RMSFE of 0.859 across all horizons, but they are outperformed by other more complex methodologies, such as factors and sparse regressions, though the differences are not large.[13]

Continuing with France, from Table 3, for early estimates of the HICP the elastic net with standard macro finance variables and the LASSO with big data indicators (EN-MF and LASSO-MF-G) are the top models. However, and as for Germany, as we approach the official release date (and more timely data on past HICP become available), AR(AIC) performs better. For IP, the univariate models are the clear

[13]It is worth noting that for simple, naïve and univariate time series models there are often changes in the criterion value between -5w, -4w and -3w, and then sometimes from -3w to -2w and -1w. This pattern is explained by publication lags. For example, consider we want to nowcast the target variable at time t. At -5w and -4w, the latest actual observation we have in our sample is for $t-3$. However, usually between -4w and -3w (particularly, closer to -4w) the next observation, for $t-2$, becomes available. Finally, between -2w and -1w (often closer to -1w), the value for $t-1$ is also released. The exact release pattern sometimes changes during the evaluation sample.

winner, again as for Germany. For UR, as for HICP, multivariate models including the uncertainty indicators are best at longer nowcast horizons (-5w, -4w, -3w). At shorter horizons, univariate models perform best, though the difference with respect to more complex specifications, with or without the big data indicators, are small.

Next, Table 4 illustrates the results for Italy. For both HICP and IP, overall the univariate models outperform all other candidate models. Yet, for HICP, an AR model augmented with the Google based uncertainty indicator (Linreg_G-L3) is often included in the top-5 models. For UR, the same model is the top performer at the -3w, -2w and -1w horizons.

Finally, from Table 5, for UK HICP specifications including big data based indicators are often in the top-5 models. At shorter horizons, the Linreg_G-L3 model is the second best (after AR(AIC)). For UR, there are sometimes models including big data based indicators among the top performers, while this is generally not the case for IP. However, for IP the differences among the best models in each category are minor.

Overall, the evaluation of the point nowcasts highlights that for some variables and horizons the inclusion of the big data based uncertainty indicator can reduce the RMSFE, particularly so for HICP and UR, and for longer horizons. This is an interesting finding both in terms of the usefulness of big data and of the predictive power of uncertainty for macroeconomic variables.

5.4.3 SSR

Next, we turn our focus on the estimated direction rather than the actual point estimate. Tables 6 to 9 present the results for the sign success ratio (SSR). For this statistic we report the actual estimates (without reference to a benchmark). All results are expressed in percentages (%).

A first general comment on the results is that some changes in the specification, like including or not the uncertainty indicator, often do not change the sign forecasts, so that alternative specifications can give the same SSR. Also, the top models in each category usually have same similar SSRs.

Going in more details, and starting with German HICP, Table 6 shows that, as for point nowcasts, AR(AIC) gives the best results at short horizons, but the differences

with respect to sparse principal component models without or with the big data uncertainty indicator, are minor (the values are in the range 55%-58%). The latter models are among the best performers at longer nowcast horizons. As for HICP, AR(AIC) is the best performer at short horizons also for IP, with more complex models with or without the uncertainty indexes in the top-5 for these horizons, and best at longer horizons. For UR the top model is the random walk, followed by other simple specifications. It is worth mentioning that targets with relatively higher volatility, as the first difference of UR, are associated with smaller SSR.

Continuing with France, from Table 7 it emerges that across all target variables the top models are: AR(AIC), ETS and Avg(24). However, most methods perform similarly (and generally better than for Germany), and models with the uncertainty indicator often appear in the top-5 lists for UR, and sometimes in that for HICP.

For Italy there is more evidence of the usefulness of the Google uncertainty indicator, according to the figures in Table 8. Specifically, starting with HICP, LinReg-G-L3 is the top model across the nowcasting horizons, with SSR starting at 56.52% (-5w) and increasing to 78.22% (-1w). For IP, the BaggedETS is the top performing model followed by "Best" strategies and univariate models. For UR, and in line with the point nowcasts, the simple LinReg-G-L3 is best at -3w, -2w and -1w.

The results for the UK are more varied. According to Table 9, for HICP the LinReg-G-L3 and LASSO-MF-G are among the top models, with SSR of about 73.5% at -2w and -1w, much higher than that of simple and univariate time series models (about 61.7%). Continuing with IP, sparse principal component models including the big data indicator are the top performers, with an SSR of 76.5% for -2w, and of 70.6% for –3w and 1w. Finally, for UR, various models perform similarly.

Overall, the evaluation of the directional nowcasts confirms that for several variables and horizons the inclusion of the big data based uncertainty indicator can yield gains.

5.4.4 Nowcast Intervals Coverage Rates

We now evaluate the empirical coverage rate of 90% nowcast intervals obtained from the various specifications. In particular, we report the absolute value of the distance between the empirical coverage rate from the nominal 90%, that is $CS =$

$\left|\widehat{CR^{90\%}} - 90\%\right|$. Consequently, we are looking for models which have a CS as low as possible. Results are reported in Tables 10 to 13.

The general message is that for the top models within each group the CS are very close to zero and comparable, with the exception of the naive and simple averaging approaches, implying that the empirical coverage rates are close to the nominal ones for several specifications with or without the Google uncertainty index.

6 Recommendations for the Use of Big Data for Economic Nowcasting

We believe our analysis so far has highlighted the potential theoretical and empirical benefits associated with the use of big data. However, there are also costs that should be considered, for example, for data collection, storage and handling, and there are also potential issues in terms of data quality, confidentiality, and reliability of provision. Overall, our suggestion is to take a pragmatic approach that balances potential gains and costs from the use of big data for nowcasting economic indicators, in addition to standard indicators. The following step-wise procedure could be helpful in the assessment of the expected net gains from the use of big data in a nowcasting context.

A preliminary step should be an a priori assessment of the potential usefulness of big data for a specific indicator of interest, such as GDP growth, inflation or unemployment, for a specific country or region. This requires to evaluate the quality of the existing nowcasts and whether any identified problems, such as bias or inefficiency or large errors in specific periods, can be fixed by adding information as potentially available in Big Data based indicators. Similarly, it should be considered whether these additional indicators could improve the timeliness, frequency of release and extent of revision of the nowcasts. Relevant information can be gathered by looking at existing empirical studies focusing on similar variables or countries.

Once big data passes the "need check" in the preliminary step, the first proper step of the big data based nowcasting exercise is a careful search for the specific big data to be collected. As we have mentioned, there are many potential providers, which can be grouped into Social Networks, Traditional Business Systems, and the

Internet of Things. Naturally, it is not possible to give general guidelines on a preferred data source, as its choice is heavily dependent on the target indicator of the nowcasting exercise.

Having identified the preferred source of big data, the second step requires to assess the availability and quality of the data. A relevant issue is whether direct data collection is needed, which can be very costly, or a provider makes the data available. In case a provider is available, its reliability (and cost) should be assessed, together with the availability of meta data, the likelihood that continuity of data provision is guaranteed, and the possibility of customization (e.g., make the data available at higher frequency, with a particular disaggregation, for a longer sample, etc.). All these aspects are particularly relevant in the context of applications in official statistical offices. As the specific goal is nowcasting, it should be also carefully checked that the temporal dimension of the big data is long and homogeneous enough to allow for proper model estimation and evaluation of the resulting nowcasts.

The third step requires an analysis of specific features of the collected Big Data. A first issue that is sometimes neglected is the amount of the required storage space and the associated need of specific hardware and software for storing and handling the big data. A second issue is the type of the big data, as it is often unstructured and may require a transformation into cross-sectional or time series observations. Even when already available in numerical format, pre-treatment of the big data is often needed to remove deterministic patterns and deal with data irregularities, such as outliers and missing observations. While standard methods can be usually applied, the size of the datasets suggests to resort to robust and computationally simple approaches, applied variable by variable.

The fourth step requires to assess the presence of a possible bias in the answers provided by the big data, due to the "digital divide" or the tendency of individuals and businesses not to report truthfully their experiences, assessments and opinions. A related problem, particularly relevant for nowcasting, is the possible instability of the relationship with the target variable. This is a common problem also with standard indicators, as the type and size of economic shocks that hit the economy vary over time. Both issues can be however tackled at the modelling and evaluation stages.

