ARTICLE IN PRESS

International Journal of Forecasting xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Using stochastic hierarchical aggregation constraints to nowcast regional economic aggregates[☆]

Gary Koop[a], Stuart McIntyre[a,*], James Mitchell[b], Aubrey Poon[c]

[a] University of Strathclyde, United Kingdom
[b] Federal Reserve Bank of Cleveland, United States of America
[c] Örebro University, Sweden

ARTICLE INFO

Keywords:
Regional data
Mixed frequency
Nowcasting
Bayesian methods
Real-time data
Vector autoregressions

ABSTRACT

Recent decades have seen advances in using econometric methods to produce more timely and higher frequency estimates of economic activity at the national level, enabling better tracking of the economy in real-time. These advances have not generally been replicated at the sub-national level, likely because of the empirical challenges that nowcasting at a regional level presents, notably, the short time series of available data, changes in data frequency over time, and the hierarchical structure of the data. This paper develops a mixed-frequency Bayesian VAR model to address common features of the regional nowcasting context, using an application to regional productivity in the UK. We evaluate the contribution that different features of our model provide to the accuracy of point and density nowcasts, in particular, the role of hierarchical aggregation constraints. We show that these aggregation constraints, imposed in stochastic form, play a crucial role in delivering improved regional nowcasts; they prove more important than adding region-specific predictors when the equivalent national data are known, but not when this aggregate is unknown.

1. Introduction

A key barrier to tracking the performance of regional economies lies in the lack of timely official economic data on key aggregates like economic output and productivity. With increasing trends of decentralization and localism, and the associated growing role of sub-national leaders in responding to economic events and fluctuations, there is a need to improve the timeliness of sub-national economic statistics to support more effective policymaking. In the UK, for example, official quarterly regional output data are released with a delay of around six months after the end of the quarter for most regions. Typically, estimates of sub-national economic output are released with a longer lag than equivalent national-level data. This is the case, for example, with US state-level GDP estimates and International Territorial Level (ITL) 1 data for the European Union. This means that local policymakers, who are often on the frontline when responding to the immediate needs created by economic downturns, do not know the severity of any downturn in economic activity for some time after it occurs.

At a national level, the situation is often much better, with GDP data typically released between four and six weeks after the end of the relevant quarter in most advanced economies. Responding to a demand for more

☆ We thank an associate editor and two anonymous referees for helpful comments. This research has been funded by the Office for National Statistics as part of the Economic Statistics Centre of Excellence (ESCoE) research program. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or the Federal Reserve System. All authors are Research Associates at the Economic Statistics Centre of Excellence.

* Corresponding author.
E-mail addresses: gary.koop@strath.ac.uk (G. Koop), s.mcintyre@strath.ac.uk (S. McIntyre), james.mitchell@clev.frb.org (J. Mitchell), aubrey.poon@oru.se (A. Poon).

https://doi.org/10.1016/j.ijforecast.2022.04.002
0169-2070/

timely and higher frequency economic data, one even sees the development and release of monthly GDP data for countries such as the UK (where it is released with a delay of around six weeks after the end of the month). Yet this is still not as timely as policymakers would like, which is why a large amount of literature has explored ways of leveraging more timely data on economic activity, including both 'hard' (e.g., unemployment claims) and 'soft' (e.g., qualitative business survey data) indicators to produce more timely estimates or 'nowcasts' of national economies. However, relatively little research has explored the challenge of sub-national nowcasting (notable exceptions which we draw on below being Koop et al., 2020a, 2020b).

Nowcasting regional economies is a different challenge than nowcasting at the national level. Firstly, longer release delays for regional data mean a longer forecast horizon; secondly, regional data typically have much shorter historical samples; thirdly, regions usually have a more limited set of potential predictors that can be used when nowcasting; and fourthly, smaller regional economies tend to exhibit greater economic volatility than national economies. At the same time, however, the hierarchical relationship between regional and national economic aggregates ought to be able to help address some of these issues. Indeed, given accounting identities, it is undoubtedly an essential feature of any regional nowcast or forecast that they are cross-sectionally consistent with, and informed by, the corresponding national estimates. Below we elaborate on data features observed for the UK and bundle them together under the banner 'measurement errors.' These data features mean that we should not always expect the regional data to 'add up' (or average) exactly to the national data, even when modeling growth rates, as we do in this paper. However, there is no systematic investigation of these issues and how this hierarchical or cross-sectional relationship can be used to produce more accurate and well-calibrated regional nowcasts and forecasts. Our paper seeks to fill this gap by considering how, within our proposed Bayesian estimation strategy, one can, in effect, 'shrink' towards aggregation constraints rather than impose them exactly.

In doing so, we focus on nowcasting productivity growth at a regional level in the UK. Central banks and other policymakers consult the latest productivity estimates when assessing how future output growth might evolve. Productivity is conventionally measured using some measure of economic output (Q) divided by some measure of labor (L) used to produce that output. In this paper, we utilize official estimates from the UK Office for National Statistics (ONS) on gross value added (GVA) and hours worked at the ITL1 (formerly the NUTS1) regions of the UK, of which there are 12 (plus extra-regio[1]) alongside the equivalent national level versions of these data. These measures of Q and L that we use are produced with delay. The exact timing of release delays differs for Q and L, and between the regions and the UK. It has also changed over time. But the general pattern is that the UK data are released more quickly than the regional data. We exploit this fact to produce and evaluate timely nowcasts of regional productivity.

