STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

NOWCASTING SCOTTISH GDP GROWTH

BY

GRANT ALLAN, GARY KOOP, STUART MCINTYRE
AND PAUL SMITH

NO 14-11

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

Nowcasting Scottish GDP Growth

Grant Allan

Gary Koop

Stuart McIntyre

Paul Smith

November 2014

"...some of our statistics are too late to be as useful as they ought to be. We are always, as it were, looking up a train in last year's Bradshaw [timetable]"

Harold MacMillan, UK Chancellor of the Exchequer, 1956.

Abstract: The delays in the release of macroeconomic variables such as GDP mean that policymakers do not know their current values. Thus, nowcasts, which are estimates of current values of macroeconomic variables, are becoming increasingly popular. This paper takes up the challenge of nowcasting Scottish GDP growth. Nowcasting in Scotland, currently a government office region within the United Kingdom, is complicated due to data limitations. For instance, key nowcast predictors such as industrial production are unavailable. Accordingly, we use data on some non-traditional variables and investigate whether UK aggregates can help nowcast Scottish GDP growth. Such data limitations are shared by many other sub-national regions, so we hope this paper can provide lessons for other regions interested in developing nowcasting models.

Keywords: nowcasting, mixed frequency data, regional macroeconomics

1 Introduction

For some purposes, policymakers are interested in future values of macroeconomic variables and, thus, forecasting is an important activity. However, for other purposes, policymakers are interested in the values of macroeconomic variables *now*. For many variables (e.g. asset and commodity prices), obtaining current values of variables is trivial. But for other variables, data must be collected and processed before release and, thus, the policymaker does not know their current values, in some cases for a significant period of time. While timeliness can certainly be an issue at the national level, it is especially acute and problematic for sub-national data and sub-national policy making.

A good example of this difficulty is evident in Scotland where the initial estimate of Scottish GDP[1] for the second quarter of 2014 was released on 19 October, 2014 (and even this initial estimate is liable to be revised in upcoming months). Thus, the policymaker in 2014Q2 did not know the current value of GDP when making decisions and would not find out what it was until 15 weeks after the end of the quarter. Such concerns motivate interest in the growing field of *nowcasting*: providing current estimates of key macroeconomic variables such as GDP. Nowcasts of major macroeconomic aggregates such as GDP are currently produced for many countries. For instance, the major online nowcasting service (www.now-casting.com) produces nowcasts for the major OECD countries as well as Brazil and China.

[1]In fact, it is gross value added which we nowcast, but we use the term GDP here for consistency with the national literature on nowcasting. GVA is one component of GDP. The Office of National Statistics describe the relationship between GVA and GDP as follows: "GVA (at current basic prices; available by industry only) plus taxes on products (available at whole economy level only) less subsidies on products (available at whole economy level only) equals GDP (at current market prices; available at whole economy level only)"`http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/gva/relationship-gva-and-gdp/gross-value-added-and-gross-domestic-product.html`. GVA, and not GDP, is what is released for the UK Government Office Regions.

There are no nowcasts for Scottish GDP growth or any other sub-national region that we are aware of. Generating nowcasts at the sub-national level raises its own particular issues. These include the reduced range of data series collected and data being released in a less timely manner that at the national level.

The purpose of this paper is to develop nowcasting models for Scotland and evaluate their performance. Using data from 1998[2] to 2014, we nowcast the growth in Scottish GDP in pseudo real-time. That is, we provide nowcasts at each point in time (say time τ) using the data available at time τ and compare the nowcasts to the actual values for GDP growth in time τ (which would not have been known until much later than time τ).

Over the upcoming years, we will nowcast in real time (as opposed to pseudo real time) and see how close our nowcasts are to the actual outcomes. If things go well, our goal is to provide regularly updated nowcasts on the Fraser of Allander Institute's website and add nowcasts to the set of forecasts produced in the Fraser of Allander Institute Economic Commentary.

In this paper, we describe the methods we use to produce nowcasts and carry out the pseudo real time nowcasting exercise. To achieve the former, this paper begins by surveying the existing methods used by nowcasters. Subsequently, we describe the distinctive challenges which occur when nowcasting in Scotland. These include the short time span for which data is available, the lack of many key variables commonly used with other nowcasts and the greater time delays in the release of data. We then discuss how we construct nowcasts in light of these challenges. The final part of this paper contains the pseudo real time forecasting exercise.

[2]Before 1998, quarterly Scottish data is unavailable.

2 Nowcasting: An Overview

2.1 Summary of the Issues

Several excellent surveys of nowcasting (or closely related topics such as short-term forecasting) have recently appeared. These include Banbura, Giannone and Reichlin (2011), Banbura, Giannone, Modugno and Reichlin (2013), Camacho, Perez-Quiros and Poncela (2013) and Foroni and Marcellino (2013). There are a number of different, related, approaches to nowcasting in the literature; which we briefly summarise here, but the interested reader in search of a fuller treatment of these methods (including bridge equations, factor models, mixed frequency VARs and MIDAS[3])[4] should refer to the cited papers for more details. Here we outline the general concepts underlying nowcasting before describing the particular set of methods that we use in this paper.

At the most general level, nowcasting methods (like many forecasting methods) seek to find explanatory variables/predictors which are useful for predicting the dependent variable to be nowcast. Nowcasts are based on an econometric model linking the predictors to the dependent variable. For GDP growth there are a myriad of such predictors. For instance, Banbura, Giannone, Modugno and Reichlin (2013) use 23 predictors in their nowcasting model of US GDP growth including both "hard" variables such as industrial production and "soft" variables such as surveys of businesses.

Important econometric issues arise when nowcasting due to the fact that nowcasters want their predictors to be as timely as possible. For instance, when nowcasting 2014Q2 GDP growth, having a predictor for which data becomes available in May or June, 2014 is very useful. A predictor which is not available until October 2014 (when the initial estimate of 2014Q2 GDP is

[3]MIDAS is an acronym for Mixed Data Sampling.

[4]New approaches to nowcasting which do not quite fit into these categories include Carriero, Clark and Massimiliano (2012) and Mazzi, Mitchell and Montana (2013).

released) is virtually useless. Furthermore, nowcasters typically update their nowcasts throughout the quarter as new information becomes available. The desire for timeliness and frequent updating of nowcasts leads to two econometric issues which are treated in different ways by the different nowcasting approaches. These are: i) the dependent and explanatory variables have different frequencies and ii) the nowcasters' data set typically has a "ragged edge".

The mixed frequency issue arises since GDP is observed quarterly whereas many potential predictors for GDP (e.g. industrial production, some labour force statistics and Purchasing Managers' Indices, PMIs) are available monthly. In this paper, we will use MIDAS methods (described below) to address the mixed frequency issue, but several other methods exist (see, in particular, Foroni and Marcellino, 2013 for a survey of the various econometric methods used with mixed frequency data).