The fifth step when nowcasting with big data requires to select the proper econo-

metric technique. Here, it is important to be systematic about the correspondence between the nature of the big data setting and use under investigation and the method that is used. There is a number of dimensions along which we wish to differentiate. The first choice is between the use of methods suited for large but not huge datasets, and therefore applied to summaries of the big data (such as Google Trends, commonly used in nowcasting applications), or of techniques specifically designed for big data. For example, nowcasting with large datasets can be based on factor models, large BVARs, or shrinkage regressions. Huge datasets can be handled by sparse principal components, linear models combined with heuristic optimization, or a variety of machine learning methods (which, though, are generally developed assuming i.i.d. variables). As we have seen, it is difficult to provide an a priori ranking of all these techniques and there are few empirical comparisons and even fewer in a nowcasting context, so that it may be appropriate to apply and compare a few of them for nowcasting the specific indicator of interest. A second dimension is the frequency of the available data. If this frequency is mixed then specific techniques for mixed frequency data become relevant. Chief among them is unrestricted MIDAS which provides a very flexible framework of analysis and can be adapted to work together with most if not all big data methods be they machine learning of econometric. Yet another dimension relates to the purpose for which large datasets are considered. Possibilities include model or indicator selection, forecasting or a more structural analysis. In this case of course each purpose is best served by different methods and the choice of method crucially depends on the purpose. Most methods can be used for forecasting and so the choice has to be case dependent. We recommend that as many methods as possible are evaluated in a forecasting context although past experience suggests that factor analysis and shrinkage methods can be of great use. For model or indicator selection penalised regression and the Multiple Testing methods seem to be appropriate and also have been reported to have good potential. Finally, for more structural analysis it is clear that it is likely that huge datasets are more difficult to accommodate. In this case, system methods that analyse the whole or a large proportion of the available data simultaneously, seem necessary for a satisfactory analytical outcome. Bayesian VAR models stand out as an appropriate method in this context.

The final step consists of a critical and comprehensive assessment of the contri-

bution of big data for nowcasting the indicator of interest. In order to avoid, or at least reduce the extent of, data and model snooping, a cross-validation approach should be followed, whereby various models and indicators are estimated over a first sample and they are selected and/or pooled according to their performance, but then the performance of the preferred approaches is re-evaluated over a second sample. This procedure, which we have implemented in the empirical evaluation, provides a reliable assessment of the gains in terms of enhanced nowcasting performance and timeliness from the use of big data.

7 Conclusions

In this paper we have provided an assessment of the usefulness of big data for nowcasting and flash estimation, both from a theoretical and an empirical point of view. A companion paper, Buono et al. (2017), presents an extensive literature review on the use of big data in economics, with a special focus on nowcasting and forecasting

After discussing a typology of big data and issues related with their use, we have proposed various ways to move from huge unstructured big data to a limited number of summary time series, which can then be used in combination with more standard economic indicators. As an illustration, we have constructed uncertainty indexes based on Google trends associated with uncertainty related queries.

Next, we have briefly surveyed a variety of econometric methods suited for large (though not huge) datasets, and implemented them in an empirical exercise on point, interval and directional nowcasting key economic variables for the four largest European countries, using a large set of macroeconomic and financial indicators combined with the big data based uncertainty indexes.

Finally, the empirical exercise has shown that for several variables and horizons the inclusion of the big data based uncertainty indicator can reduce the RMSFE for point nowcasts, increase the SSR for directional nowcasts, and improve the coverage rate for interval nowcasts. The gains are particularly evident for HICP and UR, and for longer horizons. This is an interesting finding both in terms of the usefulness of big data and of the predictive power of uncertainty.

8 References

1. Baker, S.R., Bloom, N., Davis, S.J. (2016). "Measuring Economic Policy Uncertainty". The Quarterly Journal of Economics, 131(4), 1593-1636.

2. Bates, J.M. and Granger, C.W.J. (1969). "The Combination of Forecasts", Operational Research, 20(4), 451-468.

3. Bergmeir, C., Hyndman, R.J., Benitez, J.M. (2016). "Bagging Exponential Smoothing Methods using STL Decomposition and Box-Cox Transformation." International Journal of Forecasting, 32, 303-312.

4. Bok, B., Caratelli, D., Giannone, D., Sbordone, A., Tambalotti, A. (2017). "Macroeconomic Nowcasting and Forecasting with Big Data", Federal Reserve Bank of New York, Staff Report No. 830.

5. Buono, D., Kapetanios, G., Marcellino, M., Mazzi, G.L., Papailias, F. (2017). "The Challenge of Big Data", Handbook on Rapid Estimates, Luxembourg: Publications Office of the European Union, Chapter 16, 449-462.

6. Carriero, A., Clark, T.E., Marcellino, M. (2018). "Measuring Uncertainty and Its Impact on the Economy". Review of Economics & Statistics, Forthcoming. `https://doi.org/10.1162/REST_a_00693`.

7. Diebold, F.X., Mariano, R.S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economics Statistics*, 13(3), 253-263.

8. Doornik, J. A., Hendry, D. F. (2015). "Statistical Model Selection with big data". Cogent Economics & Finance, 3(1), 2015.

9. Eckley, P. (2015). "Measuring economic uncertainty using news-media textual data". MPRA Paper No. 69784.

10. Foroni, C., Marcellino, M., Schumacher, C. (2015). "Unrestricted Mixed Data Sampling (MIDAS): MIDAS Regressions with Unrestricted Lag Polynomials". Journal of the Royal Statistical Society: Series A, 178(1), 57-82.

11. Ghysels, E., Santa-Clara, P., Valkanov, R. (2004). "The MIDAS Touch: Mixed Data Sampling Regression Models", CIRANO Working Paper, 2004s-20.

12. Giannone, D., Reichlin, L., Small, D. (2008). "Nowcasting: The Real-Time Informational Content of Macroeconomic Data", Journal of Monetary Economics, 55, 665-676.

13. Hartford, T. (2014). "Big data: Are we making a big mistake?". Financial Times, March, 28, 2014. Available at: `https://goo.gl/SMOL2L`.

14. Hepenstrich, C. and Marcellino, M. (2018). "Forecasting with Large Unbalanced Datasets: The Mixed Frequency Three-Pass Regression Filter", Journal of the Royal Statistical Society, Series A, forthcoming.

15. Jurado,K., Ludvigson, S.C., Ng, S. (2015). "Measuring Uncertainty." American Economic Review, 105(3), 1177–1216.

16. Kapetanios, G., Marcellino, M., Papailias, F. (2017a). "Big Data Conversion Techniques including their Main Features and Characteristics". Eurostat Statistical Working Papers, 2017. Available at: `https://goo.gl/X4P71T`.

17. Kapetanios, G., Marcellino, M., Papailias, F. (2017b). "Filtering techniques for big data and big data based uncertainty indexes". Eurostat Statistical Working Papers, 2017. Available at: `https://goo.gl/UoAeVk`.

18. Kapetanios, G., Marcellino, M., Papailias, F. (2017c). "Guidance and Recommendations on the Use of Big data for Macroeconomic Nowcasting", Handbook on Rapid Estimates, Luxembourg: Publications Office of the European Union, Chapter 17, 463-479.

19. Lloyd, S. P. (1982). "Least squares quantization in PCM". Information Theory, IEEE Transactions, 28(2), 129-137.

20. McCracken, M.W, Ng, S. (2015) FRED-MD: A Monthly Database for Macroeconomic Research, *Research Division, Federal Reserve Bank of St. Louis, Working Paper Series*, Working Paper 2015-012B.

21. Ng, S. (2016). "Opportunities and Challenges: Lessons from Analyzing Terabytes of Scanner Data". Working Paper.

22. Roussas, G. (1996). "Exponential Probability Inequalities with Some Applications". Statistics, Probability & Game Theory, IMS Lecture Notes – Monograph Series, 30.

23. Scott, S., Varian, H. (2013). Predicting the Present with Bayesian Structural Time Series, *Working Paper.*

24. Stock, J., Watson, M. (2002a). "Forecasting Using Principal Components from a Large Number of Predictors", Journal of the American Statistical Association, 297, 1167-1179.

25. Stock, J., Watson, M. (2002b). "Macroeconomic Forecasting using Diffusion Indexes", Journal of Business & Economics Statistics, 20, 147-162.

9 Tables

Variable Category/Type	Frequency	Transformation	Variable Category/Type	Frequency	Transformation
Bank Lending Rate	M	2	Imports	M	3
Bankruptcies	M	5	Initial Jobless Claims	M	5
Banks Balance Sheet	M	6	Three Month Interbank Rate	M	2
Business Confidence	M	2	Job Vacancies	M	2
Capital Flows	M	1	Loan Growth YoY	M	2
Car Registrations	M	3	Loans to Private Sector	M	5
Central Bank Balance Sheet	M	6	Manufacturing Production	M	1
Construction Output	M	2	Mining Production	M	1
Consumer Confidence	M	2	Money Supply M1	M	6
Consumer Credit	M	3	Money Supply M2	M	6
Core Consumer Prices	M	6	Money Supply M3	M	6
CPI Housing & Utilities	M	6	Non Farm Payrolls	W	5
CPI Transportation	M	6	Household Consumption MoM	M	1
Crude Oil Production	M	2	Producer Prices	M	6
Exchange Rate	W	3	Retail Sales MoM	M	1
Current Account	M	3	Steel Production	M	3
Export Prices	M	6	Youth Unemployment Rate	M	5
Exports	M	3	Zew Economic Sentiment Index	M	2
Food Inflation	M	2	Various Eurostat Surveys	M	2
Foreign Direct Investment	M	1	Interest Rates (Yield Curve Components)	W	2
Foreign Exchange Reserves	M	6	Various Stock Market & Volatility Indexes	W	3
Gasoline Prices	M	3	Google Trends Big Data Based Indicator	W	1
Housing Starts	M	4			
Import Prices	M	6			

Notes: The above variables are selected appropriate for each country. M: Monthly, W: Weekly. 1: no change, 2: first difference, 3: percentage change, 4: log, 5: log difference, 6: second log difference.