To produce more frequent nowcasts of regional productivity, we use a high-dimensional Mixed Frequency Vector Autoregressive (MF-VAR). MF-VARs have become a popular nowcasting tool; see, among many others, Brave et al. (2019), Eraker et al. (2015), Koop et al. (2020a, 2020b), Schorfheide and Song (2015), and McCracken et al. (2021) But none of these econometric models can be used directly for nowcasting UK regional productivity due to the specific characteristics of our data set. We have two variables, Q and L. At the UK level, we use quarterly observations on these variables going back to 1966Q1. However, regional output data are available at the annual frequency for the first part of the sample (from 1966 to 2011 for Q and 1966 to 1981 for L), but then become quarterly (from 2012Q1 for Q and 1982Q1 for L).

Thus, the timing of the switch from annual to quarterly differs for Q and L. This makes it desirable to create a high dimensional MF-VAR where Q and L (for every region and the UK) are the dependent variables instead of the simpler strategy of working with productivity (Q/L) directly and building a lower-dimensional MF-VAR. Our model is designed to deal with this complicated data setup, where the variables have different release delays and the frequency mismatch changes over time. This mismatch also helps explain why the assembled historical databases, particularly for Q, are a 'mix-and-match' of estimates measured in different ways and at different points in time. This means that our data are not always perfectly consistent in respecting accounting constraints. Therefore, our modeling approach incorporates the hierarchical relationship between the national and regional data in stochastic rather than exact form. Importantly, this still enables us to exploit the information in the more rapidly released UK data on Q and L to improve our regional estimates. Our work thus relates to the hierarchical forecasting and forecast reconciliation literature, for example, the recent contributions from Athanasopoulos et al. (2020), Eckert et al. (2021), and Panagiotelis et al. (2021). We also build on the temporal-and-spatial disaggregation literature, where the focus has been on imposing the cross-sectional aggregation constraints in exact rather than approximate form. This is appropriate if and when the accounting identities are binding. Leading examples of this literature are Di Fonzo (1990), Frale et al. (2011), Guerrero and Nieto (1999), and Proietti (2011a). Proietti (2011b) is the closest paper to ours, as he sets out a frequentist way of imposing cross-sectional aggregation constraints subject to errors. Our point of departure is to consider Bayesian methods appropriate for large dimensional VAR models. VARs are an attractive option for modeling regional data sets such as ours. They allow information in one region to spill over into others, both statically (via the error covariance matrix) and dynamically (via the lagged dependent variables). But, with two variables of interest for each of the 12 regions (as well as UK quantities), the resulting VARs are very high-dimensional. This high-dimensional setting, combined with so many latent states and the

[1] For a definition of 'extra-regio' see p.22 of https://ec.europa.eu/eurostat/documents/3859598/5937641/KS-GQ-13-001-EN.PDF/7114fba9-1a3f-43df-b028-e97232b6bac5.

extremely unbalanced nature of our (panel) data, raises overfitting concerns. Bayesian prior shrinkage can be used to overcome these.

In our empirical application, we focus primarily on evaluating the accuracy of our regional forecasts produced in pseudo-real-time at different forecast horizons. A fully real-time data analysis is not possible. Firstly, historical regional data vintages for Q do not exist. Secondly, some of the regional Q data that we exploit have only been published by ONS since 2017, precluding meaningful analysis of regional data revisions. Given release delays, we produce 3 estimates of growth for a given quarter in each region (i.e., a forecast, a nowcast, and a backcast). We utilize five different versions of our model, reflecting different combinations of our hierarchical (cross-sectional) aggregation constraint over the Q and L variables, as well as exploring the contribution of additional regional-level predictors. Our goal in this exercise is to document the contribution that each of these model features makes to the accuracy of our point and density nowcasts. We show that these hierarchical restrictions significantly improve nowcast accuracy in cases where the UK aggregate has been published and is known, but not when the corresponding UK value is unknown. Furthermore, these gains are much larger than those produced by incorporating additional regional predictors.

The rest of this paper proceeds as follows. The following section sets out the data and context for this paper, particularly the evolving data release calendar and the construction of our regional database. The nature of these data – and the reality that they are subject to measurement error(s) – helps explain why it is not appropriate to impose the hierarchical aggregation constraints in exact form. Section 3 introduces the notation used in this paper, while Section 4 sets out our econometric model with stochastic aggregation constraints appropriate for modeling the data in growth rates. Section 5 presents our results, and Section 6 concludes the paper. Technical details on the variational Bayes estimation algorithm and prior and hyperparameter choices can be found in online Appendix A. Online Appendix B presents robustness checks on the empirical results in the main paper.

2. The evolving regional output and employment data landscape in the UK

This paper aims to develop a model that produces accurate nowcasts of regional labor productivity in the UK. We must model both economic output and labor input at the regional level. Given our focus in this paper, we must also model the equivalent national (UK) aggregates. This section explains the UK data landscape and the construction of the database that we use.

Regional economic output data in the UK share several features common to sub-national data internationally. The available time series of economic output data is relatively short (1998–2019 for annual data and 2012Q1–2020Q3 for quarterly data, at the time of writing) compared to national-level data. It is also much less timely than national-level data; the typical delay in releasing annual regional output data in the UK is about a year. For the quarterly data the typical release delay is six months (with slightly more timely data – produced using a different methodology – for Scotland and Northern Ireland). In the European Union, annual regional data are released with a similar lag. In the US, quarterly state-level GDP are available but are released with a delay of around three months compared to less than one month for US GDP data. Therefore, the UK context provides a good case study to explore the efficacy of regional nowcasting models.

