ECONOMIC STATISTICS CENTRE OF EXCELLENCE

A COLLABORATION WITH

Office for National Statistics

A Quality Assessment Framework for Maintaining & Publishing New Indicators

George Kapetanios and Fotis Papailias

ESCoE Technical Report TR-18

May 2022

ISSN 2515-4664

TECHNICAL REPORT

A Quality Assessment Framework for Maintaining & Publishing New Indicators
George Kapetanios and Fotis Papailias
ESCoE Technical Report TR-18
May 2022

Abstract

This technical report builds on the research output of “National Accounts and Beyond GDP: Predictive Performance of Real-Time Indicators” ESCoE/ONS collaborative project to summarise the key findings and provide a standardised methodology to guide the editing, maintenance, publishing and potential incorporation of new real-time indicators into the development of early estimates of headline economic statistics. Empirical results from previous tasks are revisited and standardised qualitative and quantitative criteria are discussed.

Keywords: nowcasting, machine learning, real-time indicators

JEL classification: C53, E37

Fotis Papailias, King’s College London
fotis.papailias@kcl.ac.uk

Published by:
Economic Statistics Centre of Excellence
National Institute of Economic and Social Research
2 Dean Trench St
London SW1P 3HE
United Kingdom
www.escoe.ac.uk

A Quality Assessment Framework for Maintaining & Publishing New Indicators

George Kapetanios Fotis Papailias[*]
King's Business School
Data Analytics Centre for Macro and Finance
Economic Statistics Centre of Excellence

April 06, 2022

Abstract

This technical report builds on the research output of "National Accounts and Beyond GDP: Predictive Performance of Real-Time Indicators" ESCoE/ONS collaborative project to summarise the key findings and provide a standardised methodology to guide the editing, maintenance, publishing and potential incorporation of new real-time indicators into the development of early estimates of headline economic statistics. Empirical results from previous tasks are revisited and standardised qualitative and quantitative criteria are discussed.

Keywords: Nowcasting, Machine Learning, Real-Time Indicators.

[*]Corresponding author. King's Business School, King's College London, Bush House, 30 Aldwych, London, WC2B 4BG, UK. Knot Analytics Ltd. E-mail: fotis.papailias@kcl.ac.uk

JEL Codes: C53, E37.

Contents

1 Summary of the Research Output

This section summarises the "National Accounts and Beyond GDP: Predictive Performance of Real-Time Indicators" ESCoE/ONS collaborative project main tasks and output. As described below, in this project we adopt a "top-down" approach. We start from the general case (*top*) to investigate whether a large set of new predictors can lead to gains in nowcasting complementing a standard set of macroeconomic predictors and then continue (*mid*) with the examination of individual predictors and/or groups of predictors. Finally, we focus on a specific dataset (*down*) and provide a head-to-head comparison investigate gains in nowcasting using the main aggregates of this datasets versus linear and non-linear models based on the underlying disaggregate series.

1.1 RQ1: Standard Set of Predictors & ONS Real-Time Indicators

The first part of this project is concerned with the investigation of a large collection of "alternative" indicators; mainly the ONS Real-Time Indicators dataset. This set of indicators includes, among others, VAT indices, use of debit and credit cards, transport usage, business activities, online job advertisements, traffic cams data, COVID-19 surveys, online retail prices, and the traffic in UK ports. These series are available in monthly, weekly and, in some cases, even daily frequency. Their timely nature and almost no publication lags allows, in principle, their effective use in economic nowcasting and monitoring of real-time economic conditions.[1]

Therefore, a natural question to ask is "how do these indicators assist in economic nowcasting during times of crises?". The COVID-19 outbreak makes a perfect case study to focus on this dataset and examine its usefulness. Obviously, there can be many ways a researcher can evaluate a specific set of indicators and place it into different context. For example, a labour economist might not be so interested in

[1]We say "in principle" as these indicators might become available with some (minor) publication delays to the average researcher. However, ONS does have early access on the datasets which, in turn, allows for timely estimations.

economic nowcasting; instead, she might prefer to investigate the effect of COVID-19 on schooling (via platforms of online learning) and develop a more specialised framework in this context. Or, an energy economist might be willing to focus more on electricity and gas demand and consumption to provide an early warning system of gas imports and potential increase in prices. However, as here we approach this through the lens of a statistics institute, we take into account the "big picture", that is the total economy and this is why we focus on real-time monitoring and nowcasting. Of course, the framework we attempt to introduce here could be generalised and applied in different contexts.

On one hand we have the ONS Real-Time Indicators dataset. However, economists do have ways of producing real-time (or close to real-time) estimates even before the publication of timely indicators. Hence, in the first task we follow the literature by organising a dataset of standard macroeconomics and finance indicators complemented by the ONS Real-Time Indicators dataset. Our aim is first, to use the above dataset to construct a real-time coincident indicator for the UK and investigate which of the ONS Real-Time Indicators variables contribute more to this index at different times. Then, we employ this coincident indicator in a pseudo real-time out-of-sample nowcasting exercise which focuses on the COVID-19 period and investigate gains in terms of nowcasting error.

There are many difficulties in this setup. One main difficulty which is present across all the research outputs of this project is the very short availability of most of the ONS Real-Time Indicators variables which, most of them, become available from around mid-2019. Considering that the target variable is the monthly estimate of the GDP, this allows a total sample of about 30 monthly observations which is very short. Another difficulty is the difference in the publication dates; i.e. when some variable are available and, thus, have information, some others still have missing values. We attempt to solve the problem of missing values by suggesting three approaches for extracting factors: non-linear iterative partial least squares, the probabilistic PCA and the Bayesian missing value estimation. We refer the reader to Kapetanios and Papailias (2021a) for more technical details.

Our application starts with the construction of a real-time coincident indicator of economic activity during the COVID-19 outbreak by using: (i) strictly, the standard macroeconomics and financial set of indicators, (ii) both the standard and the ONS Real-Time Indicators datasets, and (iii) strictly the ONS Real-Time Indicators dataset. In all cases, the indicator is the first static principal component. More factor extraction methods and fine tuning could be considered here, however our aim is to use the simplest method to evaluate the novelty of the dataset. As it can been in Figure 1, the real-time coincident indicator based on strictly the ONS Real-Time Indicators dataset provides an accurate tracking of the real-time economic conditions already highlighting the usefulness of this dataset in economic applications.