The ragged edge problem refers to the fact that the variables in the nowcaster's data set typically have different release dates and, thus, at the end of the sample missing observations will exist for some of them. Consider, for instance, nowcasting 2014Q2 Scottish GDP growth at the end of July 2014. By this time, the value of June's Bank of Scotland's PMI was released and the nowcaster would wish to update the 2014Q2 nowcast. But data on UK exports and imports for June will not be released until mid August. Again, there are several ways of addressing this ragged edge problem, but we will address them using MIDAS methods.

A final data issue worth noting, of relevance to both forecasters and nowcasters working in real time, is that GDP is revised over time as new information is collected, leading to different vintages of data (i.e. the first vintage of GDP data is the initial release 15 weeks after the end of the quarter, the second vintage follows a quarter after that, etc.). For instance, the initial estimate of 2012 Q3 Scottish GDP growth was 0.6%, but three months later this was revised to 0.4%, later revisions occurred such that at present GDP growth in this quarter is estimated as 0.1%. In the present paper, our pseudo real-time nowcasting exercise does not address this issue since we use final vintage data. However, when we do our future nowcasting work, we will always use the most recent version of each of our variables.

2.2 A brief overview of competing approaches

This section provides a very brief overview of three competing methods of overcoming the mixed frequency and jagged edge issues in nowcasting which were discussed in the previous section. A reader in search of more details should refer to the survey papers cited at the beginning of the preceding sub-section.

Historically, bridge equation methods have been the most popular. As an example of how they work, consider Smith (2013) who uses univariate autoregressive forecasting models to 'fill in the gaps' caused by the jagged edge, before applying a bridging equation approach to transform the higher frequency data into explanatory variables to be used in a regression involving the lower frequency dependent variable being nowcast. To be precise, the bridging process in Smith (2013) involves taking an average of the higher frequency observations to produce a lower frequency variable. This average is then used to explain the lower frequency dependent variable of interest. An example of this would be averaging across the three monthly values of the PMI in a quarter and then using this average to nowcast quarterly GDP.

This is a simple and easily implemented approach, but at the cost of losing potentially useful information. By taking a simple average, recent and past values are weighted equally (possibly an undesirable feature) and the impact of a single good (or bad) month in the quarter can be ameliorated. While bridging approaches provide an intuitive and straightforward solution to the difficulties posed by mixed frequency data, in recent years more complex models have been developed to improve nowcast accuracy.

Factor models are a major alternative to bridging equations. Factor methods take a large number of explanatory variables and extract a small number of variables called factors which contain most of the information in the explanatory variables. These factors can then be used in a regression. The methods developed in Giannone et al (2008) allow for the factors to be at a higher frequency than the lower frequency variable being nowcast. Thus, this approach also deals with both of the issues identified earlier.

The third main alternative, and the one we use in this paper, is MIDAS. This is also regression-based method, initially introduced by Ghysels (2004). We will explain MIDAS in more detail in the next section. But, before doing so, we note here that it addresses both of the issues raised above. Under MIDAS, no forecasting of missing values is necessary (so the first difficulty noted above disappears) and the models are set up (as the name suggests) to deal with mixed frequency data (addressing the second issue raised above).

Within the MIDAS approach there are a number of different specifications that are possible, and a literature has built up which walks the reader through these. It is worth noting that much of the MIDAS literature is focussed on using very high frequency explanatory variables (e.g. daily financial data) to forecast monthly or quarterly data. In such cases, if the researcher uses each daily observation as an explanatory variable in a regression involving a monthly dependent variable, then the number of explanatory variables can be enormous. MIDAS surmounts the problems that result by placing restrictions on the coefficients. The different MIDAS specifications arise from the nature of these restrictions. In the case of a large frequency mis-match (e.g. daily explanatory variables and monthly dependent variables) the gains from MIDAS can be substantial. However, even for smaller frequency mis-matches (e.g. monthly explanatory variables and quarterly dependent variables), Foroni (2012) and Ghysels (2014) both argue that there are advantages to using an unrestricted MIDAS (U-MIDAS) approach. In particular, Foroni (2012) show that U-MIDAS performs better than other MIDAS specifications for this type of mixed data sampling.

Previously we have cited survey papers which discuss the practical use of MIDAS methods. For the reader interested in the econometric theory, Andreou, Ghysels and Kourtellos (2013) is a recent survey. Much pioneering work in the field was done by Eric Ghysels in several papers including Ghysels, Sinko and Valkanov (2007). Bai, Ghysels and Wright (2013) shows the close relationship between MIDAS methods and the factor methods used by nowcasters such as Giannone, Reichlin and Small (2008). The next section provides a more in-depth technical treatment and explanation of the MIDAS methods that we use in this paper.

2.3 MIDAS

GDP data (and some of the predictors we use) are available at quarterly intervals, whereas most of our predictors are available at monthly intervals.[5] MIDAS methods were developed to deal with such situations. To explain how MIDAS works in more detail, we introduce notation where y_{t_Q} is the quarterly variable we are interested in nowcasting (in our case GDP growth) for $t_Q = 1, .., T_Q$ quarters and X_{t_M} is the monthly predictor for $t_M = 1, .., T_M$. Note that the first time index counts at the quarterly frequency and the second at a monthly frequency and $T_M = 3T_Q$. One way of over-coming the frequency mismatch between dependent variable and predictor would be to transform the monthly explanatory variables to a quarterly frequency, i.e. create

$$X_{t_Q}^{Q} = \frac{X_{3\left(t_Q-1\right)+1} + X_{3\left(t_Q-1\right)+2} + X_{3\left(t_Q-1\right)+3}}{3}$$

[5]We plan on providing monthly updates of our nowcasts and, hence, work at this frequency. Some nowcasters work at the daily frequency, providing daily nowcasts so that, e.g., on 13 January, 2014, when the value of December's Bank of Scotland's PMI was released, the nowcast of GDP could be updated on 13 or 14 January. Given we are updating nowcasts montly, we would use this PMI release in our 1 February nowcast and treat all of our predictors as though they are end of month values.

and then use a standard regression model:

$$y_{t_Q} = \alpha + \beta X_{t_Q}^{Q} + \varepsilon_t.$$

Such an approach, which underlies bridge sampling methods, can be thought of as taking an average of recent values of the monthly variables and using the result as a predictor. An example of this would be creating a quarterly explanatory variable by averaging across the three monthly values of the Purchasing Managers' Index (PMI) and then using this average to nowcast quarterly GDP. This is a simple and easily implemented approach, but at the cost of losing potentially useful information, for instance by ameliorating the impact of a single good (or bad) month in the quarter and weighting more distant information equally to the most recent.

One thing that can be done to address some of the criticisms of bridge equation modelling is to allow for unequal weights so as to have more recent data receive more weight than data from the more distant past.[6] This suggests working with a regression model of the form:

$$y_{t_Q} = \alpha + \beta \sum_{j=0}^{p_{M-1}} w_j X_{3t_Q - j} + \varepsilon_t, \tag{1}$$

where the weights, w_j, sum to one and depend on unknown parameters which are estimated from the data and p_M are the number of monthly lags. This is a MIDAS model. We will not discuss estimation of such a model other than to note that nonlinear least squares can be used.