Quarterly real-terms gross value added (GVA) data are readily available for the UK. These data are measured on the output-side, and we consider data from 1966Q1 through to 2020Q3.[2] We use chain-linked volumes. In chained volume terms, real GVA and real GDP growth rates should be the same and, in practice, are very close (they have a correlation coefficient of 0.99). The UK-level GVA data are released about 45 days after the end of the quarter. In the absence, as we shall explain, of real-time data vintages for the regional GVA data, we consider only the latest-vintage UK GVA data (at the time of writing, May 2021). It will be interesting for future research, once a sufficiently large real-time regional dataset does build up, to extend our analysis to consider if and how data revisions affect the accuracy of regional nowcasts.

The first element of our corresponding regional output database is real–term annual regional GVA data for the 12 ITL1 regions of the UK from 1998–2019.[3] We use the ONS's 'preferred' single measure of regional output (GVA(B)), that balances the income and production measures of GVA.[4] These real-terms regional GVA data are chain-linked volumes. Chain-linked volume estimates are not additive, so we should not expect the regional GVA data to sum in levels, across regions, to the UK total as we move away from the base year.[5] Estimates back to 1998 were first published in 2017. This recent history means that sufficient real-time data vintages to allow a meaningful analysis of regional data revisions do not yet exist. We, therefore, use the latest-vintage estimates (at the time of writing, this was the May 2021 vintage). To extend these data back beyond 1998, we make use of historical (revised) vintage data for regional nominal GVA data for 1966–1996, released by the ONS,[6] which, in the absence of regional inflation data we deflate by the UK deflator.[7] This means that the real-terms regional GVA data before 1998 are constant-price. These historical GVA

[2] Recall that GVA plus taxes (less subsidies) on products is GDP: see https://www.ons.gov.uk/ons/rel/elmr/economic-trends--discontinued-/no--627--february-2006/methodology-notes--links-between-gross-domestic-product--gdp--and-gross-value-added--gva-.pdf2.

[3] https://www.ons.gov.uk/economy/grossvalueaddedgva/datasets/nominalandrealregionalgrossvalueaddedbalancedbyindustry.

[4] https://consultations.ons.gov.uk/national-accounts/consultation-on-balanced-estimates-of-regional-gva/supporting_documents/Development%20of%20a%20balanced%20measure%20of%20regional%20gross%20value%20added.pdf.

[5] See p. 192 in the most recent QNA manual available at https://www.imf.org/external/pubs/ft/qna/pdf/2017/chapter8.pdf.

[6] https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome/adhocs/006226historiceconomicdataforregionsoftheuk1966to1996.

[7] One further data issue in combining the historical data (1966–1996) with the more recent data (1997–2019) is that, as identified

data are also measured on the income side only, rather than being balanced estimates on the income and production sides. This switch from constant-price income-based to chain-linked balanced estimates of regional real GVA is a necessary practical step in constructing an historical regional database. These data features contribute to why, econometrically, we impose the hierarchical constraints relating the regional disaggregates to the UK aggregate in stochastic form. We should not expect the constraints to hold exactly, given that our real-terms regional GVA data and the UK GVA data to which the regional data are related are a mix-and-match of constant-price income-based and chain-linked balanced estimates. But we might hope that the constraints remain informative in practice when nowcasting when imposed in stochastic form.

We then add higher frequency quarterly regional real-terms output data with historical coverage back to 2012Q1. The ONS began producing these so-called Regional Short Term Indicators (RSTIs), which include real GVA data, in September 2019; see Koop et al. (2020a) for details. They are referred to by ONS as 'Regional GDP' and are for the nine ITL1 regions of England plus Wales, with equivalent data for Scotland[8] and Northern Ireland[9] (both ITL1 regions of the UK) provided directly by these two devolved administrations. How do these quarterly and annual output data align? Not perfectly. In principle, the quarterly and annual data from ONS should align exactly. The quarterly data should be constrained to sum to the published annual data over the period for which both exist. Delays in a data release and the timing of data revisions will not always ensure that this is the case. For example, at the time of writing, the annual data only runs to 2019, while the quarterly data are available until 2020Q3. When 2020 annual data are published, the quarterly estimates will be revised to constrain to these new annual totals.

The data for Scotland and Northern Ireland are not constrained to the totals published by the ONS, although comparison on a nominal basis suggests quite close alignment. In real terms, however, differences in the approach to deflation mean that the Scottish Government data align less well with the ONS published annual data (Koop et al., 2020a).[10] Complicating the data landscape further, the quarterly regional output data in the UK are released with different delays after the end of the relevant quarter. The Scottish Government operates with a release delay of just under three months. The data for Northern Ireland are released with a delay of just over three months. The ONS data for the English regions and Wales are released with a delay of around six months. Output data for the UK are released with a delay of approximately six weeks after the end of the relevant quarter.

In short, measurement errors (including differences in methodology) explain why temporal and cross-sectional constraints, based on accounting identities, should not be expected to hold exactly. Another factor contributing to measurement error is the reality that, due to the limited availability of vintage/historical regional data, the regional and UK data have not always been through the same revisions and benchmarking processes. This means the data do not always 'add up' (or average) exactly. This lack of consistency between the RSTI data, the regional GVA(B) data, and the UK GDP data is conveniently summarized in this extract from the ONS[11]: '...[RSTIs] will align with the annual growth rates determined by regional accounts while fitting a quarterly path based on the underlying [RSTIs] data. Since regional accounts themselves are constrained to national estimates of GDP, this benchmarking process ensures that [RSTIs] are also broadly in line with the national estimates. However, there may still be inconsistencies between our [RSTIs] data, post regional accounts benchmarking, and our short-term estimates of GDP. This is because there are clear differences in the data sources and methods used (for example, in the extent to which VAT data is used). This means that while [RSTIs] aim to produce the best estimates at a regional level, the sum of the regions (adding in published estimates for Scotland and Northern Ireland) may not equal the national total in the time period following the regional accounts benchmarking.' Indeed, even in nominal terms, as discussed in the online appendix of Koop, McIntyre and Mitchell (2020), the historical (pre-1998) regional GVA data in levels do not sum exactly to the UK total. The ONS acknowledge this via explicit publication of a statistical discrepancy in their underlying Regional Trends publications.