Having illustrated the good performance of the coincident indicator, we continue with the use of this dataset in a pseudo out-of-sample cross-validation nowcasting exercise extracting principal component factors from the standard set of variables, the ONS Real-Time Indicators and both. At this stage, we have evidence that in 4 out of 5 model types (i.e. 80% of the cases), the inclusion of the Real-Time Indicators leads to minor improvements in the nowcasting error.

1.2 RQ2: Real-Time Indicators during the COVID-19 Pandemic

Having illustrated the usefulness of the Real-Time Indicators dataset, we continue by focusing more exclusively on it. RQ2 is based on weekly Real-Time Indicators which complement a weekly set of standard financial predictors.

Our second technical report aims to present a generalised framework for assessing the predictive content of ONS real-time indicators in both dimensions: (i) individual predictors (i.e. variable-by-variable), and (ii) using machine learning techniques to build variable selection models. It is important to notice that RQ1 only focuses on a standard static principal components analysis whereas RQ2 now considers state-of-the-art econometric and machine learning techniques such as best subset selection, penalised regressions (ridge, lasso, elastic net, adaptive lasso), principal components and random forests (as a non-linear machine learning representative). We also calculate the simple correlation and predictive power scores for each indicator across time.

Our empirical findings highlight the usefulness of the fast indicators dataset as well as the importance of some individual predictors. We refer the reader to Kapetanios and Papailias (2021b) for more details.

1.3 RQ2-3: VAT and CHAPS

Given the usefulness of RQ2, we take as an intermediate step between RQ2 and RQ3 the research output RQ2-3 which is a repetition of RQ2 on a very specific subset of the ONS Real-Time Indicators dataset: (i) the monthly VAT indices, and (ii) the weekly CHAPS usage on credit and debit cards.

As in RQ2, our approach utilises nowcasting and attempts to characterise the specific dataset in terms of its estimation gains. We take a general approach where the applied researcher can choose any target of their preference and study any group of indicators. The framework includes elements of correlation, such as the time-varying correlation and the Predictive Power Score, as well as attempts to capture causality via the use of nowcasting regressions, either with individual indicators or using the whole set and employ machine learning methods.

We use the monthly GDP and the monthly VAT diffusion indexes as an illustrative example of variables of the same frequency and with the same publication lags. Our results highlight the importance of publication lags and the report concludes that the VAT diffusion indices do not yield to improved nowcasts. If we really need to consider this set of variables then, perhaps, we are better off by picking single indicators. This, again, does not yield accurate estimates as the publication lag pattern is obvious in all results. However, our aim is not to do a critique to specific groups of indicators (i.e. the VAT diffusion indices). Instead, we want to show the framework's versatility when the target and the set of indicators are expressed in the same frequency with the same lags; obviously, if publication lags for the VAT indices were shorter, the overall conclusion would be totally different. We refer the reader to Kapetanios and Papailias (2021c) for more details.

In contrast, when we focus on the monthly GDP and the daily-translated-to-weekly CHAPS set of indicators we highlight the importance of “real-time” indicators and the empirical gains we can potentially obtain in applications when we consider leading indicators (or timely indicators in general). The contribution of the CHAPS report, however, is not to praise specific groups of indicators but again highlight the framework’s versatility when the target and the set of indicators are expressed in different frequencies. We refer the reader to Kapetanios and Papailias (2021d) for more details.

1.4 RQ3: Aggregates vs. Disaggregates

Finally, in RQ3 we focus even more on a specific set of ONS Real-Time Indicator to measure gains in nowcasting employing the total aggregated variables only versus models which aim to exploit the information in all the underlying disaggregated series.

At first, this might seem like a task with a limited scope. However, the applied researcher should view this approach as only a part of an overall assessment framework for novel datasets. Our main aim is to provide a framework which answers the following question regarding a candidate dataset of new indicators: “Should a national statistics institute invest resources in organising, editing, polishing and publishing this novel dataset of indicators and why?”.

Using two datasets from the ONS Real-Time Indicators we attempt to answer this question empirically via means of economic nowcasting. In particular, we consider a linear model which uses the main aggregates and compare its nowcasting performance with various, mainly machine learning-based, models which utilise all the underlying disaggregate series. In this report linear (penalised regressions and actor-based regressions) as well as non-linear (random forests, neural networks and support vector regressions) models are included. It is important to notice that RQ3 includes the addition of state-of-the-art sophisticated neural networks (Multilayer Perceptron and Extreme Learning Machines) as well as Support Vector Regressions.

Our findings provide empirical evidence in favour of the “big data” principle; i.e. in today’s world, national statistics institutes should publish data to some, if not the highest possible, level of disaggregation as most modern econometric techniques can handle these datasets and exploit their gains in economic applications such as nowcasting or forecasting. As expected, our results show that during crises, such as the COVID-19 outbreak, non-linear models tend to perform better than linear ones, however this reverts in periods of economic stability.

2 Criteria

Having summarised the key research outputs of this project, we now discuss some criteria which could be used by the applied researcher in the evaluation of a given novel dataset or new indicators. These criteria can be used again from a "top-down" perspective; i.e. examine the indicators usefulness in generic macroeconomic context, as in RQ1, and then isolate them, compare the disaggregates/variables one-by-one, as in RQ2 and RQ2-3, and finally compare the aggregates with the disaggregates, as in RQ3. Alternatively, one could adopt a "bottom-up" approach; i.e. given a set of novel indicators, compare the aggregates to the disaggregates and examine if the full set of indicators is useful, as in RQ3, then take these indicators to an individual and group-based comparison, as in RQ2 and RQ2-3, and finally include them in a wider macroeconomic application, as in RQ1. Below we list these criteria in terms of priority, however the order could change according to the preferences of the applied researcher. Figure 2 illustrates this decision pyramid.

2.1 Economic Measurement

The first criterion, which is found at the "top" of the pyramid, is the ability of a new set of indicators to create a new economic index or improve an existing one.

In RQ1 we considered the ONS Real-Time Indicators dataset and attempted to construct a real-time coincident indicator illustrating that the version of the indicator based on the Real-Time Indicators accurately monitors the economic conditions; see Figure 1.