Table 1: Generic Variable & Transformation Table.

Country: Germany (DE)										
	Best Model					RMSFE				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Avg(12)	Avg(12)	Avg(12)	Avg(24)	Avg(24)	1.113	1.113	1.004	1.015	1.015
TS	AR(4)	AR(4)	AR(AIC)	AR(AIC)	AR(AIC)	0.992	0.992	0.785	0.770	0.770
G	LASSO-MF-G	PLS(2)-MF-G	EN-MF-G	PLS(5)-MF-G	EN-MF-G	0.966	0.966	0.813	0.820	0.819
MF	LASSO-MF	PLS(2)-MF	EN-MF	EN-MF	EN-MF	0.955	0.960	0.826	0.826	0.822
Best	Best10	Best10	Best10	Best10	Best10	1.037	1.035	1.000	0.996	0.998
Top1	LASSO-MF	PLS(2)-MF	AR(AIC)	AR(AIC)	AR(AIC)	0.955	0.960	0.785	0.770	0.770
Top2	LASSO-MF-G	PLS(2)-MF-G	EN-MF-G	PLS(5)-MF-G	EN-MF-G	0.966	0.966	0.813	0.820	0.819
Top3	EN-MF	PLS(3)-MF	SPC(5)-MF-G	SPC(5)-MF-G	PLS(5)-MF-G	0.971	0.979	0.818	0.822	0.820
Top4	EN-MF-G	PLS(3)-MF-G	EN-MF	EN-MF-G	EN-MF	0.982	0.983	0.826	0.825	0.822
Top5	AR(4)	LASSO-MF	PLS(5)-MF-G	EN-MF	SPC(5)-MF	0.992	0.987	0.826	0.826	0.823
Target: Industrial Production (IP)										
AveNaive	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	0.971	0.971	0.974	0.973	0.973
TS	AutoArima	AutoArima	BaggedETS	BaggedETS	BaggedETS	0.934	0.934	0.880	0.889	0.907
G	Spike-MF-G	Spike-MF-G	DFA2-MF-G	DFA2-MF-G	EN-MF-G	0.955	0.943	0.951	0.959	0.954
MF	Spike-MF	Spike-MF	DFA2-MF	DFA2-MF	DFA2-MF	0.956	0.940	0.947	0.956	0.955
Best	Best10	Best10	Best1	Best10	Best10	1.030	1.029	0.957	0.997	0.996
Top1	AutoArima	AutoArima	BaggedETS	BaggedETS	BaggedETS	0.934	0.934	0.880	0.889	0.907
Top2	AR(4)	AR(4)	NN	NN	NN	0.940	0.940	0.915	0.916	0.918
Top3	THETA	Spike-MF	AutoArima	AutoArima	AutoArima	0.944	0.940	0.921	0.932	0.932
Top4	Spike-MF-G	Spike-MF-G	THETA	THETA	THETA	0.955	0.943	0.945	0.945	0.945
Top5	Spike-MF	THETA	AR(4)	AR(4)	AR(4)	0.956	0.944	0.946	0.945	0.945
Target: Unemployment Rate (UR)										
AveNaive	Avg(12)	Avg(12)	Avg(24)	Avg(24)	Avg(24)	0.940	0.940	0.816	0.790	0.790
TS	AR(AIC)	AR(AIC)	THETA	BaggedETS	BaggedETS	0.932	0.932	0.842	0.791	0.799
G	SPC(1)-MF-G	SPC(1)-MF-G	LinReg-G	LinReg-G	LinReg-G	0.841	0.848	0.799	0.780	0.780
MF	SPC(1)-MF	SPC(1)-MF	SPC(1)-MF	SPC(1)-MF	SPC(1)-MF	0.841	0.846	0.813	0.794	0.791
Best	Best1	Best1	Best5	Best10	Best10	0.894	0.891	0.839	0.816	0.820
Top1	SPC(1)-MF	SPC(1)-MF	LinReg-G	LinReg-G	LinReg-G	0.841	0.846	0.799	0.780	0.780
Top2	SPC(1)-MF-G	SPC(1)-MF-G	SPC(1)-MF-G	Avg(24)	Avg(24)	0.841	0.848	0.812	0.790	0.790
Top3	PLS(2)-MF	PLS(5)-MF	SPC(1)-MF	BaggedETS	SPC(1)-MF-G	0.868	0.882	0.813	0.791	0.791
Top4	PLS(2)-MF-G	PLS(5)-MF-G	Avg(24)	SPC(1)-MF-G	SPC(1)-MF	0.871	0.886	0.816	0.793	0.791
Top5	PLS(5)-MF	PLS(2)-MF	EN-MF-G	SPC(1)-MF	BaggedETS	0.871	0.888	0.828	0.794	0.799

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 2: Germany, RMSFE relative to AR(1).

Country: France (FR)										
	Best Model					RMSFE				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	1.017	1.017	0.986	0.993	0.993
TS	AR(AIC)	AR(AIC)	AR(AIC)	AR(AIC)	AR(AIC)	0.917	0.916	0.603	0.577	0.577
G	LASSO-MF-G	LASSO-MF-G	LASSO-MF-G	LASSO-MF-G	EN-MF-G	0.910	0.852	0.868	0.872	0.873
MF	EN-MF	LASSO-MF	LASSO-MF	LASSO-MF	LASSO-MF	0.908	0.854	0.858	0.865	0.856
Best	Best10	Best5	Best1	Best1	Best1	0.949	0.891	0.709	0.684	0.684
Top1	EN-MF	LASSO-MF-G	AR(AIC)	AR(AIC)	AR(AIC)	0.908	0.852	0.603	0.577	0.577
Top2	LASSO-MF-G	LASSO-MF	Best1	Best1	Best1	0.910	0.854	0.709	0.684	0.684
Top3	EN-MF-G	EN-MF-G	Best3	Best3	Best3	0.913	0.854	0.788	0.806	0.807
Top4	AR(AIC)	EN-MF	Best5	Best5	Best5	0.917	0.863	0.839	0.851	0.850
Top5	LASSO-MF	Best5	LASSO-MF	LASSO-MF	LASSO-MF	0.920	0.891	0.858	0.865	0.856
Target: Industrial Production (IP)										
AveNaive	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	1.204	1.204	1.106	1.125	1.125
TS	BaggedETS	BaggedETS	AR(AIC)	BaggedETS	BaggedETS	0.979	0.996	0.967	0.951	0.960
G	PLS(3)-MF-G	PLS(4)-MF-G	DFA2-MF-G	DFA2-MF-G	DFA2-MF-G	1.064	1.084	1.067	1.085	1.085
MF	PLS(4)-MF	PLS(4)-MF	DFA2-MF	DFA2-MF	DFA2-MF	1.071	1.087	1.068	1.085	1.085
Best	Best1	Best10	Best5	Best3	Best1	1.049	1.065	0.997	0.989	0.991
Top1	BaggedETS	BaggedETS	AR(AIC)	BaggedETS	BaggedETS	0.979	0.996	0.967	0.951	0.960
Top2	AR(1)	AR(1)	ETS	ETS	ETS	1.000	1.000	0.970	0.969	0.969
Top3	AR(AIC)	AR(AIC)	BATS	AR(AIC)	AR(AIC)	1.003	1.003	0.971	0.970	0.970
Top4	ETS	ETS	TBATS	BATS	BATS	1.007	1.007	0.971	0.971	0.971
Top5	BATS	BATS	BaggedETS	TBATS	TBATS	1.007	1.007	0.972	0.971	0.971
Target: Unemployment Rate (UR)										
AveNaive	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	0.956	0.956	0.901	0.868	0.868
TS	NN	NN	NN	AR(AIC)	NN	0.939	0.949	0.898	0.870	0.865
G	Spike-MF-G	Spike-MF-G	Spike-MF-G	EN-MF-G	EN-MF-G	0.950	0.945	0.890	0.903	0.897
MF	Spike-MF	Spike-MF	Spike-MF	EN-MF	LASSO-MF	0.950	0.946	0.893	0.890	0.893
Best	Best10	Best10	Best10	Best10	Best10	0.989	0.982	0.938	0.951	0.946
Top1	NN	Spike-MF-G	Spike-MF-G	Avg(24)	NN	0.939	0.945	0.890	0.868	0.865
Top2	Spike-MF	Spike-MF	Spike-MF	AR(AIC)	Avg(24)	0.950	0.946	0.893	0.870	0.868
Top3	Spike-MF-G	NN	NN	THETA	AR(AIC)	0.950	0.949	0.898	0.874	0.870
Top4	Avg(24)	Avg(24)	LASSO-MF	NN	THETA	0.956	0.956	0.899	0.875	0.874
Top5	PLS(3)-MF	LASSO-MF	Avg(24)	Avg(12)	Avg(12)	0.962	0.959	0.901	0.882	0.882

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 3: France, RMSFE relative to AR(1).