To construct a measure of labor productivity, we also require a measure of labor input, either in the form of hours worked or jobs, to produce a measure of output per hours worked or output per job. Given changes in the structure of the labor market over time, the preferred measure – and the one used in this paper – is hours worked. To construct this measure, we use data on hours worked at a UK level, available at the quarterly frequency for our full sample period.[12] At the regional level, the data are not so easily available. From 1997Q2 to 2021Q1, regional hours worked data are available at the quarterly frequency.[13] Before this, however, these data are not available, so we backcast this measure over the earlier part of our sample using regional-level data on the number of jobs. The cross-sectional constraint between

in Koop, McIntyre and Mitchell (2020), these two datasets in nominal levels evidence a level shift in the value of regional output between 1996 and 1997. Such a shift is not present in the equivalent UK data. It likely reflects differences in how regional output was calculated historically with methods used now. We, therefore, elected to smooth out this spike in the 1996–1997 annual growth rate. As our regional econometric models are estimated in annual growth rates rather than (log) levels, our solution is to proxy the 1997 growth rate with the average growth rates in 1996 and 1998.

8 https://www.gov.scot/collections/economy-statistics/.

9 https://www.nisra.gov.uk/statistics/economic-output-statistics/ni-composite-economic-index.

10 For more on this, see the Scottish Government GDP methodology document here: https://www2.gov.scot/Resource/0054/00542708.pdf.

11 Taken from https://www.ons.gov.uk/economy/grossdomesticproductgdp/methodologies/introducinggdpforthecountriesoftheukandtheregionsofengland.

12 https://www.ons.gov.uk/employmentandlabormarket/peopleinwork/laborproductivity/datasets/laborproductivity.

13 https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/datasets/quarterlyregionalproductivityhoursandproductivityjobsnuts1.

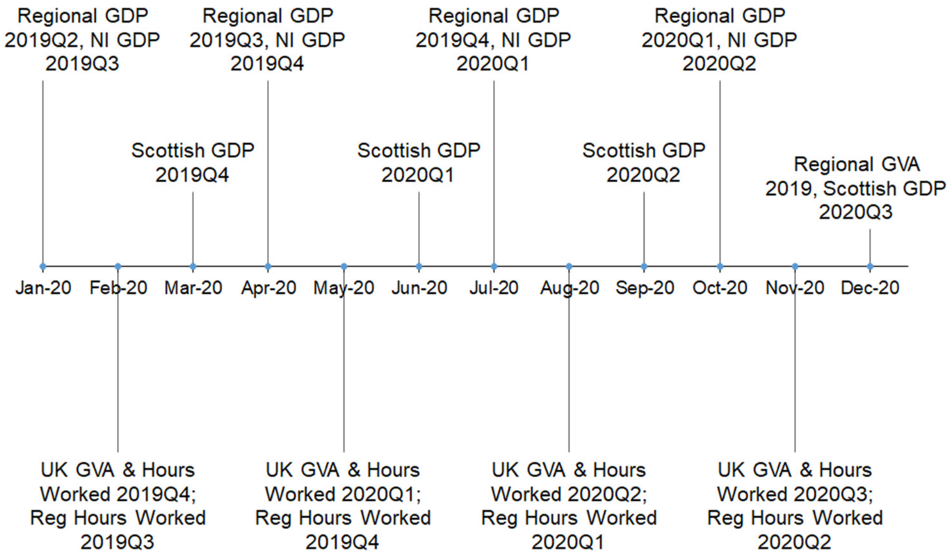

Fig. 1. Illustrative release calendar for UK regional data indicating data published through 2020.

hours worked in the UK and the ITL1 regions of the UK over the first part of our sample will therefore not hold exactly. However, these do aggregate directly in the latter part of our sample. Data on hours worked at the regional level are released three months after the release of equivalent national-level data (which, in turn, are released with a delay of around six weeks). For both Q and L, data at a national level are more readily available with a longer time series than equivalent data at a regional level, motivating our use of stochastic hierarchical aggregation constraints in this paper.

To explore the role of adding region-specific information into our model – acknowledging the limited time series of data available at a regional level – we include two additional quarterly predictors for each region into our model. The first is the claimant count measure of unemployment, and the other is the Confederation of British Industry's (CBI) Business Optimism index. In our model, we also include four UK–wide macroeconomic indicators (the Bank rate, CPI inflation, USD:GBP exchange rate, and the Brent oil price). These data are available in real-time, except for the CPI Index, which has a release delay of one month. They are omitted from Fig. 1 which summarizes the release calendar for our application. We refer to this release calendar below when explaining the timing of our nowcasts and forecasts.

3. Notation and key data features

We begin by describing some variable definitions, relationships, and notational conventions used in this paper. All changes and growth rates referred to below are exact (not log differenced).