Of course, the coincident indicator is only a working example. Our suggestion here is for the applied researcher to consider how a new set of novel variables could be combined to create a new index of some economic measurement or supplement an existing one. If this new dataset improves economic measurement, then we cam confidently argue that there is value in this dataset which should be further edited, polished, published and maintained.

On the other hand, if a new set of variables does not improve an existing economic index (or is not able to create a new one), we should not immediately discard this dataset, but we should further investigate its usefulness with some other criteria. We also refer the reader to Kapetanios and Papailias (2021a) for an illustrative example.

Based on the above discussion, we call this a "qualitative" criterion rather than a "quantitative" one. Of course the construction of an index is totally done using quantitative techniques, however the decision if the resulting index is one with favourable properties is a qualitative one.

2.2 Correlation

Another criterion which one could apply in practice is the calculation of the (time-varying) correlation between the variables of a novel set of indicators and a specific target variable.

If we find evidence that some or most of the variables are consistently correlated across time with our target, we can infer that there are possibly some underlying relationships which highlight the importance of this dataset. Furthermore, apart from contemporaneous correlations, i.e. y_t and x_t, it is worth investigating lead/lag relationships by calculating the correlation of leads, i.e. y_t and x_{t-k}, and lags y_{t-k} and x_t.

In the modern age of machine learning, some argue that the correlation coefficient does not accurately reflect non-linear relationships. We could solve this problem by calculating the (time-varying) Predictive Power Score. The Predictive Power Score can identify patterns in the data and feature selection even if these are non-linear, however it cannot be interpreted as easily as the correlation coefficient because it does not indicate anything about the type of the relationship. We refer the reader to Kapetanios and Papailias (2021b, 2021c, 2021d) for some illustrative examples.

The above measures are what we call pure “quantitative” criteria since they really illustrate some underlying relationships without the need of action from the researcher.

2.3 Econometric Modelling I: All Indicators

Another type of criterion is the performance of various econometric models when the underlying dataset is in use. This can be examined in various economics applications such as nowcasting, forecasting, backcasting, etc., depending on the relevant research context.

The performance of various models with and without the dataset of interest in a pseudo out-of-sample nowcasting exercise is measured in all Kapetanios and Papailias (2021a to 2021d and 2022) reports. Statistics such as the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in the nowcasting are reported and a conclusion on the usefulness of the underlying dataset can be drawn. Therefore, this is another quantitative criterion. More statistics, such as the sign success ratio and nowcast density estimation, could further be considered.

Based on the above, the researcher could further construct a version of “count” statistic. For example, one could count the number of models which improve in terms of nowcasting MAE or RMSE when the underlying dataset is used. Then, if the number of models which improve exceeds a user threshold, we can argue that the dataset has quality components which can be exploited by the majority of models and, thus, should be further edited, polished, published and maintained.

2.4 Econometric Modelling II: Aggregates vs. Disaggregates

Another type of criterion similar to Econometric Modelling I is the use of statistics in a relevant economic application, such as nowcasting, comparing the performance of simple models using the total aggregates of the dataset to complex models which aim to exploit the information which is hidden in the disaggregated data.

Again, one could further construct a count statistic to measure the number of models which improve when the disaggregated variables are used. This provides ample quantitative support to the argument in favour of the editing, polishing, publishing and maintaining all the appropriate variables of the underlying dataset. We refer the reader to Kapetanios and Papailias (2022) for some illustrative examples.

3 Working Examples

In this section, we revisit some examples from RQ1 to RQ3 and apply the criteria discussed in the previous section to assess the quality and usefulness of the underlying datasets.

3.1 Example 1: A Real-Time Coincident Indicator

As we have also mentioned earlier, Figure 1 shows the construction of a coincident indicator in real-time (red line) using most of the ONS Real-Time Indicators variables. We clearly see that this index closely captures the movement of the monthly Gross Value Added (which is the relevant target variable).

This application -on its own- without the need of any other statistic clearly illustrates the quality and usefulness of the Real-Time Indicators dataset in economic applications. Obviously, one cannot argue that this is the purpose that this set of indicators has been organised. Instead, this type of economic application is an illustrative example; perhaps, more applications could be considered with even better results.

Then, Table 1 shows the performance of this real-time coincident indicator in a pseudo out-of-sample nowcasting exercise. In this case, we can construct a count statistic to measure the number of times a factor regression model has smaller nowcast error when including the Real-Time Indicators dataset. In this case, we see that 21 out of the 24 models which include the Real-Time Indicators dataset improve in terms of RMSE; this corresponds to 87.5% if the models indicating the underlying quality of the dataset.

It is important to notice that the above count statistic is not based on 12 different types of models with different number of factors, factor selection method and lags of the dependent variable. This ensures that this result is not biased towards a specific model type but improves the estimation across different (possibly misspecified) models.

3.2 Example 2: Correlation Across Time and Nowcasting using Individual Variables

Next, we revisit the results in Kapetanios and Papailias (2021b) report and, particularly, the correlation indicators reported in Table 2. This table reports the *absolute* correlation between each variable and the disaggregated monthly-to-weekly target as we are mainly interested in the existence of an underlying relationship and not necessarily in its direction. We present the summary statistics (mean, median, min and max) for the correlation for each indicator *across* all the rounds in the out-of-sample cross-validation exercise. Large mean and median values indicate that the corresponding variable is correlated or can predict the target variable. On top of the summary statistics, we also report the Top1, Top5, Top10, Top15 and Top20 percentage statistics which measure the number of times a variable ranks in the TopX positions of the corresponding statistic; for example, if a variable has a Top1 of 25% in the absolute correlation means that this variable is ranked with the largest correlation in 25% of the cross-validation rounds.

Table 2 shows that, across time, 55 out of 60 (91.67%) variables have an absolute correlation greater than 0.05, 47 out of 60 (78.33%) variables have an absolute correlation greater than 0.1 and 18 out of 60 (30%) variables have an absolute correlation greater than 0.2. These are encouraging results since they indicate a significant absolute correlation in a large number of indicators. It is important to highlight that these results correspond to the GVA target variable; if the researcher changes the target might find other even more encouraging results. Obviously, this measure can also be used in the Aggregate vs. Disaggregate comparison.