Country: Italy (IT)										
	Best Model					RMSFE				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Avg(12)	Avg(12)	Avg(24)	Avg(24)	Avg(24)	0.976	0.976	0.989	0.991	0.991
TS	AR(AIC)	AR(AIC)	AR(AIC)	NN	NN	0.681	0.681	0.421	0.250	0.262
G	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	0.727	0.722	0.443	0.401	0.401
MF	LASSO-MF	EN-MF	LASSO-MF	LASSO-MF	EN-MF	0.928	0.899	0.898	0.915	0.915
Best	Best3	Best3	Best3	Best3	Best3	0.724	0.722	0.541	0.463	0.463
Top1	AR(AIC)	AR(AIC)	AR(AIC)	NN	NN	0.681	0.681	0.421	0.250	0.262
Top2	NN	NN	NN	AR(AIC)	AR(AIC)	0.699	0.687	0.436	0.292	0.292
Top3	Best3	Best3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	0.724	0.722	0.443	0.401	0.401
Top4	AR(4)	LinReg-G-L3	AR(4)	Best3	Best3	0.726	0.722	0.529	0.463	0.463
Top5	LinReg-G-L3	AR(4)	Best3	AR(4)	AR(4)	0.727	0.726	0.541	0.466	0.466
Target: Industrial Production (IP)										
AveNaive	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	1.000	1.000	0.995	1.008	1.008
TS	BaggedETS	BaggedETS	BaggedETS	BaggedETS	BaggedETS	0.812	0.819	0.864	0.883	0.880
G	LASSO-MF-G	LASSO-MF-G	LASSO-MF-G	LASSO-MF-G	LASSO-MF-G	0.970	0.955	0.906	0.927	0.933
MF	LASSO-MF	LASSO-MF	LASSO-MF	EN-MF	LASSO-MF	0.974	0.966	0.889	0.931	0.917
Best	Best1	Best1	Best10	Best5	Best10	0.846	0.853	0.917	0.940	0.944
Top1	BaggedETS	BaggedETS	BaggedETS	BaggedETS	BaggedETS	0.812	0.819	0.864	0.883	0.880
Top2	Best1	ETS	LASSO-MF	ETS	ETS	0.846	0.851	0.889	0.900	0.900
Top3	ETS	Best1	ETS	BATS	BATS	0.851	0.853	0.896	0.904	0.904
Top4	BATS	BATS	BATS	TBATS	TBATS	0.855	0.855	0.900	0.904	0.904
Top5	TBATS	TBATS	TBATS	LASSO-MF-G	LASSO-MF	0.855	0.855	0.900	0.927	0.917
Target: Unemployment Rate (UR)										
AveNaive	Avg(12)	Avg(12)	Avg(24)	Avg(24)	Avg(24)	0.979	0.979	0.999	0.981	0.981
TS	NN	NN	AR(4)	AR(4)	AR(4)	0.906	0.906	0.918	0.931	0.931
G	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	0.933	0.915	0.898	0.925	0.925
MF	SPC(4)-MF	LASSO-MF	PLS(4)-MF	PLS(2)-MF	PLS(2)-MF	0.998	0.949	1.000	1.001	1.001
Best	Best3	Best5	Best3	Best3	Best10	0.991	0.979	1.020	1.037	1.038
Top1	NN	NN	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	0.906	0.906	0.898	0.925	0.925
Top2	ETS	LinReg-G-L3	AR(4)	AR(4)	AR(4)	0.925	0.915	0.918	0.931	0.931
Top3	BaggedETS	BaggedETS	ETS	ETS	ETS	0.930	0.923	0.939	0.932	0.932
Top4	LinReg-G-L3	ETS	NN	NN	BaggedETS	0.933	0.925	0.940	0.943	0.940
Top5	THETA	LASSO-MF	BaggedETS	BaggedETS	NN	0.957	0.949	0.954	0.945	0.943

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 4: Italy, RMSFE relative to AR(1).

Country: United Kingdom (UK)										
	Best Model					RMSFE				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Avg(12)	Avg(12)	Avg(12)	Avg(12)	Avg(12)	0.958	0.959	0.881	0.873	0.873
TS	AR(AIC)	AR(AIC)	AR(AIC)	AR(AIC)	AR(AIC)	0.935	0.933	0.723	0.718	0.718
G	EN-MF-G	EN-MF-G	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	0.793	0.800	0.807	0.794	0.796
MF	Spike-MF	EN-MF	EN-MF	Spike-MF	LASSO-MF	0.798	0.800	0.856	0.830	0.825
Best	Best10	Best10	Best3	Best3	Best3	0.811	0.818	0.820	0.843	0.841
Top1	EN-MF-G	EN-MF-G	AR(AIC)	AR(AIC)	AR(AIC)	0.793	0.800	0.723	0.718	0.718
Top2	Spike-MF-G	EN-MF	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	0.795	0.800	0.807	0.794	0.796
Top3	LASSO-MF-G	Spike-MF	Best3	LASSO-MF-G	EN-MF-G	0.797	0.809	0.820	0.801	0.807
Top4	Spike-MF	Spike-MF-G	Best5	EN-MF-G	LASSO-MF-G	0.798	0.811	0.839	0.808	0.819
Top5	EN-MF	LASSO-MF-G	LASSO-MF-G	Spike-MF-G	LASSO-MF	0.801	0.811	0.839	0.830	0.825
Target: Industrial Production (IP)										
AveNaive	Avg(12)	Avg(12)	Avg(12)	Avg(12)	Avg(12)	0.905	0.904	0.968	0.991	0.991
TS	ETS	ETS	AR(4)	ETS	ETS	0.902	0.902	0.940	0.934	0.934
G	LASSO-MF-G	DFA5-MF-G	SPC(5)-MF-G	SPC(5)-MF-G	PLS(2)-MF-G	0.960	0.911	0.955	0.941	0.961
MF	LASSO-MF	DFA5-MF	PLS(2)-MF	SPC(5)-MF	SPC(5)-MF	0.948	0.912	0.969	0.956	0.942
Best	Best5	Best10	Best10	Best10	Best10	1.024	1.025	0.958	0.954	0.957
Top1	ETS	ETS	AR(4)	ETS	ETS	0.902	0.902	0.940	0.934	0.934
Top2	BATS	BATS	ETS	BATS	BATS	0.903	0.903	0.941	0.935	0.935
Top3	TBATS	TBATS	BATS	TBATS	TBATS	0.903	0.903	0.942	0.935	0.935
Top4	Avg(12)	Avg(12)	TBATS	AutoArima	AutoArima	0.905	0.904	0.942	0.936	0.936
Top5	AutoArima	AutoArima	AutoArima	SPC(5)-MF-G	BaggedETS	0.905	0.905	0.945	0.941	0.936
Target: Unemployment Rate (UR)										
AveNaive	Avg(4)	Avg(4)	Avg(4)	Avg(4)	Avg(4)	0.825	0.823	0.863	0.888	0.888
TS	AutoArima	AutoArima	AutoArima	AutoArima	AutoArima	0.847	0.845	0.891	0.904	0.904
G	PLS(4)-MF-G	PLS(4)-MF-G	Spike-MF-G	PLS(3)-MF-G	PLS(3)-MF-G	0.874	0.890	0.941	0.893	0.896
MF	PLS(4)-MF	PLS(4)-MF	Spike-MF	PLS(3)-MF	PLS(3)-MF	0.884	0.898	0.941	0.893	0.896
Best	Best1	Best1	Best3	Best10	Best10	0.882	0.881	0.874	0.887	0.889
Top1	Avg(4)	Avg(4)	Avg(4)	Best10	Avg(4)	0.825	0.823	0.863	0.887	0.888
Top2	AutoArima	AutoArima	Best3	Avg(4)	Best10	0.847	0.845	0.874	0.888	0.889
Top3	PLS(4)-MF-G	Best1	Best1	Best5	PLS(3)-MF	0.874	0.881	0.883	0.892	0.896
Top4	PLS(1)-MF-G	Spline	Best5	PLS(3)-MF	PLS(3)-MF-G	0.877	0.881	0.884	0.893	0.896
Top5	Best1	Avg(12)	AutoArima	PLS(3)-MF-G	PLS(5)-MF	0.882	0.882	0.891	0.893	0.896

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 5: UK, RMSFE relative to AR(1).