- $t = 1, \ldots, T$ runs at the *quarterly* frequency.
- $r = 1, \ldots, R$ denotes the R regions in the UK.
- $j = Q, L$ indexes output (Q) and labor (L).
- $Y_t^{UK,j}$ are Q and L for the UK in quarter t. $y_t^{UK,j}$ are the corresponding quarterly growth rates. These are observed throughout the sample. $y_t^{UK} = (y_t^{UK,Q}, y_t^{UK,L})'$ is the vector of quarterly growth rates for UK quantities.
- $Y_t^{r,j}$ are Q and L for region r. $y_t^{r,j}$ are the corresponding quarterly growth rates. For $j = Q$ these are not observed before 2012 and for $j = L$ these are not observed before 1997. $y_t^Q = \left(y_t^{1,Q}, y_t^{1,L}, \ldots, y_t^{R,L}\right)'$ is the vector of quarterly growth rates for regional quantities.
- $Y_t^{r,j,A} = Y_t^{r,j} + Y_{t-1}^{r,j} + Y_{t-2}^{r,j} + Y_{t-3}^{r,j}$ are annual Q and L for region r.[14] $y_t^{r,j,A}$ are the corresponding annual growth rates. These are observed in quarter 4 of each year throughout the sample, but not in other quarters. $y_t^A = \left(y_t^{1,Q,A}, y_t^{1,L,A}, \ldots, y_t^{R,L,A}\right)'$ is the vector of annual growth rates for regional quantities.

The MF-VAR we work with uses $y_t = \left(y_t^{UK}, y_t^{Q\prime}\right)'$ as the vector of dependent variables.[15] Note that the regional variables in our MF-VAR are not observed for some periods of time. These unobserved quarterly quantities are constrained to add up to observed annual quantities via the following temporal constraint (see equation C.1 in the online appendix to Koop et al. 2020a):

$$y_t^{r,j,A} = c^{r,j} + \frac{1}{4}y_t^{r,j} + \frac{1}{2}y_{t-1}^{r,j} + \frac{3}{4}y_{t-2}^{r,j} + y_{t-3}^{r,j} + \frac{3}{4}y_{t-4}^{r,j} + \frac{1}{2}y_{t-5}^{r,j} + \frac{1}{4}y_{t-6}^{r,j} + \eta_{t,j}^r. \quad (1)$$

14 This identity in fact holds as the annual average, rather than the annual sum of the quarterly values, when Y is defined as an index rather than in levels, as it is for L and for real-terms (chained-linked volume) regional GVA after 1998.

15 In our empirical work, as described earlier, we augment this vector with four additional UK quarterly predictors and include two additional regional predictors as exogenous variables.

This constraint, familiar in the MF-VAR literature as a Mariano and Murasawa (2003)-type temporal aggregation constraint when applied to log differenced data, is an approximation to avoid the need for a nonlinear measurement equation.[16] The approximation becomes poorer in more volatile periods. For this reason, and reflecting the data inconsistency issues discussed in the preceding section, we impose the constraint stochastically, allowing for an error, $\eta^r_{t,j}$. To allow for bias, we include intercepts, $c^{r,j}$, in (1), and for them we use relatively non-informative priors centered over 0. For the error variances, $\eta^r_{t,j} \sim N(0, \sigma^2_{j,r})$, we also use relatively non-informative priors. But for robustness, as summarized below (with detailed forecasting results in online Appendix B), we do consider alternative priors that reflect a belief that $\sigma^2_{j,r}$ is small.

Note that the temporal constraint involves seven quarters. This motivates our choice of seven lags in the VAR model (see online Appendix A). However, as described below, our global-local shrinkage prior can shrink extraneous parameters to zero if, as is likely, our lag length is too long. This type of shrinkage prior can be thought of as an automatic way of selecting lag length. It has the advantage over conventional methods because it allows different equations in the VAR to have different lag lengths, and the included lags need not be sequential (e.g., it could consist of first and fourth lags, but not the second and third).

In addition, we exploit the constraints that UK growth rates are approximately, due to the data characteristics explained above, the weighted sums of regional growth, where the weights are the shares of each region in the national total. We refer to these as cross-sectional aggregation constraints. For $j = Q, L$ these can be shown to be (see equation C.2 in the online appendix to Koop et al., 2020b):

$$y^{UK,j}_t = c^j + \sum_{r=1}^{R} w^{r,j}_t y^{r,j}_t + \eta_{t,j}, \tag{2}$$

where $w^{r,j}_t = \left(\frac{Y^{r,j}_t}{\sum_{r=1}^{R} Y^r_t}\right)$ is the share of the regional quantity in the aggregate quantity in quarter t. This share is not observable for L at the quarterly frequency for much of our sample, while it is never observed quarterly, only annually for Q. Therefore, we proxy $w^{r,j}_t$ based on the observed averages from 1999–2013 (our out-of-sample evaluation period starts in 2014). This use of an average is (yet) another reason not to expect the cross-sectional constraint to hold exactly.

We again allow for intercepts and errors in the aggregation constraints as emphasized above. We assume $\eta_{t,j} \sim N(0, \sigma^2_j)$.[17] Below, we refer to robustness checks undertaken to explore the sensitivity of results to how (1) and (2) are specified.

We remind the reader that we are modeling regional growth rates; and seeking to test, when nowcasting, the empirical utility of imposing stochastic hierarchical constraints. We do not aim to study the underlying levels data, nor do we assess to what degree the implied level nowcasts of L and Q are aggregation-consistent. Finally, details of our priors, including prior hyperparameter choice, are given in Appendix A.