Continuing with the results in Table 3, one could again calculate a count statistic to measure the number of models which improve when the Real-Time Indicators dataset is included. This table shows that 14 out 20 (70%) of the models improve in terms of MAE and RMSE in nowcasting when the weekly Real-Time Indicators dataset is included solely or together with the standard set of weekly indicators. This reaches to a safe conclusion regarding the usefulness of the underlying dataset.

It is again important to highlight that this results comes from models across various methodologies such as best subset selection, ridge, lasso, elastic net and adaptive lasso, random forests and principal components. This really ensures that the gains in nowcasting due to the underlying dataset are not model-specific and, therefore, can be attributed the quality of the dataset.

This can be further seen in Tables 4 and 5. Considering individual variables out of the standard set of weekly indicators, Table 4, we see that the range of relative MAE is (0.932, 0.946) with an average of 0.942. THE corresponding relative RMSE ranges in (0.968, 0.975) with an average of 0.973. However, when we consider individual variables from the weekly Real-Time Indicators dataset we see that the relative MAE ranges in (0.369, 1.199) with an average of 0.736. Similarly, the relative RMSE is in the range of (0.257, 1.098) with an average of 0.662. We can argue that the Real-Time Indicators, on average, reduce the MAE by 21.95% and the RMSE by 31.93% indicating gains in nowcasting. This result, in turn, highlight the quality and usefulness of the underlying dataset.

3.3 Example 3: Aggregates vs. Disaggregates

The final example we revisit comes from the third research output, RQ3, where our aim is to evaluate whether an official statistics institute should publish an aggregate variable out of a dataset or if there is research potential publishing disaggregated variables of a dataset. Kapetanios and Papailias (2022) use two examples: (i) one based on the online job advertisements, and (ii) one based on the UK ports traffic.

Table 6 is concerned with the example of the online job advertisements. In this case we see that there are three benchmark models: the standard autoregressive models, AR(1) and AR(P), as well as the “Aggregate” model. The purpose of this task is to evaluate if a linear model which uses the "Aggregate" as the only predictor should be preferred against models from various methodologies which use all the dissagregates. Of course, a critique would be on the modelling process; i.e. the researcher could be biased towards a specific methodology which might result in nowcasting gains due to data snooping. This argument can be easily addressed by the large variety of linear and non-linear we well as simple and more complex machine learning methods which are used in this task. In particular, our linear modelling strategy includes the best subset selection, ridge, lasso and elastic net, adaptive lasso using the ridge initial estimates, adaptive lasso using the lasso initial estimates, as well as nowcasting regressions using PCA-based factor extraction. Our non-linear modelling strategy includes random forests, two sophisticated neural networks approaches, the Multilayer Perceptron and the Extreme Learning Machines as well as the Support Vector Regression using a radial basis function.

As in the previous examples, one could construct a "count" statistic which measures how many models, which are based on the entire universe of disaggregate variables, improve in terms of nowcasting. For example, looking at Table 6 we can see that in the full sample case 12 out of 13 models (92.31%) provide a smaller MAE and RMSE than the aggregate linear model. Similarly, in the subsample case we have that 11 out of 13 models (84.62%) improve. These results are encouraging and indicate that there is scientific reasoning on why ONS choose to publish and maintain the disaggregates of this dataset.

4 Concluding Remarks

The "National Accounts and Beyond GDP: Predictive Performance of Real-Time Indicators" ESCoE/ONS collaborative project consists of four separate tasks. Each task is concerned with different aspects of the predictive performance of alternative and more timely indicators; illustrative examples are provided in Kapetanios and Papailias (2021a to 2021d and 2022).

In the first task, RQ1, we investigate whether a wide dataset of timely indicators can provide new insights in a macroeconomic application. In particular, we use the ONS Real-Time Indicators dataset to construct a real-time coincident indicator which is shown to accurately illustrate the economic conditions during the COVID-19 pandemic. Also, a nowcasting exercise using PCA factors extracted from this dataset shows some weak signs of nowcasting gains.

Then, RQ2 is concerned with the closer examination of the ONS Real-Time Indicators dataset in economic nowcasting using various state-of-the-art linear and non-linear models based on machine learning methodologies. This task provides the wider context of assessing the quality of a dataset across various models ensuring that the conclusion is not biased towards one specific method. Also, it ensures that if various models improve in nowcasting, then this source of improvement must come from the specific dataset in use and not the idiosyncratic components of each model. RQ2-3 is heavily based on RQ2 and allows the researcher to focus on a specific category or type of variables and investigate each predictor individually as well as group of specific predictors.

Then, RQ3 further expands the wide range of models, including three non-linear models based on neural networks and support vector machines, and attempts to guide the official statistics institute on whether disaggregates of a specific dataset should be edited, polished, published and maintained or if it is better to simply publish some aggregated variables out of the underlying dataset.

Overall, this project provides a generic framework for quality assessment of new timely indicators. As argued throughout the tasks, and also earlier in this report, the main criterion used here is the predictive ability of various indicators in nowcasting. Of course, the framework can be further generalised to contemporaneous regressions, forecasting or even backcasting.

5 References

1. Kapetanios G., Papailias F., 2021a. UK Economic Conditions during the Pandemic: Assessing the Economy using ONS Faster Indicators. ESCoE Working Paper, DP 2021-10. Available here.

2. Kapetanios G., Papailias F., 2021b. Faster Indicators during the COVID-19 Pandemic: Individual Predictors & Selection. ESCoE Technical Report.

3. Kapetanios G., Papailias F., 2021c. An Evaluation Framework for Targeted Indicators, Part I: Monthly VAT. ESCoE Technical Report.

4. Kapetanios G., Papailias F., 2021d. An Evaluation Framework for Targeted Indicators, Part II: Weekly CHAPS. ESCoE Technical Report.

5. Kapetanios G., Papailias F., 2022. An Evaluation Framework for Targeted Indicators, Aggregates vs. Disaggregates. ESCoE Technical Report.