Given the importance of timing issues in nowcasting, we elaborate on what exactly MIDAS nowcasts involve for the Scottish case. Note that, for any quarter's GDP growth, there are five nowcasts of interest. Consider, for instance, GDP growth in 2014Q2. During this quarter, we do not know its

[6] Andreou, Ghysels and Kourtellos (2013) also show some econometric problems of the equal weight specification used in bridge sampling, including the potential for asymptotic bias or inefficiency.

value and, thus, nowcasts made on 1 May and 1 June, 2014 will be needed. But the initial estimate of GDP growth in 2014Q2 will not be released until mid October and, hence, nowcasts[7] made on 1 July , 1 August and 1 September, 2014 are also required. We do not produce nowcasts on 1 October, 2014 since the initial release will occur shortly, but this can be done if desired. These nowcasts can be produced using a slight alteration to (1) by introducing an index h to denote these five nowcasts through the following specification:

$$y_{t_Q} = \alpha + \beta \sum_{j=0}^{p_M - 1} w_j X_{3t_Q - j - h} + \varepsilon_t. \tag{2}$$

To understand the properties of this specification, we will continue using 2014Q2 as an example. If $h = 0$, then the explanatory variables are all dated June 2014 (or earlier). Given a one month delay in releasing data on the explanatory variables, this data would be available by the end of July 2014. Thus, nowcasts of 2014Q2 GDP growth made on 1 August, 2014 can be obtained by setting $h = 0$. By similar reasoning, setting $h = 1$ produces nowcasts using explanatory variables dated May 2014 which come available during June. This is what we would want when making nowcasts on 1 July, 2014, etc. We can even set h to be a negative number. This is called MIDAS with leads. Setting $h = -1, 0, 1, 2, 3$ will produce the five nowcasts referred to at the beginning of this paragraph.

Another issue that we need to address is the role played by lags of the dependent variable. That is, it is common, even after controlling for explanatory variables, for macroeconomic aggregates such as GDP growth to exhibit autocorrelation. Thus, including lags of the dependent variable has the potential to improve nowcast performance. This can easily be accommodated

[7] One could call these "backcasts" instead of nowcasts.

by adding lags of the dependent variable to the MIDAS model:

$$y_{t_Q} = \alpha + \sum_{j=1}^{q} \rho_j y_{t_Q - j} + \beta \sum_{j=0}^{p_M - 1} w_j X_{3t_Q - j - h} + \varepsilon_t. \tag{3}$$

This is what we do in this paper. However, we have to be careful since Scottish GDP figures are released with a 15 week delay. Consider, again, the five nowcasts of 2014Q2 GDP growth obtained by setting $h = -1, 0, 1, 2, 3$ in (3). For the first three (made on 1 May, 1 June and 1 July 2014), the initial release of 2014Q1 GDP figures would not be available. Thus, we would not yet know what $y_{t_Q - 1}$ is and it cannot be used as a predictor. Accordingly, the lags must begin with $y_{t_Q - 2}$ (or, equivalently, we must set $\rho_1 = 0$ in (3) for the first three out of the five nowcasts.

The following table summarizes the timing of the data[8] for each of our nowcasts.

Table 1: Timing of Data and Nowcast Releases

h	Month data relates to:	e.g. for Q2 GDP	Nowcast released on 1st day of	e.g. for Q2 GDP
-1	First month of following quarter	July	Third month of following quarter	September
0	Third month of quarter	June	Second month of following quarter	August
1	Second month of quarter	May	First month of following quarter	July
2	First month of quarter	April	Third month of quarter	June
3	Last month of preceding quarter	March	Second month of quarter	May

MIDAS is commonly used with financial data where daily data is used to forecast monthly or quarterly variables. In such a case, parsimony is a major concern since there can be so many weights to estimate. That is, instead of our three months in a quarter (leading to three weights in the case where we lag variables up to a quarter), there are 31 days in a month and 122 days in a quarter. This has led to wide range of distributed lag specifications being proposed. However, for our relatively parsimonious case, we do not

[8]This timing is relevant for monthly variables which are released within a month. As noted in the appendix, a small number of our variables are released with a delay of more than a month and, for these, the timing convention is adjusted appropriately.

consider such specifications but, instead, work with the unrestricted MIDAS specification of Foroni, Marcellino and Schumacher (2013). The interested reader is referred to, e.g., Andreou, Ghysels and Kourtellos (2013) for a discussion of other specifications.

We also need to extend the basic MIDAS model given in (1) to account for the fact that we do not have a single explanatory variable, X_t, but rather 40. Given that our data span is very short, beginning in 1998Q1, simply including all of them would lead to a very non-parsimonious model. There are two main ways to get around this problem, the first of these is through use of the factor MIDAS model and the second is through model averaging. In this paper, we will use model averaging. Instead of working with one single MIDAS model, we use 40 models each of which uses one of the predictors. We use as our nowcast a weighted average of all of the individual nowcasts. A similar strategy is used in Mazzi, Mitchell and Montana (2013).

We consider various weighting schemes. In particular, if we have N models and p_{it} is the weight attached to model i at time t for $i = 1, .., N$, then we consider:

- Equal weights:

$$p_{it} = \frac{1}{N}$$

- BIC based weights:

$$p_{it} = \frac{\exp\left(BIC_{it}\right)}{\sum_{j=1}^{N} \exp\left(BIC_{jt}\right)}$$

- MSFE based weights:

$$p_{it} = \frac{\left(MSFE_{it}\right)^{-1}}{\sum_{j=1}^{N} \left(MSFE_{jt}\right)^{-1}}.$$

In these weights BIC_{it} stands for Bayesian information criterion of model i at time t and MSFE is mean squared forecast error. Both are calculated using the data available at time t. BIC is a popular model selection device and BIC based weights put more weight on models which have scored well on this metric. MSFE is a measure of forecast performance, so using MSFE based weights results in more weight being put on models which have forecast well in the past.

3 Nowcasting in Scotland

For the reasons outlined in Section 2.1, the goal of the nowcaster is to find variables: i) which help predict GDP growth, ii) which are timely and iii) which are updated frequently (e.g. at a monthly frequency). Typically, this has lead researchers to use a variety of hard and soft predictors. Industrial production (and its components) is commonly used as one of the main hard variables. Variables reflecting the labour market, employment, sales and consumption are also popular hard variables. Soft variables are based on surveys of various sorts (i.e. surveys of business, consumers, etc.). However, many of these (and, in particular, many of the hard variables) are unavailable for Scotland. This is a problem facing many regions. Accordingly, we have collected a data set of predictors containing some traditional nowcasting predictors, but also some non-traditional ones. In addition, we include some conventional hard nowcasting variables for the UK as a whole to investigate whether these have enough explanatory power to help improve nowcasts of Scottish GDP growth. Furthermore, it is possible that there is information in other UK regions which our nowcasts can exploit. For this reason, some of our predictors are for the other regions of the UK.