Country: Germany (DE)										
	Best Model					SSR				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Avg(4)	Avg(4)	Avg(24)	Avg(24)	Avg(24)	50.000	50.000	38.240	38.240	38.240
TS	NN	NN	AR(AIC)	AR(AIC)	AR(AIC)	47.060	47.060	55.880	58.820	58.820
G	PLS(3)-MF-G	PLS(3)-MF-G	SPC(5)-MF-G	SPC(5)-MF-G	SPC(5)-MF-G	52.940	50.000	55.880	55.880	55.880
MF	PLS(3)-MF	PLS(3)-MF	SPC(5)-MF	SPC(5)-MF	SPC(5)-MF	50.000	47.060	55.880	55.880	55.880
Best	Best5	Best1	Best10	Best10	Best1	44.120	41.180	52.940	52.940	50.000
Top1	PLS(3)-MF-G	Avg(4)	AR(AIC)	AR(AIC)	AR(AIC)	52.940	50.000	55.880	58.820	58.820
Top2	Avg(4)	PLS(3)-MF-G	SPC(5)-MF	SPC(5)-MF	SPC(5)-MF	50.000	50.000	55.880	55.880	55.880
Top3	PLS(3)-MF	NN	SPC(5)-MF-G	SPC(5)-MF-G	SPC(5)-MF-G	50.000	47.060	55.880	55.880	55.880
Top4	PLS(4)-MF	PLS(3)-MF	EN-MF-G	LASSO-MF	LASSO-MF-G	50.000	47.060	55.880	55.880	55.880
Top5	NN	PLS(5)-MF	AR(4)	AR(4)	AR(4)	47.060	47.060	52.940	52.940	52.940
Target: Industrial Production (IP)										
AveNaive	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	55.880	55.880	35.290	32.350	32.350
TS	AR(4)	AR(4)	AR(AIC)	AR(AIC)	AR(AIC)	61.760	61.760	50.000	50.000	50.000
G	PLS(4)-MF-G	PLS(4)-MF-G	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	67.650	64.710	52.940	50.000	50.000
MF	PLS(4)-MF	PLS(4)-MF	DFA5-MF	DFA5-MF	DFA5-MF	67.650	64.710	52.940	50.000	50.000
Best	Best3	Best10	Best1	Best1	Best5	55.880	64.710	47.060	47.060	44.120
Top1	PLS(4)-MF	PLS(4)-MF	LinReg-G-L3	AR(AIC)	AR(AIC)	67.650	64.710	52.940	50.000	50.000
Top2	PLS(4)-MF-G	PLS(4)-MF-G	DFA5-MF	LinReg-G-L3	LinReg-G-L3	67.650	64.710	52.940	50.000	50.000
Top3	PLS(5)-MF	PLS(5)-MF	DFA5-MF-G	DFA5-MF	DFA5-MF	67.650	64.710	52.940	50.000	50.000
Top4	PLS(5)-MF-G	PLS(5)-MF-G	SPC(5)-MF-G	DFA5-MF-G	DFA5-MF-G	67.650	64.710	52.940	50.000	50.000
Top5	AR(4)	Best10	AR(AIC)	SPC(5)-MF-G	EN-MF-G	61.760	64.710	50.000	50.000	50.000
Target: Unemployment Rate (UR)										
AveNaive	Naive	Naive	Naive	Naive	Naive	38.240	38.240	41.180	38.240	38.240
TS	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	32.350	32.350	32.350	32.350	32.350
G	LinReg-G	LinReg-G	LinReg-G	LinReg-G	LinReg-G	32.350	32.350	32.350	32.350	32.350
MF	PLS(1)-MF	DFA2-MF	DFA2-MF	DFA2-MF	DFA2-MF	32.350	32.350	32.350	32.350	32.350
Best	Best1	Best1	Best1	Best1	Best1	38.240	38.240	38.240	38.240	38.240
Top1	Naive	Naive	Naive	Naive	Naive	38.240	38.240	41.180	38.240	38.240
Top2	Best1	Best1	Best1	Best1	Best1	38.240	38.240	38.240	38.240	38.240
Top3	Best3	Best3	Avg(4)	Avg(4)	Avg(4)	35.290	35.290	35.290	35.290	35.290
Top4	Best5	Best5	Best10	Best10	Best10	35.290	35.290	35.290	35.290	35.290
Top5	Best10	Best10	Avg(12)	Avg(12)	Avg(12)	35.290	35.290	32.350	32.350	32.350

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 6: Germany, SSR.

Country: France (FR)										
	Best Model					SSR				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Avg(12)	Avg(12)	Avg(24)	Avg(24)	Avg(24)	64.710	64.710	52.940	50.000	50.000
TS	BATS	BATS	AR(AIC)	AR(AIC)	AR(AIC)	67.650	67.650	79.410	82.350	82.350
G	PLS(1)-MF-G	EN-MF-G	LASSO-MF-G	LASSO-MF-G	LASSO-MF-G	70.590	76.470	73.530	67.650	67.650
MF	PLS(1)-MF	EN-MF	LASSO-MF	EN-MF	LASSO-MF	70.590	73.530	73.530	70.590	73.530
Best	Best1	Best5	Best5	Best1	Best1	64.710	73.530	79.410	79.410	79.410
Top1	PLS(1)-MF	EN-MF-G	AR(AIC)	AR(AIC)	AR(AIC)	70.590	76.470	79.410	82.350	82.350
Top2	PLS(1)-MF-G	LASSO-MF-G	Best5	Best1	Best1	70.590	73.530	79.410	79.410	79.410
Top3	LASSO-MF-G	EN-MF	Best1	Best3	Best3	70.590	73.530	76.470	79.410	79.410
Top4	EN-MF-G	Best5	Best3	Best5	LASSO-MF	70.590	73.530	76.470	73.530	73.530
Top5	BATS	LASSO-MF	LASSO-MF	EN-MF	Best5	67.650	70.590	73.530	70.590	73.530
Target: Industrial Production (IP)										
AveNaive	Avg(24)	Avg(24)	Avg(12)	Avg(12)	Avg(12)	70.590	70.590	50.000	44.120	44.120
TS	AR(AIC)	AR(AIC)	ETS	ETS	ETS	79.410	79.410	73.530	73.530	73.530
G	SPC(3)-MF-G	PLS(2)-MF-G	PLS(2)-MF-G	EN-MF-G	PLS(2)-MF-G	76.470	76.470	61.760	61.760	58.820
MF	SPC(3)-MF	PLS(2)-MF	EN-MF	LASSO-MF	EN-MF	76.470	76.470	64.710	58.820	58.820
Best	Best1	Best3	Best1	Best10	Best10	70.590	73.530	67.650	67.650	70.590
Top1	AR(AIC)	AR(AIC)	ETS	ETS	ETS	79.410	79.410	73.530	73.530	73.530
Top2	AR(1)	AR(1)	BaggedETS	BATS	BaggedETS	76.470	76.470	73.530	73.530	73.530
Top3	ETS	ETS	BATS	TBATS	BATS	76.470	76.470	73.530	73.530	73.530
Top4	BATS	BATS	TBATS	BaggedETS	TBATS	76.470	76.470	73.530	70.590	73.530
Top5	TBATS	TBATS	THETA	THETA	THETA	76.470	76.470	70.590	70.590	70.590
Target: Unemployment Rate (UR)										
AveNaive	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	32.350	32.350	38.240	35.290	35.290
TS	AutoArima	AutoArima	AR(AIC)	AR(AIC)	AR(AIC)	32.350	32.350	35.290	35.290	35.290
G	PLS(1)-MF-G	PLS(5)-MF-G	Spike-MF-G	LinReg-G-L1	LinReg-G-L1	32.350	32.350	35.290	29.410	29.410
MF	PLS(1)-MF	DFA3-MF	LASSO-MF	Spike-MF	SPC(2)-MF	32.350	29.410	38.240	29.410	29.410
Best	Best3	Best3	Best3	Best3	Best3	32.350	29.410	35.290	29.410	29.410
Top1	Avg(24)	Avg(24)	Avg(24)	Avg(24)	Avg(24)	32.350	32.350	38.240	35.290	35.290
Top2	AutoArima	AutoArima	Naive	AR(AIC)	AR(AIC)	32.350	32.350	38.240	35.290	35.290
Top3	PLS(1)-MF	PLS(5)-MF-G	LASSO-MF	Naive	Naive	32.350	32.350	38.240	29.410	29.410
Top4	PLS(1)-MF-G	AR(AIC)	AR(AIC)	LinReg-G-L1	LinReg-G-L1	32.350	29.410	35.290	29.410	29.410
Top5	PLS(2)-MF	ETS	Spike-MF-G	LinReg-G-L3	LinReg-G-L3	32.350	29.410	35.290	29.410	29.410

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 7: France, SSR.