4. Overview of econometric methods

This section provides an overview of the econometric methods used in this paper. Complete details are given in online Appendix A. We use a MF-VAR, which as noted above is simply a VAR using $y_t = \left(y^{UK}_t, y^{Q\prime}_t\right)'$ as the vector of dependent variables. It differs from a standard VAR in that some variables are not observed (at least for some periods). These are treated as unobserved latent states, and the MF-VAR is a state space model. The measurement equations in the state space model are given by the temporal constraints given in (1), which link the observed annual regional data to the unobserved quarterly counterparts. The model just described is a standard MF-VAR as used, among other things, in Schorfheide and Song (2015). Standard Bayesian methods exist for estimating this model. Our model differs in several ways. First, the frequency mismatch changes at different points in time for different variables. Methods for handling such a change are developed in Koop et al. (2020a) and described for the present case in online Appendix A. Intuitively, at points in time when regional data switch from annual to quarterly, the appropriate blocks of the model switch to becoming VARs instead of MF-VARs.

A second difference from most existing MF-VARs is that our model is much larger, with many more latent variables to estimate (and a long lag length). For instance, Schorfheide and Song (2015) use an MF-VAR with 11 variables, of which three are at the lower frequency. We have 24 low-frequency variables (Q and L for each of the 12 regions) and only six high-frequency variables

[16] Nonlinear methods that allow for exact treatment of the temporal constraint have been developed; e.g., see Proietti (2011a). Our preference in this paper is to stick with linear methods appropriate for VAR models and instead consider the utility of stochastic imposition of the constraints.

[17] Another reason for allowing for an error in this aggregation constraint is because of our modeling choice not to include extra-regio, sometimes referred to in the UK as the 'UK continental shelf' (UKCS) in our vector of regional outputs, given its idiosyncratic time series properties. UKCS mostly reflects oil and gas output from the North Sea. Since both the quantity of oil and gas produced and their price have fluctuated greatly over time, it is a very volatile series, with time series properties that are very different from other regions of the UK. Koop, McIntyre and Mitchell (2020) similarly preferred to omit UKCS from their model, finding that its inclusion did not help deliver improved regional nowcasts (which is not surprising given the global nature of the drivers of oil and gas activity and UKCS activity's relatively small share of UK GDP (0.8% in 2018)). But output from the UKCS is included in the UK GVA total. This means UK output is not the sum of the regions' output in levels (even when we ignore measurement errors). Note that it is not possible to remove UKCS activity from the overall estimates of UK quarterly GVA because it is not separately identified within quarterly GDP and includes activity in multiple sectors of the economy. It is common for countries to have an 'extra regio' element within their national accounts (capturing, for example, activities in overseas embassies, military bases, territorial enclaves, and so on. For a more formal definition, see https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/gvadataforextraregio).

(Q and L for the UK plus four additional UK variables). In such a high-dimensional model with so many latent states to estimate, the role of prior information becomes important to avoid over-fitting. In the extensive Bayesian VAR literature, various global-local or variable selection priors have been shown to work well. They automatically shrink superfluous coefficients to zero with minimal subjective prior input from the researcher (see, e.g., Gefang, 2014; George et al., 2008; Koop, 2013; Korobilis, 2013, and Kastner & Huber, 2020). In this paper, we use a popular prior in this class: the adaptive Lasso prior.[18]

To explain the adaptive Lasso, consider equation i of the VAR which involves k_i right hand side variables. The adaptive Lasso prior is a conditionally Normal prior with mean zero (thus ensuring shrinkage towards zero) and covariance matrix:

$$\mathbf{V}_i = \mathrm{diag}(\tau_{i,1}, \ldots, \tau_{i,k_i}). \tag{3}$$

The $\tau_{i,j}$, which control the strength of shrinkage of the jth coefficient in the ith equation, are treated as unknown parameters that are estimated in a data-based fashion. These shrinkage parameters have a prior of the form:

$$\tau_{i,j} \sim Exp(\tfrac{\lambda_{i,j}}{2}), \quad \text{for } j = 1, \ldots, k_i \tag{4}$$

with

$$\lambda_{i,j} \sim G(\underline{a}_0, \underline{b}_0), \tag{5}$$

where Exp denotes the exponential distribution and G denotes the Gamma distribution. Note that the only prior hyperparameters to be chosen are $\underline{a}_0$ and $\underline{b}_0$. See online Appendix A for more details on the specification of the hyperparameters. Zou (2006) demonstrates that the adaptive Lasso satisfies so-called oracle properties. These include asymptotically selecting the correct variables and having the optimal estimation rate.

The model we have specified thus far is identified, and, in our empirical section, we will show it produces sensible nowcasts of regional productivity growth. However, we can exploit other features not available with the conventional MF-VAR. These are the aggregation constraints given in (2). Note that there are two of these, one for Q and one for L. One of the issues we explore in our empirical application is whether exploiting either or both of these can help improve the accuracy of our nowcasts. As shown in Koop et al. (2020b), an aggregation constraint can be treated as an additional measurement equation in the state space model that defines the MF-VAR. Bayesian estimation and nowcasting in state space models are achieved using Markov Chain Monte Carlo (MCMC) methods, but these can be computationally slow in MF-VARs of high dimensions such as we have. Accordingly, we use Variational Bayes (VB) methods. VB methods are approximate but are much faster than MCMC methods. Gefang et al. (2022) develop VB methods for MF-VARs, and we adopt similar methods in this paper. Further computational details are in online Appendix A.