The specific variables that we have collected and used are briefly described below, in addition we explain why these have been chosen. Further details on each variable (including definitions, timeliness, sources, transformations and release dates) are given in the Data Appendix.

Some of these are available for Scotland alone, while others are for other regions of the UK or the UK as a whole. For the reasons described above we have taken the stance that data which may be a useful predictor of Scottish economic activity should be included, even if the data relates to a wider geographic grouping, such as the UK. Additionally, many of the data used in nowcasting at a national level are simply either not available for regions, or are only available for the regions with a greater lag.

We should note that the quarterly Gross Value Added (GVA) growth index for Scotland was first produced for 1999Q1 and (at the time of writing) runs to 2014Q2 (produced on the 19th of October 2014). This is the index of economic activity for which we are seeking to nowcast. We are especially keen to include variables which would be available over the same period, and have not included some series that are available only for part of this time period. Quarterly variables included therefore run from 1998Q1 to 2014Q2, while monthly variables run from January 1998 to September 2014 (although as the Data Appendix explains, some of the monthly variables are released with a longer delays and so are only currently available for earlier months).

In all, we employ a total of forty predictors, across a range of hard and soft indicators. We begin by describing the (thirty-one) monthly variables available. We have twelve Purchasing Managers Index (PMI) variables for the government office regions of the UK, including Scotland[9]. These are a widely used– including by the Bank of England – tracker of economic activity, also produced for nations and national groupings outside the UK (such as the Eurozone). Recent evidence suggests that the UK PMI measure has tracked well with recent UK economic performance, suggesting they may also be useful for nowcasting Scottish performance. Additionally, their short publication lag – produced 10 days after the end of the month – merits their inclusion in our analysis. We include PMI measures for other regions of the UK (PMILON, PMISE, PMISW, PMIEAST, PMIWALES, PMIWMID,

[9]This series are reported by Lloyds Banking Group, and are known as: Bank of Scotland PMI Scotland.

PMIEMID, PMIYORK, PMINE, PMINW AND PMINI) in addition to Scotland (PMISCOT) firstly as these data are available, and secondly as the rest of the UK is the primary and principle destination for Scottish exports. Additionally, we include three variables which are PMIs for the UK, Eurozone and world (PMIUK, PMIEZ and PMIWORLD).

We include eight variables related to VAT receipts for the UK. Such figures are likely to track with the level of spending, and, with consumption spending a significant portion of GDP, it is useful to include these measures. Five variables (VATPAY, VATREPAY, VATRCPT, IMPVAT and TOTALVAT) will track such receipts on a monthly basis. A further three variables relate to the number of firms registered for VAT purposes (NEWVATREG, VATDEREG and TRADEPOP).

There are a further ten monthly variables. The paucity of regional data means that only three soft indicators – GFKCC, a measure of Scottish consumer confidence, and BOSJOBS_PL and BOSJOB_ST - relate to Scottish activity specifically. Consumer confidence measures are widely used as nowcasting predictors as they give an indication about the "direction of travel" for consumption spending and so are often good predictors of sales revenues which are critical for economic activity in service-dominated economies. The two other measures mentioned above are monthly measures of the labour market in Scotland for job placements and staff demand, respectively. As such these may both be useful predictors of employment growth and economic activity. The only other "hard" Scottish data series comes as the (claimant count) unemployment rate for Scotland (UNEM).

As UK-wide hard indicators we use industrial production (UKIP) and the services-output index (IOSG). IOSG might be a good predictor as this shows the movements in gross value added for the service industries, which overall account for around 78 per cent of UK GDP. UKCPI shows the rate of inflation for the UK as a whole which is also typically included in nowcasting analysis. Two indicators refer to the level of exports (UKEXPORTS) and imports (UKIMPORTS) for the UK economy as a whole. Both these series could be useful predictors, in particular as Scotland is likely to contribute a greater share of UK exports than its share in UK GDP, through specific products such as whisky and refined petroleum. For this latter product, we also additionally include a (UK) measure of total throughput of refined petroleum – UKREFINE.

Turning to the (seven) quarterly variables, each of these specifically relate to Scotland. Firstly, we have as hard indicators the Scottish government-produced Retail Sales Index for Scotland (RSI) and HMRC data on total Scottish exports and imports to the rest of the world (EXP and IMP respectively).[10] The RSI data series is likely to be a useful predictor of retail and consumer spending, while both trade variables may be important for the strength of external (and domestic) demand and Scottish economic activity. We include four survey variables drawn from two respected quarterly surveys of the Scottish economy: the Scottish Business Monitor and Scottish Chambers Business Survey. From the former we use a measure of output by Scottish firms (SBM). From the latter, we use variables which measure the volume of business by firms in the manufacturing, construction and tourism sectors (SCBSMAN, SCBSCON and SCBSTOUR, respectively).

4 Nowcasting in Pseudo Real-Time

The following tables contain MSFE's and sums of log predictive likelihoods for the nowcasts for the five different months (labelled $h = -1, 0, 1, 2, 3$ as described above). MSFEs are a common metric to evaluate the quality of point forecasts with lower values indicating better performance. Predictive likelihoods are a common metric for evaluating the quality of the entire predictive distribution with higher values indicating better nowcast performance.

[10] There are only annual surveys of total exports from Scotland, while the quarterly survey of exports produced by the Scottish government covers only manufacturing exports, which constitute a declining share of total exports.

A predictive likelihood is the predictive density produced by a nowcasting model, evaluated at the actual outcome. Our MIDAS methods produce a predictive mean (the point nowcast) and a predictive standard deviation. We use a Normal approximation to the predictive density. Our nowcasts are recursive, i.e. each nowcast is calculated using data from the beginning of the sample to the time the nowcast is made. We experimented with the use of rolling methods, but these performed slightly worse than recursive methods.

Each nowcast is produced using the specification given in (3) with two lags of the dependent variable and a single explanatory variable. There are 40 such nowcasts for our 40 explanatory variables. We also present nowcasts which average over all models. Our results use two lags of the dependent variable ($q = 2$) and, thus, all our models add to the AR(2) process commonly used with GDP growth. For the monthly explanatory variables MIDAS is done over the three quarters in the month ($p_M = 3$). For the quarterly explanatory variables we use a single lag which is the most recent value of the variable which is available at the time the nowcast is made. We evaluate the nowcasts beginning with the first month of 2005.

Tables 2 and 3 presents the MSFEs and sums of log predictive likelihoods, respectively. The row of Table 2 labelled "No change nowcasts" contains MSFEs for a benchmark we hope to beat. It simply uses as the nowcast the most recent value of GDP growth that is available. Given delays in release of data, this will be GDP growth two quarters ago for the three months of the quarter ($h = 3, 2, 1$) and last quarter's GDP growth for the first two months of the following quarter ($h = 0, -1$).

MSFEs and sums of log predictive likelihoods are telling similar stories and there are two main stories that emerge. First, we are finding that what we might be called current quarter nowcasts ($h = 1, 2, 3$, e.g. nowcasts for 2014Q2 made on 1 July or earlier) are substantially better than no change nowcasts. Results for what can be called following quarter nowcasts ($h = -1, 0$, e.g. nowcasts for 2014Q2 made on 1 August and 1 September) are less encouraging. Second, model averaging is a great help in improving nowcast performance. We elaborate on these stories and offer some additional details in the following material.