Country: Italy (IT)										
	Best Model					SSR				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Avg(12)	Avg(12)	Naive	Naive	Naive	30.430	30.430	47.830	52.170	52.170
TS	AR(AIC)	AR(AIC)	AR(4)	AR(4)	AR(4)	52.170	52.170	69.570	73.910	73.910
G	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	56.520	56.520	73.910	78.260	78.260
MF	LASSO-MF	DFA5-MF	PLS(1)-MF	PLS(1)-MF	PLS(1)-MF	39.130	43.480	56.520	56.520	56.520
Best	Best5	Best5	Best5	Best3	Best3	39.130	52.170	69.570	73.910	78.260
Top1	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	56.520	56.520	73.910	78.260	78.260
Top2	AR(AIC)	AR(AIC)	AR(4)	AR(4)	Best3	52.170	52.170	69.570	73.910	78.260
Top3	BATS	Best5	AR(AIC)	AR(AIC)	AR(4)	47.830	52.170	69.570	73.910	73.910
Top4	TBATS	BATS	Best5	NN	AR(AIC)	47.830	47.830	69.570	73.910	73.910
Top5	AR(4)	TBATS	Best10	Best3	NN	43.480	47.830	69.570	73.910	73.910
Target: Industrial Production (IP)										
AveNaive	Avg(4)	Avg(4)	Avg(24)	Avg(24)	Avg(24)	69.570	69.570	39.130	34.780	34.780
TS	AR(AIC)	AR(AIC)	BaggedETS	BaggedETS	BaggedETS	73.910	73.910	73.910	69.570	69.570
G	DFA2-MF-G	DFA2-MF-G	LASSO-MF-G	DFA2-MF-G	SPC(1)-MF-G	73.910	69.570	60.870	60.870	60.870
MF	DFA2-MF	PLS(5)-MF	LASSO-MF	DFA2-MF	SPC(1)-MF	73.910	73.910	65.220	60.870	60.870
Best	Best3	Best10	Best5	Best5	Best5	65.220	69.570	69.570	69.570	69.570
Top1	AR(AIC)	AR(AIC)	BaggedETS	BaggedETS	BaggedETS	73.910	73.910	73.910	69.570	69.570
Top2	BaggedETS	BaggedETS	Best5	Best5	Best5	73.910	73.910	69.570	69.570	69.570
Top3	DFA2-MF	PLS(5)-MF	Best10	AR(4)	Best10	73.910	73.910	69.570	65.220	69.570
Top4	DFA2-MF-G	Avg(4)	AR(4)	AR(AIC)	AR(4)	73.910	69.570	65.220	65.220	65.220
Top5	DFA3-MF-G	Avg(24)	AR(AIC)	ETS	AR(AIC)	73.910	69.570	65.220	65.220	65.220
Target: Unemployment Rate (UR)										
AveNaive	Avg(4)	Avg(4)	Avg(24)	Avg(24)	Avg(24)	56.520	56.520	43.480	43.480	43.480
TS	AR(1)	AR(1)	AR(4)	AR(4)	AR(4)	73.910	73.910	65.220	56.520	56.520
G	LinReg-G	LinReg-G	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	69.570	69.570	69.570	60.870	60.870
MF	PLS(1)-MF	DFA3-MF	DFA2-MF	DFA2-MF	DFA2-MF	69.570	69.570	43.480	39.130	39.130
Best	Best1	Best3	Best1	Best1	Best1	65.220	65.220	60.870	52.170	52.170
Top1	AR(1)	AR(1)	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	73.910	73.910	69.570	60.870	60.870
Top2	AR(4)	AR(4)	AR(4)	AR(4)	AR(4)	69.570	69.570	65.220	56.520	56.520
Top3	AutoArima	AutoArima	AutoArima	AutoArima	AutoArima	69.570	69.570	60.870	52.170	52.170
Top4	NN	NN	Best1	Best1	BaggedETS	69.570	69.570	60.870	52.170	52.170
Top5	THETA	THETA	Best5	Best10	Best1	69.570	69.570	60.870	52.170	52.170

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 8: Italy, SSR.

Country: United Kingdom (UK)										
	Best Model					SSR				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Naive	Naive	Avg(24)	Avg(24)	Avg(24)	61.760	61.760	55.880	61.760	61.760
TS	AR(1)	AR(1)	AR(1)	AR(1)	AR(1)	58.820	58.820	55.880	61.760	61.760
G	SPC(2)-MF-G	PLS(4)-MF-G	LinReg-G-L1	LinReg-G-L3	LinReg-G-L3	61.760	58.820	70.590	73.530	73.530
MF	SPC(2)-MF	SPC(2)-MF	Spike-MF	Spike-MF	Spike-MF	61.760	58.820	67.650	70.590	67.650
Best	Best1	Best1	Best5	Best5	Best5	64.710	61.760	64.710	70.590	70.590
Top1	Best1	Naive	LinReg-G-L1	LinReg-G-L3	LinReg-G-L3	64.710	61.760	70.590	73.530	73.530
Top2	Naive	Best1	LinReg-G-L3	LASSO-MF-G	LASSO-MF-G	61.760	61.760	70.590	73.530	73.530
Top3	SPC(2)-MF	AR(1)	LASSO-MF-G	EN-MF-G	LinReg-G-L1	61.760	58.820	70.590	70.590	70.590
Top4	SPC(2)-MF-G	PLS(4)-MF-G	EN-MF-G	Spike-MF	EN-MF-G	61.760	58.820	67.650	70.590	70.590
Top5	AR(1)	SPC(2)-MF	Spike-MF	Spike-MF-G	Spike-MF-G	58.820	58.820	67.650	70.590	70.590
Target: Industrial Production (IP)										
AveNaive	Avg(12)	Avg(12)	Avg(12)	Avg(12)	Avg(12)	58.820	58.820	50.000	52.940	52.940
TS	BaggedETS	BaggedETS	AR(AIC)	AR(AIC)	BaggedETS	67.650	64.710	58.820	58.820	61.760
G	LinReg-G-L3	SPC(2)-MF-G	SPC(5)-MF-G	SPC(5)-MF-G	DFA5-MF-G	58.820	61.760	70.590	76.470	67.650
MF	DFA4-MF	DFA2-MF	DFA5-MF	DFA5-MF	DFA5-MF	55.880	58.820	67.650	70.590	70.590
Best	Best5	Best5	Best5	Best10	Best1	55.880	50.000	55.880	61.760	58.820
Top1	BaggedETS	BaggedETS	SPC(5)-MF-G	SPC(5)-MF-G	DFA5-MF	67.650	64.710	70.590	76.470	70.590
Top2	Spline	Spline	DFA5-MF	DFA5-MF	DFA5-MF-G	61.760	61.760	67.650	70.590	67.650
Top3	Avg(12)	SPC(2)-MF-G	SPC(5)-MF	SPC(4)-MF	SPC(4)-MF-G	58.820	61.760	67.650	70.590	67.650
Top4	AR(4)	Avg(12)	DFA5-MF-G	DFA5-MF-G	SPC(5)-MF	58.820	58.820	64.710	67.650	67.650
Top5	AutoArima	AR(4)	DFA3-MF	SPC(5)-MF	SPC(4)-MF	58.820	58.820	61.760	67.650	64.710
Target: Unemployment Rate (UR)										
AveNaive	Naive	Naive	Avg(12)	Avg(12)	Avg(12)	44.120	44.120	64.710	64.710	64.710
TS	AR(AIC)	AR(AIC)	AutoArima	AutoArima	AutoArima	38.240	38.240	58.820	58.820	58.820
G	PLS(5)-MF-G	PLS(5)-MF-G	PLS(2)-MF-G	PLS(2)-MF-G	PLS(2)-MF-G	38.240	38.240	64.710	64.710	64.710
MF	PLS(5)-MF	PLS(5)-MF	PLS(2)-MF	PLS(2)-MF	PLS(2)-MF	38.240	38.240	64.710	64.710	64.710
Best	Best10	Best10	Best10	Best5	Best5	38.240	38.240	58.820	58.820	58.820
Top1	Naive	Naive	Avg(12)	Avg(12)	Avg(12)	44.120	44.120	64.710	64.710	64.710
Top2	Avg(4)	Avg(4)	Avg(24)	Avg(24)	Avg(24)	38.240	38.240	64.710	64.710	64.710
Top3	AR(AIC)	AR(AIC)	PLS(2)-MF	PLS(2)-MF	PLS(2)-MF	38.240	38.240	64.710	64.710	64.710
Top4	BaggedETS	BaggedETS	PLS(2)-MF-G	PLS(2)-MF-G	PLS(2)-MF-G	38.240	38.240	64.710	64.710	64.710
Top5	PLS(5)-MF	PLS(5)-MF	PLS(4)-MF	PLS(4)-MF	PLS(4)-MF	38.240	38.240	64.710	64.710	64.710

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 9: UK, SSR.