Mixed frequency models are particularly useful for providing timely updates of low-frequency variables. Given release delays in the regional data, as described in Section 2, for any quarter τ of the year, we produce three estimates of any quarter's regional growth rate as new information is released. We refer to these as the forecast, the nowcast, and the backcast. Take the 2nd quarter of the year as an example (it may be useful for the reader to refer to Fig. 1 while reading this explanation). We receive official estimates of Q2 UK GVA growth in August of that year, at which time the latest regional data that we have (for expositional purposes, setting aside complications that arise from the more timely data for Scotland and Northern Ireland) is for Q4 of the previous year.

More formally, our timing convention is that we run the model immediately upon receipt of the latest UK output data for each quarter, which takes place for quarter $\tau - 1$ mid-way through quarter τ, and we incorporate the latest available data on all other indicators. When we do this, we have a two-quarter gap between the available UK and regional output data, given the six-month release delay which exists for the regional output data. We, therefore, produce a 'nowcast' for each region for $\tau - 1$, as well as a 'backcast' for each region for $\tau - 2$, and a 'forecast' for each region for τ. For regional labor market data (hours worked in our case) these are released three months after the release of equivalent data for the UK as a whole, and so, we only need a 'nowcast' and a 'forecast' for these data, as at time τ we already have data on hours worked at a regional level for $\tau - 2$ making the production of a backcast redundant.

Prediction is carried out using the simulation methods set out in online Appendix A. But, in summary, our VB methods provide us with an approximation of the posterior of all the parameters and states in the model. We simulate draws from this posterior and, conditional on each draw, produce a one-step-ahead forecast from the VAR. This procedure is repeated for 1000 draws from the posterior. This procedure will produce draws of the growth of Q and L (for the UK and the regions). These draws are then differenced to produce draws of productivity growth.

We also present results from an AR(1) benchmark, which uses a relatively non-informative prior, based on taking 10,000 posterior draws with a 5,000 burn-in. The timing convention for the AR(1) model is as follows. Again we produce 'forecasts,' 'nowcasts,' and 'backcasts' from this model coincident with the release of UK output data (i.e., four times a year, mid-quarter τ). At τ, the latest regional output data relate to $\tau - 3$. So to compute forecasts we estimate an AR(1) model of regional Q for τ on $\tau - 3$, using the sample $\tau = 1$ through $\tau - 3$. Then, given these parameter estimates, we use the latest (in real-time) known regional Q value for $\tau - 3$ to forecast τ (i.e., the current quarter).

To compute nowcasts, we estimate an AR(1) regression of regional Q for τ on $\tau - 2$, using the sample $\tau = 1$ through $\tau - 3$. Then we use the latest (in real-time) known value for $\tau - 3$ to estimate $\tau - 1$. To compute backcasts, we estimate an AR(1) regression of regional Q for quarter τ

[18] Results using the Horseshoe prior are available in the online Appendix B. These tend to be very similar to the adaptive Lasso.

on $\tau-1$, using the sample τ= 1 through $\tau-3$. Then we use the latest (in real-time) known value for $\tau-3$ to estimate $\tau-2$.

AR(1) estimation for L proceeds in a similar way, but we only need to produce forecasts and nowcasts, given L is published more rapidly than Q. The forecast now involves estimating L for τ on L at $\tau-2$, τ=1, …, $\tau-2$, and using L at $\tau-2$ to estimate L for τ. The nowcast involves estimating L for τ on $\tau-1$, τ=1, …, $\tau-2$, and using L at $\tau-2$ to estimate L at $\tau-1$.

5. Empirical results: Quarterly regional growth estimates

5.1. Overview

This section investigates how our mixed frequency methods perform in forecasting, nowcasting, and backcasting regional productivity growth. We consider five different variants of the MF-VARs. Four of these differ in how we treat the stochastic cross-sectional aggregation constraint (2). In particular, we consider versions of the MF-VAR, which impose (2) both on Q and L, just on Q, just on L, and on neither Q nor L. The fifth model imposes the cross-sectional aggregation constraint on both Q and L but does not include the region-specific exogenous predictors. We include this model to investigate whether including additional regional predictors does help improve our estimates of quarterly regional productivity. The temporal aggregation constraint (1) is imposed in each of these five models to ensure temporal consistency. Our primary interest is in assessing the utility of the cross-sectional constraint. We wish to test if and when there are accuracy gains in conditioning the regional estimates on those for the UK. Hence, our interest in these five models. Results also include a benchmark AR(1) model, as described in the previous section. In online Appendix B, we present two sets of results from our MF-VAR models when we impose restricted variants of the two stochastic constraints, (1) and (2). First, we set $c^{r,j}=0$ in (1) and $c^{j}=0$ in (2). Second, in addition to setting $c^{r,j}=0$ in (1) and $c^{j}=0$ in (2), we set $\sigma^2_{j,r}=0$, and thus allow for only the cross-sectional constraint to be stochastic, for the reasons elucidated in Section 2. Results are similar to those presented below. As an additional robustness check, Appendix B also presents results using a prior that is more informative than that used in the body of the paper and reflects a belief that $\sigma^2_{j,r}$ is small. And we present results when using the Horseshoe rather than the adaptive Lasso prior. Again, in both cases, results are qualitatively similar to those presented below.

We use root mean squared forecasts errors (RMSEs) and continuous ranked probability scores (CRPSs) to evaluate our point and density estimates. Our out-of-sample evaluation period is 2014Q2 through 2020Q3.[19] Note that this choice is made to cover the time period the RSTI data are available for, allowing for sufficient lags to apply the temporal constraint in (1) throughout the evaluation period.