With current quarter nowcasts, averaging over all models is producing MSFEs which tend to be much lower than individual nowcasts using a particular predictor. Furthermore, MSFEs are being reduced by roughly a quarter relative to no change nowcasts. But most of these gains are driven by a small number of our predictors. This illustrates an advantage of our approach: we can include a large number of potential predictors, most of which provide little explanatory power, and let the econometric methodology decide which ones should be used to form the nowcasts. In our case, it is sometimes the case that non-obvious variables receive a lot of weight. For instance, PMI for Northern Ireland is the best predictor for Scottish GDP growth for several nowcasts. A careful examination of the data reveals the reason: Northern Ireland's PMI fell much further after the financial crisis than PMI for the other regions. This improved the nowcasts after the financial crisis when actual GDP growth fell dramatically. In general, some of the PMI variables do tend to be good predictors. Among the PMI variables, one would expect Scottish PMI to be the best predictor for Scottish GDP growth. It does often nowcast well. However, as noted, at some nowcast horizons Northern Ireland's PMI is a better predictor. And for $h = 0$ (i.e. nowcasts released on the first day of the second month of the following quarter (e.g. on the 1st August) using data from the third month of the previous quarter (so e.g. June)), PMI for the UK as a whole is a very good predictor.

Among the remaining soft variables (which often nowcast better for the current quarter), GFKCC (a survey of consumer confidence) tends to nowcast well. Variables from the Bank of Scotland's Report on Jobs, are also often reasonably good predictors.

Some of the hard variables nowcast well in the following quarter. Given that hard variables are often released more slowly than soft variables this is not surprising. For instance, the index of services for the UK as a whole (IOSG) is released with nearly a two month delay, but is often an excellent nowcasting variable. For our final nowcast before the new GDP data release ($h = -1$) it is the best predictor.

Most of the other predictors rarely nowcast well and obtain little weight in most of our averaged nowcasts. But most of them at least occasionally make an impact. For instance, most of our variables relating to VAT do not nowcast well, but for one nowcast horizon ($h = 0$) new VAT registrations is a good predictor. Our methods can automatically adjust to such findings, giving substantial weight to the nowcasting model with NEWVATREG when $h = 0$, and giving very little weight to this model for other values of h.

Table 2: MSFEs for nowcasts

	$h=3$	$h=2$	$h=1$	$h=0$	$h=-1$
No change nowcasts	0.611	0.611	0.611	0.405	0.405
Ave. MSFE weights	0.510	0.515	0.507	0.442	0.439
Ave. BIC weights	0.517	0.521	0.511	0.445	0.445
Ave. equal weights	0.517	0.521	0.512	0.445	0.445
PMISCOT	0.535	0.520	0.470	0.482	0.565
PMILON	0.641	0.629	0.655	0.564	0.596
PMISE	0.633	0.608	0.538	0.527	0.548
PMISW	0.576	0.590	0.605	0.512	0.555
PMIEAST	0.557	0.546	0.520	0.531	0.491
PMIWALES	0.631	0.676	0.622	0.574	0.571
PMIWMID	0.711	0.634	0.613	0.545	0.499
PMIEMID	0.611	0.639	0.619	0.507	0.486
PMIYORK	0.687	0.657	0.595	0.550	0.556
PMINE	0.675	0.650	0.681	0.511	0.545
PMINW	0.668	0.632	0.607	0.578	0.542
PMINI	0.463	0.457	0.588	0.422	0.366
VATPAY	0.705	0.702	0.671	0.570	0.586
VATREPAY	0.734	0.704	0.661	0.532	0.593
VATRCPT	0.639	0.694	0.668	0.549	0.498
IMPVAT	0.575	0.683	0.731	0.573	0.546
TOTALVAT	0.611	0.711	0.706	0.540	0.493
NEWVATREG	0.652	0.612	0.573	0.456	0.547
VATDEREG	0.656	0.604	0.559	0.522	0.555
TRADPOP	0.685	0.555	0.530	0.545	0.531
UKIP	0.641	0.672	0.651	0.484	0.532
UKCPI	0.683	0.646	0.588	0.520	0.529
UNEMP	0.618	0.674	0.682	0.541	0.462

IOSG	0.510	0.549	0.570	0.444	0.385
GFKCC	0.498	0.516	0.537	0.660	0.636
UKREFINE	0.544	0.647	0.598	0.475	0.504
UKEXPORTS	0.651	0.648	0.624	0.557	0.558
UKIMPORTS	0.616	0.608	0.594	0.537	0.553
RSI	0.592	0.592	0.592	0.488	0.488
EXP	0.551	0.551	0.551	0.488	0.488
IMP	0.517	0.517	0.517	0.452	0.452
SBM	0.541	0.551	0.551	0.415	0.426
SCBSMAN	0.573	0.573	0.573	0.480	0.480
SCBSCON	0.519	0.519	0.519	0.472	0.472
SCBSTOUR	0.559	0.559	0.559	0.507	0.507
PMIUK	0.529	0.536	0.502	0.432	0.427
PMIEZ	0.562	0.571	0.558	0.468	0.441
PMIWORLD	0.530	0.551	0.518	0.454	0.453
BOSJOBS_PL	0.538	0.575	0.574	0.444	0.494
BOSJOBS_ST	0.569	0.555	0.528	0.464	0.475

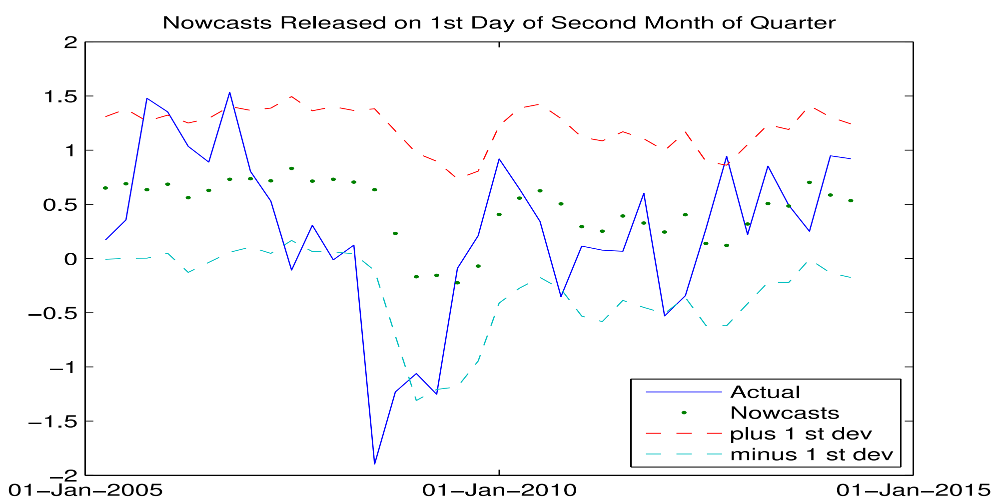
Table 3: Sums of log Predictive Likelihoods for nowcasts