Country: Germany (DE)										
	Best Model					Coverage Rate				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	4.710	7.650	4.710	4.710	4.710
TS	AR(1)	ETS	THETA	THETA	THETA	1.760	1.760	1.180	1.180	1.180
G	PLS(2)-MF-G	PLS(2)-MF-G	SPC(5)-MF-G	SPC(5)-MF-G	DFA4-MF-G	1.760	1.760	1.180	1.180	1.760
MF	PLS(2)-MF	PLS(2)-MF	DFA4-MF	SPC(5)-MF	SPC(5)-MF	4.710	1.760	1.760	1.180	1.180
Best	Best5	Best5	Best10	Best3	Best3	7.650	1.760	10.590	10.590	10.590
Top1	AR(1)	ETS	THETA	THETA	THETA	1.760	1.760	1.180	1.180	1.180
Top2	AR(4)	BATS	SPC(5)-MF-G	SPC(5)-MF	SPC(5)-MF	1.760	1.760	1.180	1.180	1.180
Top3	ETS	TBATS	AutoArima	SPC(5)-MF-G	AR(1)	1.760	1.760	1.760	1.180	1.760
Top4	BATS	THETA	DFA4-MF	AR(4)	AutoArima	1.760	1.760	1.760	1.760	1.760
Top5	TBATS	PLS(2)-MF	DFA4-MF-G	AR(AIC)	DFA4-MF	1.760	1.760	1.760	1.760	1.760
Target: Industrial Production (IP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	10.000	7.060	7.060	7.060	7.060
TS	NN	BaggedETS	BaggedETS	BaggedETS	NN	1.180	1.180	1.760	1.760	1.760
G	LASSO-MF-G	LASSO-MF-G	PLS(1)-MF-G	PLS(1)-MF-G	PLS(1)-MF-G	1.760	1.180	1.180	1.180	1.180
MF	LASSO-MF	LASSO-MF	PLS(1)-MF	PLS(1)-MF	PLS(1)-MF	1.760	1.180	1.180	1.180	1.180
Best	Best1	Best1	Best1	Best3	Best1	1.180	4.120	1.180	1.180	1.180
Top1	NN	BaggedETS	PLS(1)-MF	PLS(1)-MF	PLS(1)-MF	1.180	1.180	1.180	1.180	1.180
Top2	Best1	LASSO-MF	PLS(1)-MF-G	PLS(1)-MF-G	PLS(1)-MF-G	1.180	1.180	1.180	1.180	1.180
Top3	BaggedETS	LASSO-MF-G	PLS(2)-MF	PLS(2)-MF	PLS(2)-MF	1.760	1.180	1.180	1.180	1.180
Top4	LASSO-MF	EN-MF	PLS(2)-MF-G	PLS(2)-MF-G	PLS(2)-MF-G	1.760	1.180	1.180	1.180	1.180
Top5	LASSO-MF-G	EN-MF-G	SPC(1)-MF	SPC(1)-MF	SPC(1)-MF	1.760	1.760	1.180	1.180	1.180
Target: Unemployment Rate (UR)										
AveNaive	Avg(4)	Avg(4)	Avg(4)	Avg(4)	Avg(4)	7.060	7.060	10.000	7.060	10.000
TS	NN	NN	BATS	BATS	BATS	1.180	1.180	4.120	4.120	4.120
G	EN-MF-G	SPC(5)-MF-G	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	1.760	1.180	4.120	4.120	4.120
MF	LASSO-MF	Spike-MF	DFA4-MF	DFA4-MF	DFA4-MF	1.180	1.180	4.120	4.120	4.120
Best	Best1	Best1	Best1	Best1	Best1	10.000	7.060	1.180	1.180	4.120
Top1	NN	NN	Best1	Best1	BATS	1.180	1.180	1.180	1.180	4.120
Top2	LASSO-MF	SPC(5)-MF-G	BATS	BATS	TBATS	1.180	1.180	4.120	4.120	4.120
Top3	BATS	Spike-MF	TBATS	TBATS	LinReg-G-L3	1.760	1.180	4.120	4.120	4.120
Top4	TBATS	BATS	LinReg-G-L3	LinReg-G-L3	DFA4-MF	1.760	1.760	4.120	4.120	4.120
Top5	EN-MF-G	TBATS	DFA4-MF	DFA4-MF	DFA4-MF-G	1.760	1.760	4.120	4.120	4.120

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 10: DE, Reporting $CS = \left|\widehat{CR^{90\%}} - 90\%\right|$.

Country: France (FR)										
	Best Model					Coverage Rate				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	16.470	16.470	4.710	7.650	4.710
TS	Spline	AutoArima	AR(1)	THETA	AR(1)	10.590	10.590	1.180	1.180	1.180
G	LinReg-G	SPC(5)-MF-G	PLS(1)-MF-G	PLS(1)-MF-G	PLS(1)-MF-G	16.470	13.530	4.710	4.710	4.710
MF	DFA2-MF	SPC(5)-MF	PLS(1)-MF	PLS(1)-MF	PLS(1)-MF	16.470	13.530	4.710	4.710	4.710
Best	Best1	Best10	Best1	Best1	Best3	25.290	22.350	1.760	4.710	4.710
Top1	Spline	AutoArima	AR(1)	THETA	AR(1)	10.590	10.590	1.180	1.180	1.180
Top2	AR(1)	Spline	AR(AIC)	AR(1)	THETA	13.530	10.590	1.180	1.760	1.180
Top3	AR(4)	AR(4)	AutoArima	AR(AIC)	AutoArima	13.530	13.530	1.180	1.760	1.760
Top4	AutoArima	ETS	THETA	AutoArima	ETS	13.530	13.530	1.180	1.760	1.760
Top5	ETS	BATS	ETS	ETS	BATS	13.530	13.530	1.760	1.760	1.760
Target: Industrial Production (IP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	4.710	1.180	1.760	1.760	1.180
TS	AR(AIC)	NN	BaggedETS	BaggedETS	BaggedETS	1.180	1.180	1.180	1.180	1.180
G	DFA3-MF-G	LinReg-G	PLS(3)-MF-G	PLS(3)-MF-G	PLS(3)-MF-G	1.180	1.180	1.180	1.180	1.180
MF	DFA3-MF	DFA3-MF	PLS(3)-MF	PLS(3)-MF	PLS(3)-MF	1.180	1.180	1.180	1.180	1.180
Best	Best1	Best3	Best1	Best1	Best1	1.180	1.180	1.180	1.180	1.180
Top1	AR(AIC)	Naive	BaggedETS	BaggedETS	Naive	1.180	1.180	1.180	1.180	1.180
Top2	BaggedETS	NN	PLS(3)-MF	NN	BaggedETS	1.180	1.180	1.180	1.180	1.180
Top3	NN	LinReg-G	PLS(3)-MF-G	PLS(3)-MF	PLS(3)-MF	1.180	1.180	1.180	1.180	1.180
Top4	DFA3-MF	DFA3-MF	PLS(4)-MF	PLS(3)-MF-G	PLS(3)-MF-G	1.180	1.180	1.180	1.180	1.180
Top5	DFA3-MF-G	DFA3-MF-G	PLS(4)-MF-G	PLS(4)-MF	PLS(4)-MF	1.180	1.180	1.180	1.180	1.180
Target: Unemployment Rate (UR)										
AveNaive	Naive	Naive	Naive	Naive	Naive	4.710	1.760	4.120	4.120	7.060
TS	THETA	THETA	ETS	AR(1)	AR(4)	1.180	1.180	1.180	1.180	1.180
G	LinReg-G	LinReg-G	DFA3-MF-G	DFA3-MF-G	DFA3-MF-G	4.710	4.710	1.180	1.180	1.180
MF	SPC(1)-MF	SPC(1)-MF	DFA3-MF	DFA3-MF	DFA3-MF	4.710	4.710	1.180	1.180	1.180
Best	Best5	Best5	Best10	Best3	Best1	13.530	16.470	1.180	7.650	7.650
Top1	THETA	THETA	ETS	AR(1)	AR(4)	1.180	1.180	1.180	1.180	1.180
Top2	Naive	Naive	Spline	AR(4)	AutoArima	4.710	1.760	1.180	1.180	1.180
Top3	AR(1)	BATS	DFA3-MF	AutoArima	ETS	4.710	4.710	1.180	1.180	1.180
Top4	AutoArima	TBATS	DFA3-MF-G	ETS	Spline	4.710	4.710	1.180	1.180	1.180
Top5	BATS	LinReg-G	DFA4-MF	Spline	DFA3-MF	4.710	4.710	1.180	1.180	1.180

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 11: FR, Reporting $CS = \left|\widehat{CR^{90\%}} - 90\%\right|$.