Figures

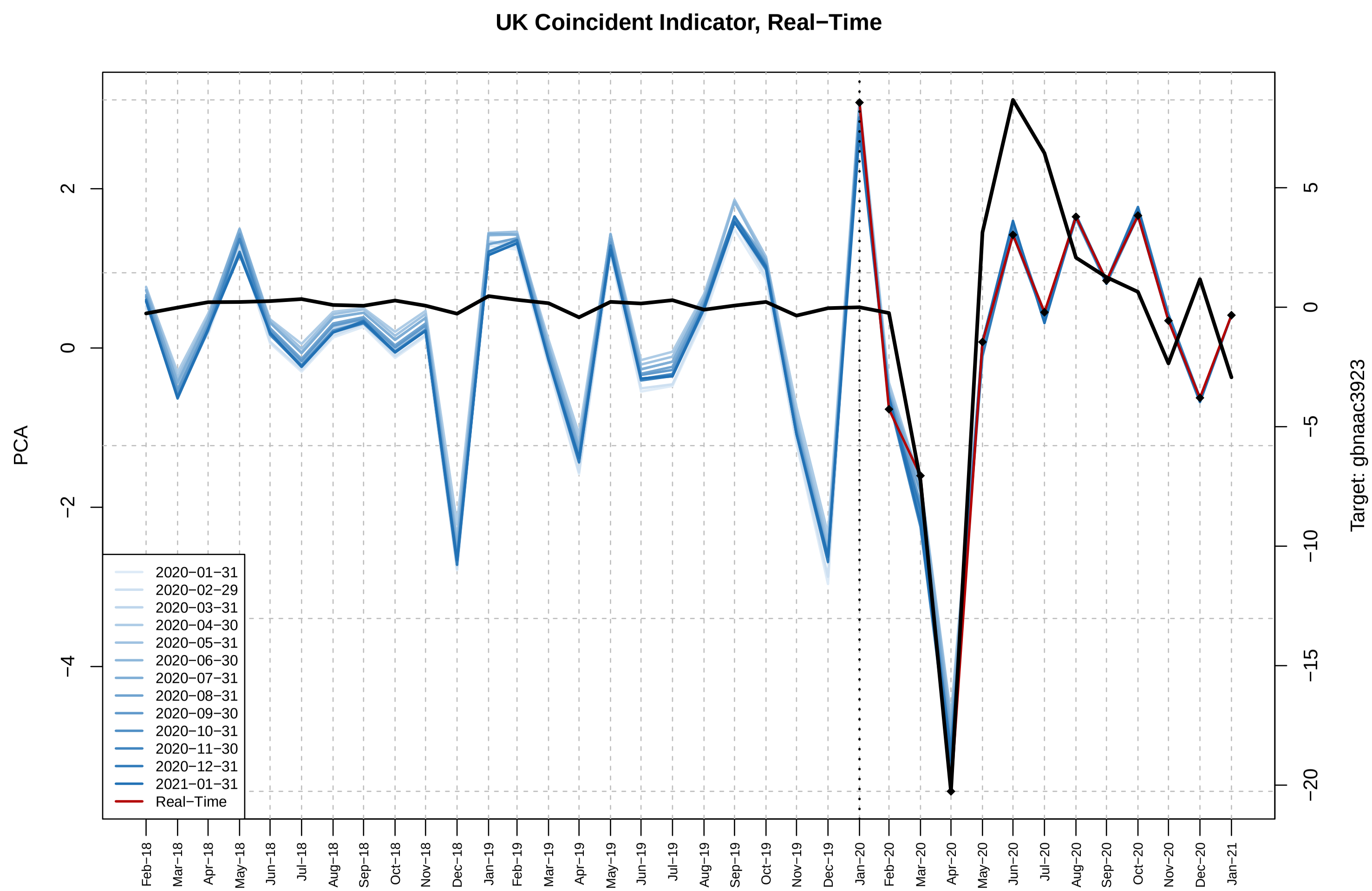

Figure 1: Figure 5 of Kapetanios and Papailias (2021a). The coincident indicator for the UK extracted using *only* the real-time indicators, FI, variables; target variable: monthly GDP growth. The black solid line denotes the actual values of the target variable. The various blue lines denote the estimate of the coincident indicator during the out-of-sample dates; light colours indicate the vintage in older dates and darker colours indicate the vintage in more recent dates. The red solid line (with black diamonds) indicates the estimate of the current value of the coincident indicator obtained at each round in the out-of-sample evaluation.