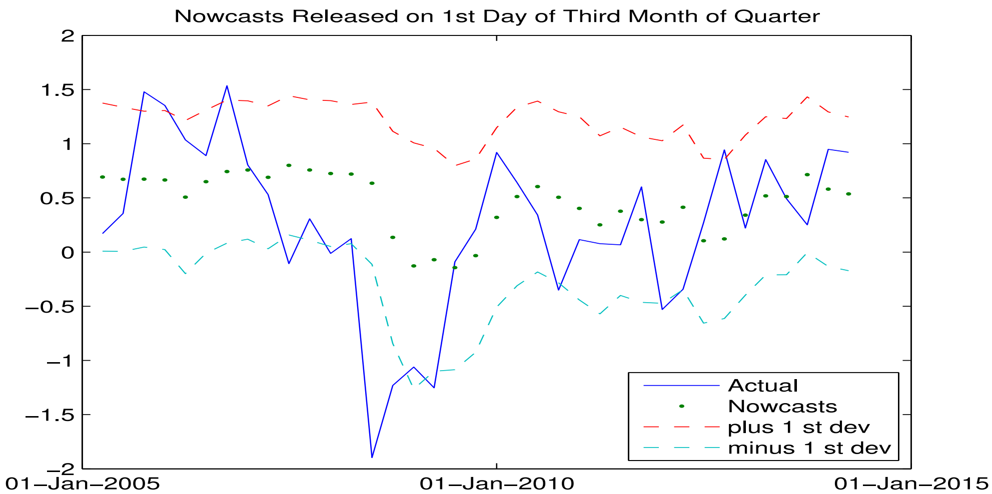
	$h=3$	$h=2$	$h=1$	$h=0$	$h=-1$
Ave. MSFE weights	130.73	130.64	130.09	132.74	129.92
Ave. BIC weights	130.57	130.48	129.99	132.67	129.72
Ave. equal weights	130.56	130.47	129.98	132.66	129.70
PMISCOT	129.16	130.45	132.24	131.88	127.13
PMILON	125.63	125.78	125.50	128.66	126.83
PMISE	126.55	126.07	127.65	131.60	128.80
PMISW	127.90	128.10	125.98	128.68	127.87
PMIEAST	129.70	130.02	127.92	130.00	127.37
PMIWALES	129.66	127.60	127.75	130.86	128.33
PMIWMID	127.21	127.05	123.76	129.04	128.46
PMIEMID	124.19	126.31	127.03	130.26	130.53
PMIYORK	123.13	123.27	126.56	128.40	126.04
PMINE	122.30	125.92	125.39	131.42	130.73
PMINW	127.60	127.56	125.40	130.50	129.35
PMINI	132.86	133.31	128.44	135.67	130.85
VATPAY	124.40	123.35	123.48	125.83	126.13
VATREPAY	120.65	120.62	122.54	127.22	127.19
VATRCPT	124.27	123.63	124.08	128.03	129.53
IMPVAT	125.62	122.39	119.86	125.48	126.84
TOTALVAT	125.42	124.13	124.41	129.51	129.36
NEWVATREG	124.65	126.05	127.88	132.03	129.82
VATDEREG	126.17	128.15	126.85	128.72	127.76
TRADPOP	126.37	129.67	129.60	129.25	130.79
UKIP	124.18	124.96	125.57	130.51	131.03
UKCPI	126.54	126.53	127.45	130.55	130.89
UNEMP	126.83	126.63	125.60	131.58	133.85
IOSG	131.02	129.28	129.34	132.36	135.53

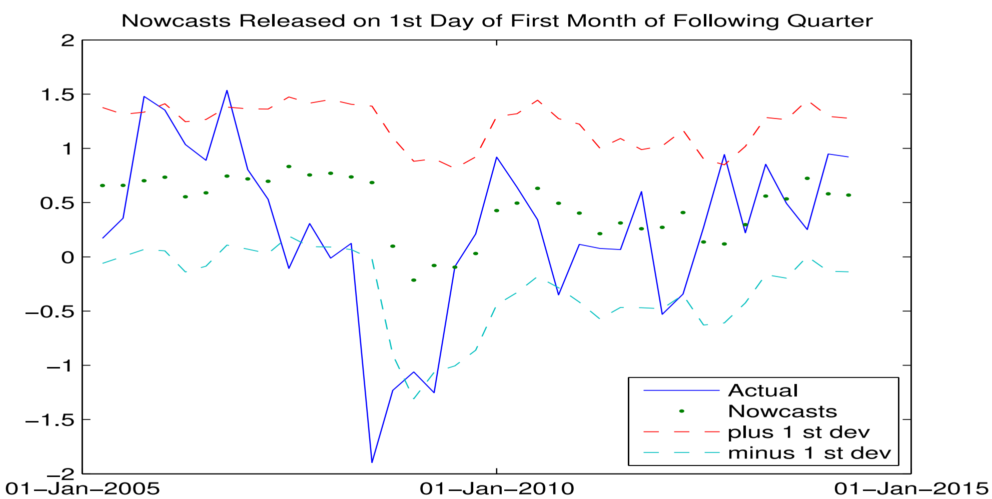
GFKCC	131.65	131.24	131.39	127.73	128.32
UKREFINE	127.34	124.06	125.88	130.23	128.64
UKEXPORTS	124.97	123.95	123.99	128.12	127.38
UKIMPORTS	121.55	122.18	123.60	127.76	127.03
RSI	125.50	125.50	125.50	129.89	129.89
EXP	127.05	127.05	127.05	130.25	130.25
IMP	127.56	127.56	127.56	131.76	131.77
SBM	129.40	127.21	127.21	135.61	131.09
SCBSMAN	126.47	126.47	126.47	130.07	130.07
SCBSCON	128.29	128.29	128.29	131.15	131.15
SCBSTOUR	127.74	127.74	127.74	130.98	130.98
PMIUK	129.04	128.91	129.72	133.42	130.08
PMIEZ	126.52	126.43	127.17	130.83	129.80
PMIWORLD	128.00	127.33	129.09	131.73	129.87
BOSJOBS_PL	127.21	126.15	126.35	131.56	129.63
BOSJOBS_ST	125.94	126.42	127.80	130.73	129.36