Country: Italy (IT)										
	Best Model					Coverage Rate				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	33.480	29.130	11.740	20.430	3.040
TS	AutoArima	AR(1)	AR(1)	AR(1)	AR(1)	24.780	24.780	1.300	1.300	1.300
G	LinReg-G-L3	LinReg-G-L1	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	37.830	37.830	3.040	1.300	1.300
MF	Spike-MF	SPC(5)-MF	EN-MF	LASSO-MF	LASSO-MF	42.170	42.170	11.740	7.390	7.390
Best	Best3	Best3	Best1	Best3	Best3	33.480	33.480	3.040	1.300	1.300
Top1	AutoArima	AR(1)	AR(1)	AR(1)	AR(1)	24.780	24.780	1.300	1.300	1.300
Top2	AR(1)	AutoArima	AR(4)	AutoArima	AR(4)	29.130	24.780	1.300	1.300	1.300
Top3	Spline	Naive	AutoArima	LinReg-G-L3	LinReg-G-L3	29.130	29.130	3.040	1.300	1.300
Top4	Naive	Spline	LinReg-G-L3	Best3	Best3	33.480	29.130	3.040	1.300	1.300
Top5	AR(4)	AR(4)	Best1	Best5	Best5	33.480	33.480	3.040	1.300	1.300
Target: Industrial Production (IP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	5.650	10.000	1.300	10.000	5.650
TS	AR(AIC)	AR(AIC)	BaggedETS	AR(AIC)	AR(AIC)	3.040	3.040	1.300	1.300	1.300
G	LinReg-G-L3	LinReg-G-L3	DFA2-MF-G	DFA2-MF-G	DFA2-MF-G	1.300	3.040	1.300	1.300	1.300
MF	EN-MF	DFA2-MF	DFA2-MF	DFA2-MF	DFA2-MF	1.300	5.650	1.300	1.300	1.300
Best	Best1	Best1	Best1	Best1	Best1	5.650	5.650	1.300	1.300	1.300
Top1	LinReg-G-L3	AR(AIC)	Naive	AR(AIC)	AR(AIC)	1.300	3.040	1.300	1.300	1.300
Top2	DFA4-MF-G	LinReg-G-L3	BaggedETS	BaggedETS	BaggedETS	1.300	3.040	1.300	1.300	1.300
Top3	SPC(4)-MF-G	AR(1)	DFA2-MF	DFA2-MF	DFA2-MF	1.300	5.650	1.300	1.300	1.300
Top4	EN-MF	BaggedETS	DFA2-MF-G	DFA2-MF-G	DFA2-MF-G	1.300	5.650	1.300	1.300	1.300
Top5	EN-MF-G	LinReg-G	PLS(3)-MF	PLS(3)-MF	PLS(3)-MF	1.300	5.650	1.300	1.300	1.300
Target: Unemployment Rate (UR)										
AveNaive	Naive	Naive	Naive	Naive	Naive	11.740	16.090	3.040	3.040	3.040
TS	BATS	BATS	AR(4)	AR(4)	AR(1)	3.040	3.040	1.300	1.300	1.300
G	LinReg-G-L1	LinReg-G-L1	LinReg-G-L3	LinReg-G-L3	LinReg-G-L3	7.390	7.390	1.300	3.040	1.300
MF	SPC(5)-MF	SPC(4)-MF	DFA2-MF	DFA2-MF	DFA2-MF	7.390	7.390	3.040	3.040	3.040
Best	Best3	Best1	Best3	Best3	Best3	7.390	7.390	1.300	1.300	1.300
Top1	BATS	BATS	AR(4)	AR(4)	AR(1)	3.040	3.040	1.300	1.300	1.300
Top2	TBATS	TBATS	ETS	ETS	AR(AIC)	3.040	3.040	1.300	1.300	1.300
Top3	Spline	Spline	BATS	BATS	ETS	3.040	3.040	1.300	1.300	1.300
Top4	AR(4)	AR(1)	TBATS	TBATS	BATS	7.390	7.390	1.300	1.300	1.300
Top5	AR(AIC)	AR(AIC)	Spline	Spline	TBATS	7.390	7.390	1.300	1.300	1.300

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 12: IT, Reporting $CS = \left|\widehat{CR^{90\%}} - 90\%\right|$.

Country: United Kingdom (UK)										
	Best Model					Coverage Rate				
	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$	$-5w$	$-4w$	$-3w$	$-2w$	$-1w$
Target: Harmonised Index of Consumer Prices (HICP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	1.180	1.760	1.180	1.180	1.180
TS	AR(1)	ETS	AR(1)	AR(4)	AR(4)	1.180	1.180	1.180	1.180	1.180
G	LinReg-G	LinReg-G-L3	LinReg-G	LinReg-G	LinReg-G	1.760	1.180	1.180	1.180	1.180
MF	DFA2-MF	SPC(4)-MF	DFA4-MF	DFA4-MF	DFA4-MF	1.760	1.180	1.180	1.180	1.180
Best	Best3	Best5	Best5	Best3	Best3	1.760	10.590	1.180	1.180	1.180
Top1	Naive	ETS	Naive	Naive	Naive	1.180	1.180	1.180	1.180	1.180
Top2	AR(1)	BATS	AR(1)	AR(4)	AR(4)	1.180	1.180	1.180	1.180	1.180
Top3	ETS	TBATS	AR(4)	AR(AIC)	AR(AIC)	1.180	1.180	1.180	1.180	1.180
Top4	BATS	Spline	ETS	AutoArima	ETS	1.180	1.180	1.180	1.180	1.180
Top5	TBATS	LinReg-G-L3	BATS	ETS	BATS	1.180	1.180	1.180	1.180	1.180
Target: Industrial Production (IP)										
AveNaive	Naive	Naive	Naive	Naive	Naive	4.120	4.120	7.060	7.060	7.060
TS	BaggedETS	BaggedETS	AR(AIC)	AR(AIC)	AR(AIC)	1.180	1.180	1.180	1.180	1.760
G	LinReg-G-L3	LinReg-G	LinReg-G	LinReg-G	LinReg-G	1.180	4.120	7.060	7.060	7.060
MF	DFA2-MF	DFA2-MF	DFA2-MF	DFA2-MF	DFA2-MF	4.120	4.120	7.060	7.060	7.060
Best	Best1	Best1	Best1	Best1	Best1	4.120	4.120	7.060	7.060	7.060
Top1	BaggedETS	BaggedETS	AR(AIC)	AR(AIC)	AR(AIC)	1.180	1.180	1.180	1.180	1.760
Top2	LinReg-G-L3	Naive	BaggedETS	Naive	Naive	1.180	4.120	4.120	7.060	7.060
Top3	AR(AIC)	NN	Naive	AR(4)	AR(1)	1.760	4.120	7.060	7.060	7.060
Top4	Naive	LinReg-G	AR(1)	AutoArima	AR(4)	4.120	4.120	7.060	7.060	7.060
Top5	NN	LinReg-G-L1	AR(4)	ETS	AutoArima	4.120	4.120	7.060	7.060	7.060
Target: Unemployment Rate (UR)										
AveNaive	Naive	Naive	Naive	Naive	Naive	19.410	28.240	1.180	1.180	4.120
TS	Spline	AR(4)	AR(1)	AR(AIC)	AR(AIC)	7.650	7.650	1.180	1.760	1.180
G	SPC(4)-MF-G	PLS(1)-MF-G	PLS(3)-MF-G	PLS(3)-MF-G	PLS(3)-MF-G	25.290	25.290	1.760	1.180	1.180
MF	PLS(5)-MF	PLS(1)-MF	PLS(1)-MF	PLS(1)-MF	PLS(1)-MF	28.240	28.240	1.760	1.180	1.180
Best	Best3	Best3	Best5	Best10	Best5	31.180	31.180	1.180	1.180	1.760
Top1	Spline	AR(4)	Naive	Naive	AR(AIC)	7.650	7.650	1.180	1.180	1.180
Top2	ETS	Spline	AR(1)	PLS(1)-MF	PLS(1)-MF	10.590	7.650	1.180	1.180	1.180
Top3	BATS	AutoArima	AutoArima	PLS(3)-MF	PLS(3)-MF	10.590	10.590	1.180	1.180	1.180
Top4	TBATS	ETS	ETS	PLS(3)-MF-G	PLS(3)-MF-G	10.590	10.590	1.180	1.180	1.180
Top5	AR(4)	BATS	BATS	PLS(4)-MF	PLS(4)-MF	13.530	10.590	1.180	1.180	1.180

Notes: We report the top model in each category and the top five models overall. AveNaive: simple averages and naïve forecasts, TS: univariate time series models, G: any model which includes Google-based indicators, MF: models which only include standard Macro and Financial indicators, Best: best model selection strategies, Top1-Top5: top performing models overall.

Table 13: UK, Reporting $CS = \left|\widehat{CR^{90\%}} - 90\%\right|$.