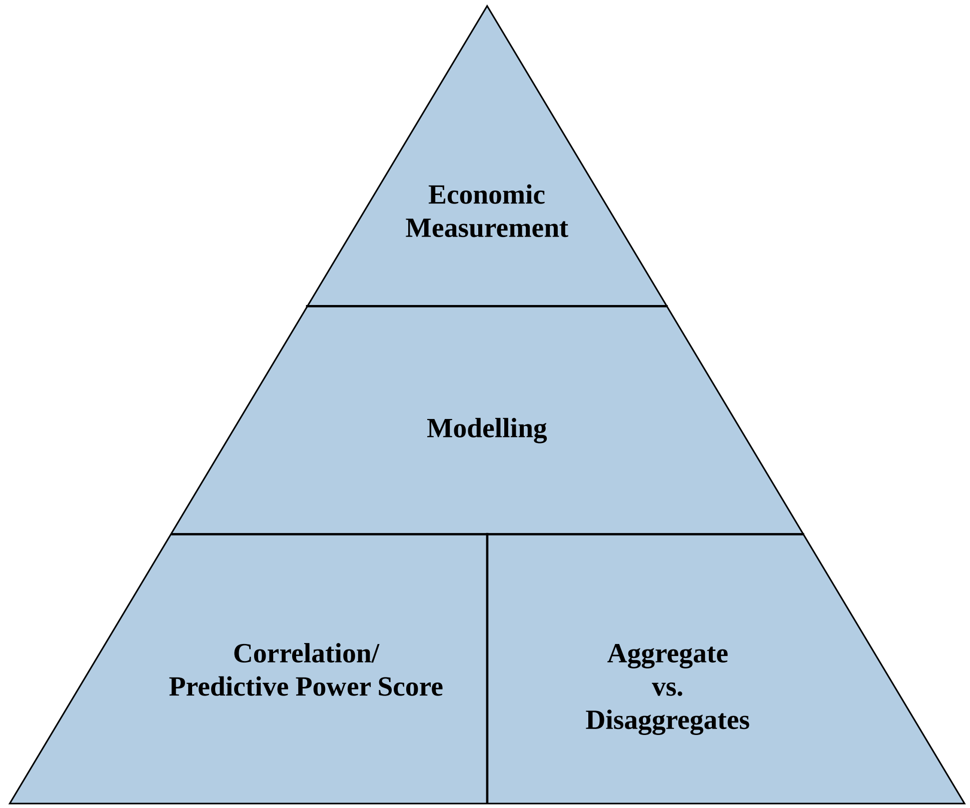

Figure 2: Decision pyramid of the relevant criteria for assessing the quality of a new set of timely economic indicators. The figure shows a “top-down” approach starting from the usefulness of the dataset in a wide economic application (Economic Measurement) ending to simple statistics to identify lead/lag relationships and the usefulness of the underlying disaggregated data. As discussed in Section 2, the researcher could apply a “top-down” approach as well as a “bottom-up” approach; i.e. starting from the usefulness of the disaggregates and/or simple statistics to illustrate the underlying relationships escalating with the use of the dataset in a wide economic application.

Tables

Model	GVA	Model	GVA
PCA(1)-(SD)	0.995	**PCA(4)-(SD)**	1.002
PCA(1)-(FI)	0.985	**PCA(4)-(FI)**	0.978
PCA(1)-(SD,FI)	0.983	**PCA(4)-(SD,FI)**	0.989
PCA(1)-(SD)-YL	1.031	**PCA(4)-(SD)-YL**	0.985
PCA(1)-(FI)-YL	**0.975**	**PCA(4)-(FI)-YL**	**0.965**
PCA(1)-(SD,FI)-YL	0.996	**PCA(4)-(SD,FI)-YL**	0.976
PCA(2)-(SD)	0.994	**PCA(5)-(SD)**	1.007
PCA(2)-(FI)	0.993	**PCA(5)-(FI)**	0.970
PCA(2)-(SD,FI)	0.979	**PCA(5)-(SD,FI)**	0.992
PCA(2)-(SD)-YL	1.007	**PCA(5)-(SD)-YL**	0.998
PCA(2)-(FI)-YL	0.983	**PCA(5)-(FI)-YL**	**0.944**
PCA(2)-(SD,FI)-YL	**0.975**	**PCA(5)-(SD,FI)-YL**	0.981
PCA(3)-(SD)	0.989	**PCA(r_{90})-(SD)**	1.006
PCA(3)-(FI)	0.994	**PCA(r_{90})-(FI)**	0.969
PCA(3)-(SD,FI)	0.974	**PCA(r_{90})-(SD,FI)**	0.985
PCA(3)-(SD)-YL	**0.967**	**PCA(r)-(SD)-YL**	1.000
PCA(3)-(FI)-YL	0.977	**PCA(r_{90})-(FI)-YL**	**0.945**
PCA(3)-(SD,FI)-YL	0.973	**PCA(r_{90})-(SD,FI)-YL**	0.973

Table 1: Table 2 of Kapetanios and Papailias (2021a). Nowcasting results averaged across 2020M01 to 2021M01 (13 periods). (SD) indicates that factors are extracted from the traditional SD set of variables only. (FI) indicates that factors are extracted from the FI set of variables only. (SD,FI) indicates that factors are extracted using the full set of SD and FI variables. YL indicates the inclusion of the first lag of the target variable as a potential predictor. Reported is the RMSFE as defined in Equation (4) of the paper relative to the AR(1) benchmark. r_{90} means that we extract as many components which explain up to 90% of the total variance.