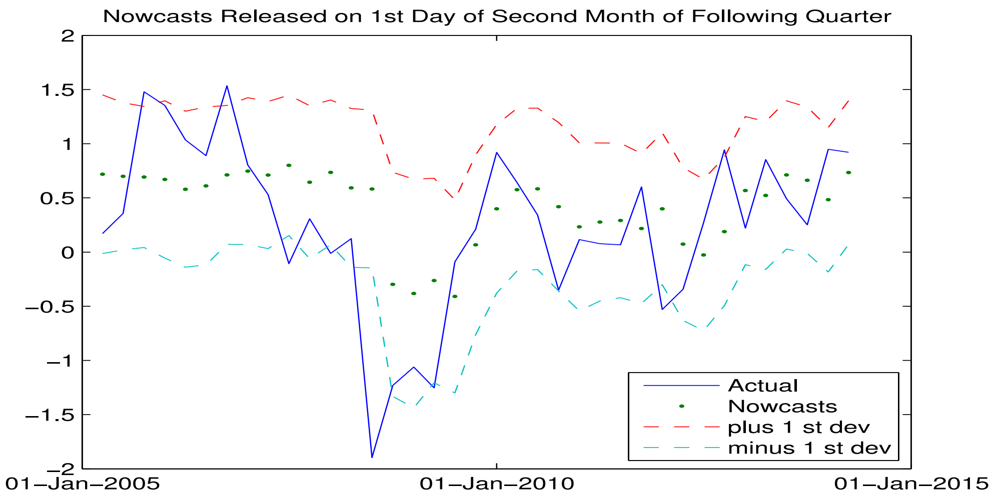
Tables 2 and 3 present forecast metrics averaged over the entire period from 2005 through the end of the sample. To gain insight into how our nowcasts perform over time, Figures 1 through 5 plot nowcasts for our preferred approach (averaged nowcasts using MSFE weights) over time for the five nowcast horizons. On the whole, our nowcasts match the actual outcomes quite well. The Great Recession began in the middle of our nowcast evaluation period. It can be seen that our methods were slightly late in capturing the fall in GDP growth and never quite predicted its magnitude. Perhaps this is unsurprising given the short sample that was being used to estimate the models and the fact that the Great Recession was quite different than anything else seen previously in our data.

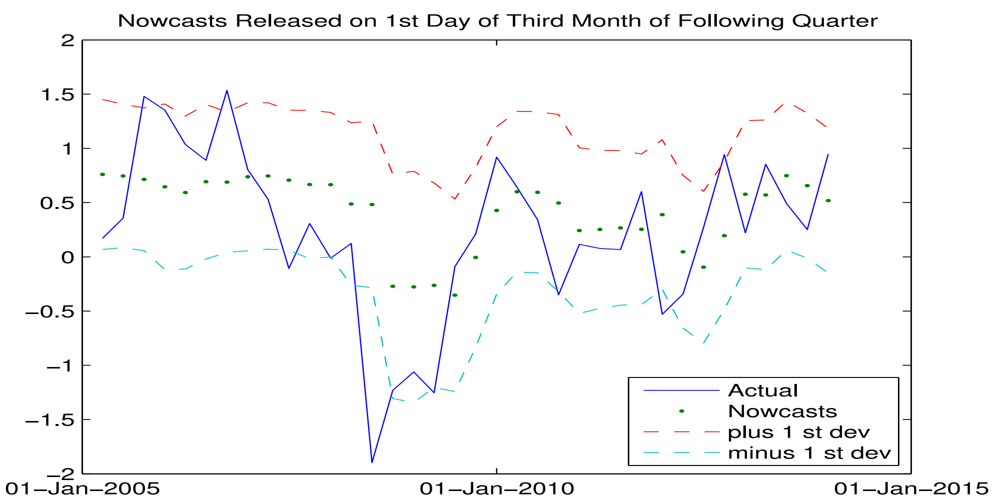
Another pattern is that the nowcasts, as expected, tend to improve over time. For instance, if one examines the stuttering recovery which began in 2010, it can be seen that the first nowcasts we produce tended to be below the eventual realization of GDP growth. However, by the second quarter of the months, the nowcasts were tracking the actual realizations much better.




Nowcasts Released on 1st Day of Third Month of Quarter
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
01−Jan−2005
01−Jan−2010
01−Jan−2015
Actual
Nowcasts
plus 1 st dev
minus 1 st dev


Nowcasts Released on 1st Day of First Month of Following Quarter
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
01−Jan−2005
01−Jan−2010
01−Jan−2015
Actual
Nowcasts
plus 1 st dev
minus 1 st dev





5 Conclusions

In this paper, we have discussed the challenges facing the researcher interested in nowcasting within a sub-national region such as Scotland. These include the longer delays in release of key variables, the lack of data on variables commonly-used to nowcast at the national level and the shortness of the time span for which data is available. To try and overcome these challenges, we have collected a large data set containing a wide variety of variables. We find that, by using MIDAS methods and averaging over results for our many models, we can nowcast fairly successfully, particularly in the quarter being nowcast. Our plan is to use these variables and econometric methods in the future, to nowcast Scottish GDP growth and release monthly updates of the current state of the economy in Scotland.

References

Andreou, E., Ghysels, E. and Kourtellos, A. (2013). “Forecasting with mixed-frequency data,” chapter 8 in the Oxford Handbook of Economic Forecasting, edited by M. Clements and D. Hendry. Oxford University Press: Oxford.

Bai, J., Ghysels, E. and Wright, J. (2013). “State space models and MIDAS regressions,” Econometric Reviews, 32, 779-813.

Banbura, M., Giannone, D., Modugno, M. and Reichlin, L. (2013). “Nowcasting and the real time data flow,” chapter 4 in the Handbook of Economic Forecasting, Vol 2A, edited by G. Elliott and A. Timmermann. Elsevier-North Holland: Amsterdam.

Banbura, M., Giannone, D. and Reichlin, L. (2011). “Nowcasting,” chapter 7 in the Oxford Handbook of Economic Forecasting, edited by M. Clements and D. Hendry. Oxford University Press: Oxford.

Camacho, M., Perez-Quiros, G. and Poncela, P., (2013). “Short-term forecasting for empirical economists: A survey of the recently proposed algorithms,” Bank of Spain Working Paper 1318.

Carriero, A., Clark, T. and Massimiliano, M. (2012). “Real-time nowcasting with a Bayesian mixed frequency model with stochastic volatility,” Federal Reserve Bank of Cleveland Working Paper 12-27.

Foroni, C. and Marcellino, M. (2013). “A survey of econometric methods for mixed frequency data,” Norges Bank Research Working Paper 2013-06.

Foroni, C., Marcellino, M. and Schumacher, C. (2013). “Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials,” Journal of the Royal Statistical Society, Series A, DOI: 10.1111/rssa.12043.

Ghysels, E., Sinko, A. and Valkanov, R. (2007). “MIDAS regressions: Further results and new directions,” Econometric Reviews, 26, 53-90

Giannone, D., Reichlin, L. and Small, D. (2008). “Nowcasting GDP and inflation: The real-time informational content of macroeconomic data releases,” Journal of Monetary Economics, 55, 665-676.

Mazzi, G., Mitchell, J. and Montana, G. (2013). "Density nowcasts and model combination: Nowcasting Euro-area GDP growth over the 2008–09 recession," Oxford Bulletin of Economics and Statistics, DOI: 10.1111/obes.12015.

Smith, P (2013), unpublished PhD Thesis, University of Strathclyde.

Data Appendix

No.	Variable name	Definition	Source	URL for latest data	Transformation	Approximate release date	Monthly (M) or quarterly (Q)	UK, Scottish or other region
-	GVASCOT	Quarterly GVA series for Scotland, constant price, chained volume measure	Scottish Government	`http://www.scotland.gov.uk/Topics/Statistics/Browse/Economy/GDP/Findings`	$\ln(X_t)$-$\ln(X_{t-1})$	15 weeks after end of quarter	Q	Scotland
1	PMISCOT	Headline PMI (output) for Scotland	Bank of Scotland PMI Scotland	—	No transformation	10 days after end of month	M	Scotland
2	PMILON	Headline PMI (output) for London	Markit	—	No transformation	10 days after end of month	M	Other
3	PMISE	Headline PMI (output) for South East England	Markit	—	No transformation	10 days after end of month	M	Other
4	PMISW	Headline PMI (output) for South West England	Markit	—	No transformation	10 days after end of month	M	Other
5	PMIEAST	Headline PMI (output) for East of England	Markit	—	No transformation	10 days after end of month	M	Other
6	PMIWALES	Headline PMI (output) for Wales	Markit	—	No transformation	10 days after end of month	M	Other
7	PMIWMID	Headline PMI (output) for the West Midlands	Markit	—	No transformation	10 days after end of month	M	Other
8	PMIEMID	Headline PMI (output) for the East Midlands	Markit	—	No transformation	10 days after end of month	M	Other
9	PMIYORK	Headline PMI (output) for Yorkshire and the Humber	Markit	—	No transformation	10 days after end of month	M	Other
10	PMINE	Headline PMI (output) for North East England	Markit	—	No transformation	10 days after end of month	M	Other
11	PMINW	Headline PMI (output) for North West England	Markit	—	No transformation	10 days after end of month	M	Other
12	PMINI	Headline PMI (output) for Northern Ireland	Markit	—	No transformation	10 days after end of month	M	Other
13	PMIUK	UK Purchasing Managers Index Output for the UK	Markit	—	No transformation	10 days after end of month	M	UK
14	PMIEZ	Eurozone Purchasing Managers Index Output for the Eurozone	Markit	—	No transformation	10 days after end of month	M	Other
15	PMIWORLD	World Purchasing Managers Index Output for the World	Markit	—	No transformation	10 days after end of month	M	Other

16	VATPAY	(Home) Value Added Tax payments	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
17	VATREPAY	(Home) Value Added Tax repayments	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
18	VATRCPT	(Home) net VAT receipts	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
19	IMPVAT	Import VAT receipts	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
20	TOTALVAT	Total VAT receipts	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
21	NEWVATREG	New registrations for VAT	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
22	VATDEREG	Deregistrations for Value Added Tax	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
23	TRADEPOP	Live (VAT-registered) trader population	HM Revenue and Customs (HMRC)	https://www.uktradeinfo.com/Statistics/Pages/TaxAndDutyBulletins.aspx	$\ln(X_t)$-$\ln(X_{t-12})$	21 days after end of month	M	UK
24	GFKCC	Monthly consumer confidence barometer	Markit	—	No transformation	Approximately 2 weeks after end of month	M	Scotland
25	BOSJOBS_PL	Index of permanent staff placements (seasonally adjusted)	Markit	—	No transformation	21 days after end of month	M	Scotland
26	BOSJOBS_ST	Index of permanent staff demand (seasonally adjusted)	Markit	—	No transformation	21 days after end of month		Scotland
27	UNEM	Claimant count rate (i.e. number of those receiving Jobseekers Allowance divided by those receiving JA plus the number of workforce jobs)	Office for National Statistics	http://www.ons.gov.uk/ons/rel/subnational-labour/regional-labour-market-statistics/index.html	$\ln(X_t)$-$\ln(X_{t-1})$	15 days after end of month	M	Scotland
28	UKIP	Index of Production	Office for National Statistics	http://www.ons.gov.uk/ons/rel/iop/index-of-production/index.html	$\ln(X_t)$-$\ln(X_{t-1})$	Approximately 5 weeks after end of month	M	UK

29	IOSG	UK Index of Services	Office for National Statistics	`http://www.ons.gov.uk/ons/rel/ios/index-of-services/index.html`	$\ln(X_t)$-$\ln(X_{t-1})$	Approximately 6 weeks after end of month	M	UK
30	UKCPI	UK consumer price inflation (CPI) index	Office for National Statistics	`http://www.ons.gov.uk/ons/taxonomy/index.html?nscl=Consumer+Prices+Index`	$\ln(X_t)$-$\ln(X_{t-1})$	Approximately 2 weeks after end of month	M	UK
31	UKEXPORTS	Total UK exports (million, seasonally adjusted)	Office for National Statistics	`http://www.ons.gov.uk/ons/rel/uktrade/uk-trade/index.html`	$\ln(X_t)$-$\ln(X_{t-1})$	Approximately 6 weeks after end of month	M	UK
32	UKIMPORTS	Total UK imports (million, seasonally adjusted)	Office for National Statistics	`http://www.ons.gov.uk/ons/rel/uktrade/uk-trade/index.html`	$\ln(X_t)$-$\ln(X_{t-1})$	Approximately 6 weeks after end of month	M	UK
33	UKREFINE	Throughput of crude and process oil at UK refineries, Table 3.12	Department of Energy and Climate Change	`https://www.gov.uk/government/statistics/oil-and-oil-products-section-3-energy-trends`	$\ln(X_t)$-$\ln(X_{t-12})$	Approximately 8 weeks after end of month	M	UK
34	RSI	Index of retail sales volume at basic prices	Scottish Government	`http://www.scotland.gov.uk/Topics/Statistics/Browse/Economy/PubRSI`	$\ln(X_t)$-$\ln(X_{t-1})$	Approximately 5 weeks after end of quarter	Q	Scotland
35	EXP	Total value of exports from Scotland (current prices)	HM Revenue and Customs (HMRC)	`https://www.uktradeinfo.com/Statistics/RTS/Pages/default.aspx`	$\ln(X_t)$-$\ln(X_{t-4})$	Approximately 9 weeks after end of quarter	Q	Scotland
36	IMP	Total value of imports to Scotland (current prices)	HM Revenue and Customs (HMRC)	`https://www.uktradeinfo.com/Statistics/RTS/Pages/default.aspx`	$\ln(X_t)$-$\ln(X_{t-4})$	Approximately 9 weeks after end of quarter	Q	Scotland
37	SBM	Index of trends in total volume of business	Scottish Business Monitor	—	No transformation	Approximately 4 weeks after end of quarter	Q	Scotland
38	SCBSMAN	Index of manufacturing orders	Scottish Chambers Business Survey	—	No transformation	Approximately 3 weeks after end of quarter	Q	Scotland
39	SCBSCON	Index of construction orders	Scottish Chambers Business Survey	—	No transformation	Approximately 3 weeks after end of quarter	Q	Scotland
40	SCBSTOUR	Index of tourism orders	Scottish Chambers Business Survey	—	No transformation	Approximately 3 weeks after end of quarter	Q	Scotland