	Absolute Correlation									
	Mean	**Median**	**SD**	**Min**	**Max**	**Top1**	**Top5**	**Top10**	**Top15**	**Top20**
gbbank6839	0.200	0.204	0.127	0.005	0.366	0.0%	22.9%	61.4%	78.6%	84.3%
gbtran0354	0.142	0.159	0.052	0.021	0.216	0.0%	7.1%	10.0%	32.9%	47.1%
gbtran0347	0.209	0.228	0.132	0.001	0.408	14.3%	37.1%	62.9%	68.6%	74.3%
gbtran0346	0.133	0.071	0.118	0.003	0.412	0.0%	12.9%	30.0%	60.0%	62.9%
gbtran0348	0.136	0.058	0.139	0.001	0.424	0.0%	18.6%	34.3%	52.9%	61.4%
gbbust1597	0.109	0.109	0.042	0.013	0.182	0.0%	1.4%	21.4%	30.0%	40.0%
gbbust1596	0.142	0.129	0.076	0.002	0.236	0.0%	2.9%	11.4%	27.1%	62.9%
gblama0103	0.179	0.254	0.109	0.000	0.293	0.0%	2.9%	14.3%	51.4%	80.0%
gblama0083	0.163	0.208	0.075	0.007	0.272	1.4%	1.4%	11.4%	35.7%	55.7%
gblama0081	0.255	0.338	0.146	0.000	0.448	7.1%	47.1%	72.9%	84.3%	92.9%
gblama0076	0.157	0.187	0.053	0.013	0.215	0.0%	12.9%	27.1%	38.6%	41.4%
gblama0091	0.219	0.251	0.059	0.022	0.287	10.0%	27.1%	41.4%	71.4%	90.0%
gblama0101	0.083	0.077	0.031	0.030	0.186	1.4%	5.7%	31.4%	34.3%	38.6%
gblama0097	0.048	0.044	0.030	0.004	0.184	1.4%	7.1%	8.6%	11.4%	27.1%
gblama0093	0.206	0.247	0.110	0.018	0.332	0.0%	11.4%	35.7%	72.9%	87.1%
gblama0082	0.200	0.227	0.059	0.004	0.294	1.4%	27.1%	35.7%	42.9%	74.3%
gblama0100	0.266	0.360	0.139	0.028	0.397	5.7%	65.7%	81.4%	91.4%	100.0%
gblama0079	0.105	0.134	0.053	0.004	0.192	0.0%	2.9%	5.7%	10.0%	32.9%
gblama0080	0.198	0.241	0.076	0.006	0.302	0.0%	24.3%	32.9%	44.3%	85.7%
gblama0077	0.196	0.265	0.100	0.003	0.290	0.0%	2.9%	32.9%	74.3%	91.4%
gbtran0460	0.260	0.332	0.174	0.008	0.468	47.1%	52.9%	60.0%	77.1%	94.3%
gbtran0464	0.210	0.267	0.152	0.003	0.417	0.0%	28.6%	44.3%	67.1%	90.0%
gbtran0463	0.259	0.300	0.118	0.003	0.376	3.8%	32.1%	71.7%	83.0%	86.8%
gbtran0454	0.311	0.332	0.096	0.000	0.424	0.0%	25.5%	80.9%	89.4%	93.6%
gbtran0458	0.276	0.310	0.103	0.030	0.368	2.1%	14.9%	74.5%	83.0%	87.2%
gbtran0457	0.275	0.331	0.118	0.004	0.387	0.0%	25.5%	70.2%	76.6%	80.9%
gbtran0476	0.184	0.199	0.069	0.036	0.285	0.0%	0.0%	2.1%	12.8%	25.5%
gbtran0479	0.178	0.215	0.079	0.000	0.275	0.0%	0.0%	0.0%	0.0%	19.1%
gbtran0478	0.163	0.189	0.078	0.005	0.448	2.1%	2.1%	2.1%	2.1%	4.3%
gbtran0481	0.209	0.228	0.070	0.052	0.296	0.0%	0.0%	11.1%	31.1%	53.3%
gbtran0484	0.216	0.242	0.062	0.065	0.277	0.0%	0.0%	0.0%	17.8%	68.9%
gbtran0483	0.176	0.197	0.068	0.022	0.255	0.0%	0.0%	0.0%	2.2%	13.3%
gbtran0471	0.156	0.164	0.055	0.012	0.248	0.0%	0.0%	0.0%	4.7%	7.0%
gbtran0474	0.128	0.096	0.090	0.021	0.268	0.0%	0.0%	0.0%	4.7%	4.7%
gbtran0473	0.122	0.088	0.073	0.012	0.285	0.0%	0.0%	2.3%	2.3%	9.3%
gbtran0466	0.172	0.179	0.082	0.008	0.295	0.0%	0.0%	2.4%	23.8%	26.2%
gbtran0469	0.222	0.201	0.057	0.140	0.457	2.5%	5.0%	10.0%	20.0%	40.0%
gbtran0468	0.146	0.144	0.071	0.013	0.251	0.0%	0.0%	0.0%	0.0%	0.0%
gbsurv00257	0.075	0.070	0.040	0.014	0.187	0.0%	0.0%	0.0%	0.0%	2.5%
gbsurv00260	0.244	0.245	0.065	0.103	0.463	5.4%	8.1%	8.1%	21.6%	54.1%
gbsurv00264	0.060	0.047	0.043	0.013	0.194	0.0%	0.0%	0.0%	0.0%	0.0%
gbsurv00262	0.048	0.039	0.040	0.004	0.146	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0101	0.130	0.127	0.050	0.065	0.259	0.0%	0.0%	0.0%	2.7%	2.7%
gbpric0102	0.094	0.095	0.045	0.001	0.220	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0105	0.092	0.094	0.058	0.000	0.191	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0124	0.088	0.053	0.064	0.026	0.260	0.0%	0.0%	0.0%	5.6%	5.6%
gbpric0129	0.096	0.087	0.050	0.016	0.181	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0135	0.073	0.072	0.036	0.004	0.148	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0142	0.046	0.028	0.046	0.002	0.190	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0161	0.123	0.109	0.049	0.056	0.240	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0178	0.145	0.148	0.047	0.015	0.235	0.0%	0.0%	0.0%	0.0%	5.9%
gbpric0189	0.127	0.134	0.037	0.051	0.210	0.0%	0.0%	0.0%	0.0%	0.0%
gbpric0212	0.111	0.039	0.112	0.001	0.283	0.0%	0.0%	0.0%	2.9%	11.8%
gbpric0221	0.100	0.128	0.053	0.003	0.167	0.0%	0.0%	0.0%	0.0%	0.0%
gbtran0356	0.118	0.122	0.014	0.099	0.140	0.0%	0.0%	0.0%	0.0%	0.0%
gbtran0367	0.141	0.136	0.023	0.114	0.216	0.0%	0.0%	0.0%	0.0%	2.9%
gbtran0361	0.029	0.027	0.006	0.019	0.039	0.0%	0.0%	0.0%	0.0%	0.0%
gbtran0359	0.248	0.249	0.023	0.216	0.307	0.0%	0.0%	8.8%	41.2%	58.8%
gbtran0360	0.040	0.037	0.014	0.027	0.074	0.0%	0.0%	0.0%	0.0%	0.0%
gbtran0390	0.145	0.147	0.015	0.102	0.163	0.0%	0.0%	0.0%	0.0%	0.0%

Table 2: Table 3 of Kapetanios and Papailias (2021b). Absolute correlation statistics across the nowcasting exercise on individual weekly real-time Indicators.

	MAE	RMSFE		MAE	RMSFE
AR(1)	1.000	1.000	**ALasso1_W_FW**	0.923	0.962
AR($\widehat{P}$)	1.262	1.284	**ALasso2_W**	0.933	0.967
BFW_W	0.929	0.966	**ALasso2_FW**	0.935	0.982
BFW_FW	1.026	1.044	**ALasso2_W_FW**	0.898	0.958
BFW_W_FW	0.916	0.978	**RF_W**	0.939	0.970
Ridge_W	0.945	0.974	**RF_FW**	0.918	0.969
Ridge_FW	0.939	0.969	**RF_W_FW**	0.912	0.965
Ridge_W_FW	0.939	0.968	**PCA(1)_W**	0.942	0.972
Lasso_W	0.940	0.973	**PCA(1)_FW**	0.946	0.970
Lasso_FW	0.953	0.984	**PCA(1)_W_FW**	0.947	0.971
Lasso_W_FW	0.908	0.960	**PCA($\widehat{k}_1$)_W**	0.940	0.971
EN_W	0.942	0.974	**PCA($\widehat{k}_1$)_FW**	0.964	0.973
EN_FW	0.934	0.969	**PCA($\widehat{k}_1$)_W_FW**	0.985	0.973
EN_W_FW	0.909	0.959	**PCA($\widehat{k}_2$)_W**	0.953	0.979
ALasso1_W	0.936	0.969	**PCA($\widehat{k}_1$)_FW**	0.928	0.967
ALasso1_FW	0.904	0.955	**PCA($\widehat{k}_1$)_W_FW**	0.943	0.972

Table 3: Table 8 of Kapetanios and Papailias (2021b). Nowcast MAE and RMSE statistics relative to AR(1) across the nowcasting exercise (where data is available) based on models using W, FW and both W and FW indicators. $\widehat{P}$ is chosen via AIC, ALasso1 is the adaptive Lasso using the Ridge coefficients in the adaptive penalty, ALasso2 is the adaptive Lasso using the original Lasso coefficients in the adaptive penalty, RF is the random forests, $\widehat{k}_1$ is chosen via cross-validation and $\widehat{k}_2$ is chosen in a fashion similar to Bai (2003).

	MAE	RMSFE		MAE	RMSFE
Rec-AR(1)	1.000	1.000	**gbboeexri0002**	0.943	0.972
gbeqin0001	0.946	0.974	**gbpboeusdfix**	0.945	0.974
gbeqin0001__pe	0.932	0.969	**gbpboeeurfix**	0.946	0.974
gbeqin0005	0.941	0.972	**gb3mgov**	0.944	0.973
gbeqin0005__pe	0.946	0.975	**gb1ygov**	0.943	0.973
gbeqin0206	0.941	0.972	**gb3ygov**	0.943	0.972
agbp10y	0.946	0.973	**gb5ygov**	0.943	0.973
bbbgbp10y	0.943	0.973	**gb10ygov**	0.940	0.971
gbpusd1m__outr__boe	0.945	0.974	**gb20ygov**	0.936	0.970
gbpusd3m__outr__boe	0.945	0.974	**gb30ygov**	0.933	0.968
gbpusd6m__outr__boe	0.944	0.974	**gbp2yswap**	0.942	0.973
gbpusd12m__outr__boe	0.944	0.974	**gbp5yswap**	0.942	0.972
gbboeexri0001	0.943	0.972	**gbp10yswap**	0.942	0.972

Table 4: Table 9 of Kapetanios and Papailias (2021b). Nowcast MAE and RMSE statistics relative to AR(1) across the nowcasting exercise (where data is available) based on models using W indicators.

	MAE	RMSFE		MAE	RMSFE
Rec-AR(1)	1	1	**gbtran0484**	0.569	0.449
gbbank6839	1.199	1.098	**gbtran0483**	0.569	0.445
gbtran0354	0.950	0.973	**gbtran0471**	0.555	0.429
gbtran0347	0.758	0.610	**gbtran0474**	0.559	0.440
gbtran0346	0.755	0.612	**gbtran0473**	0.597	0.449
gbtran0348	0.753	0.611	**gbtran0466**	0.592	0.442
gbbust1597	0.944	0.972	**gbtran0469**	0.568	0.431
gbbust1596	0.949	0.974	**gbtran0468**	0.564	0.434
gblama0103	0.933	0.968	**gbsurv00257**	0.770	0.629
gblama0083	0.937	0.970	**gbsurv00260**	0.745	0.612
gblama0081	0.934	0.967	**gbsurv00264**	0.766	0.643
gblama0076	0.936	0.970	**gbsurv00262**	0.744	0.620
gblama0091	0.929	0.969	**gbpric0101**	0.422	0.292
gblama0101	0.942	0.972	**gbpric0102**	0.429	0.300
gblama0097	0.941	0.972	**gbpric0105**	0.430	0.298
gblama0093	0.934	0.970	**gbpric0124**	0.407	0.287
gblama0082	0.939	0.970	**gbpric0129**	0.409	0.283
gblama0100	0.926	0.967	**gbpric0135**	0.396	0.273
gblama0079	0.939	0.971	**gbpric0142**	0.449	0.302
gblama0080	0.933	0.968	**gbpric0161**	0.440	0.295
gblama0077	0.932	0.969	**gbpric0178**	0.369	0.257
gbtran0460	0.735	0.598	**gbpric0189**	0.403	0.290
gbtran0464	0.752	0.606	**gbpric0212**	0.411	0.281
gbtran0463	0.758	0.607	**gbpric0221**	0.428	0.293
gbtran0454	0.768	0.605	**gbtran0356**	0.947	0.973
gbtran0458	0.784	0.617	**gbtran0367**	0.943	0.973
gbtran0457	0.760	0.607	**gbtran0361**	0.942	0.972
gbtran0476	0.765	0.634	**gbtran0359**	0.944	0.972
gbtran0479	0.776	0.636	**gbtran0360**	0.943	0.973
gbtran0478	0.760	0.625	**gbtran0390**	0.944	0.971
gbtran0481	0.553	0.425			

Table 5: Table 10 of Kapetanios and Papailias (2021b). Nowcast MAE and RMSE statistics relative to AR(1) across the nowcasting exercise (where data is available) based on models using FW indicators.

	Full Sample		Subsample	
	MAE	**RMSE**	**MAE**	**RMSE**
AR(1)	1	1	1	1
AR(P)	0.999	1	0.998	1
Aggregate	0.864	0.913	0.920	0.961
BFW	0.835	0.902	0.778	0.797
Ridge	0.843	0.908	0.822	0.866
Lasso	0.841	0.906	0.788	0.823
EN	0.842	0.906	0.796	0.831
Ad. Lasso, V1	0.837	0.904	0.770	0.800
Ad. Lasso, V2	0.834	0.904	0.752	0.786
PCA(1)	0.860	0.911	0.834	0.909
PCA(A1)	0.836	0.902	0.916	0.957
PCA(A2)	0.858	0.908	0.815	0.867
Random Forests	0.831	0.902	0.932	0.982
MLP	1.068	1.018	1.221	1.165
ELM	0.862	0.909	0.877	0.915
SVR	0.839	0.905	0.835	0.889

Table 6: Table 1 of Kapetanios and Papailias (2022). Nowcast MAE and RMSE using the Online Job Ads. Values correspond to statistics relative to the AR(1) benchmark. The Aggegate is Online Job Ads Index across all industries. The Full Sample case spans from 2020-03-31 to 2021-10-31 (20 obs.) and includes the COVID-19 outbreak (March to May, 2020) in the evaluation. The Subsample case spans from 2020-11-30 to 2021-10-31 (12 obs.).