5.2. Forecasts, nowcasts, and backcasts of regional productivity growth

Tables 1 and 2 evaluate our forecasts, nowcasts, and backcasts of regional productivity growth for the regions of the UK. In relation to point forecast performance, as measured by RMSEs, the MF-VAR forecasts with aggregation constraints on Q, and both Q and L at the same time, do not beat the simple AR(1) benchmark. Remember that the forecasts for time t reflect information known at that time, which at a regional level is output growth for $(t-2)$, and data on hours worked for $(t-1)$; while, at a national level, only data for time t are available. Thus, we still have to forecast the 'aggregate' to which the cross-sectional constraint applies. It is therefore unsurprising that the aggregation constraints do not add anything in this case since we do not know the quarter $t+1$ UK quantities yet. Thus, the aggregation constraints are of little benefit. Without any aggregation constraints, and when the aggregation constraint only applies to the hours worked data, we do see some improvements over the AR(1) model.[20]

When we turn to the nowcasts and backcasts, the MF-VAR methods are delivering point forecast performance which is substantially better than the AR(1). Why might this be? Recall that at time t we produce regional 'nowcasts' for t conditional on the UK realization for time t, and this appears to help produce more accurate nowcasts. Specifically, we see the utility of our aggregation constraints, (2), that link observed UK growth to our regional estimates. At the same time, when producing our 'backcasts' at time t, we are conditioning on time t and $t-1$ data for the UK (at time t the latter will also incorporate data revisions relative to its first release value). This UK information allows us to refine our estimates of regional productivity growth and, given release delays, do so with a significant timing advantage.

When we consider density forecast performance, as measured by CRPS, we obtain a similar story, but even more favorable to the MF-VARs. That is, with some exceptions to be discussed below, all three of our estimates (i.e., forecasts, nowcasts, and backcasts) now beat the AR(1) benchmark, often by a substantial margin. Regarding the cross-sectional aggregation constraints, which are one of the main links in the model between regional Q and L and UK Q and L, the overall picture is that it is beneficial to impose both of them. However, there are some subtle differences between backcasts, nowcasts, and forecasts worth exploring. We have strong evidence from both RMSEs and CRPSs that imposing both cross-sectional constraints leads to substantial improvement in the backcasts. For instance, looking at Table 1 (Table 2),

[19] Note that, at the time of writing, 2020Q3 is the latest data available, and it includes some of the pandemic period. We have repeated all our empirical work using a data set that ends with 2019Q4, and the main conclusions are very similar.

[20] Our short evaluation period precludes a strong interpretation, but Tables B.25 through B.30 in online Appendix B.5 show that these gains over the AR(1) can be statistically significant. Following Diebold and Mariano (2002) and Giacomini and White (2006), we use a t-statistic, assuming asymptotic normality and serially uncorrelated errors (expected for *optimal* one-step-ahead forecasts, nowcasts, and backcasts), and implement a two-sided test of equal forecast accuracy. These tests assume the impact of decreasing parameter uncertainty due to recursive estimation is negligible.

Table 1
RMSFE (multiplied by 100) for productivity growth, from an AR(1) and five MF-VAR models differing in whether they impose the cross-sectional aggregation constraint (2) on output (Q) and/or labor (L).

	NE	YH	EM	EE	LON	SE	SW	WM	NW	WA	SCOT	NI	Average
AR(1) model													
Forecast	2.35	3.11	4.00	2.16	1.91	3.18	2.51	2.48	2.12	3.80	2.04	1.81	**2.62**
Nowcast	1.63	1.86	1.96	2.10	1.78	1.86	1.85	2.03	1.51	1.66	1.51	1.03	**1.73**
Backcast	1.66	1.91	1.89	1.85	1.73	1.70	1.69	1.87	1.35	1.71	1.32	1.24	**1.66**
MF-VAR model - (with both aggregation constraints)													
Forecast	4.21	2.74	1.73	3.02	3.63	2.99	1.89	1.59	3.51	3.82	2.54	1.84	**2.79**
Nowcast	2.50	2.20	2.15	1.22	1.87	1.55	0.82	0.58	2.54	2.06	0.83	0.74	**1.59**
Backcast	1.07	1.17	1.15	0.24	0.94	0.92	0.45	0.33	1.19	0.81	0.20	0.22	**0.72**
MF-VAR model - (aggregation constraint only in Q)													
Forecast	4.22	2.79	1.73	3.12	3.72	3.18	1.98	1.73	3.53	3.83	2.59	1.86	**2.86**
Nowcast	2.05	1.64	1.97	1.16	1.70	1.63	0.68	0.57	2.04	1.81	0.84	0.72	**1.40**
Backcast	1.04	1.16	1.31	0.22	0.88	1.02	0.45	0.32	1.23	0.82	0.21	0.22	**0.74**
MF-VAR model - (aggregation constraint only in L)													
Forecast	2.06	1.43	1.56	1.13	1.32	1.35	1.20	1.25	1.32	1.75	0.72	1.71	**1.40**
Nowcast	1.61	1.57	2.12	2.09	1.08	1.56	1.54	1.92	1.47	1.29	1.51	1.24	**1.58**
Backcast	0.91	1.05	1.23	1.17	0.87	1.05	1.06	1.15	1.10	0.78	0.93	0.65	**1.00**
MF-VAR model - (No aggregation constraints)													
Forecast	2.01	1.46	1.55	1.10	1.31	1.35	1.20	1.27	1.17	1.72	0.72	1.57	**1.37**
Nowcast	2.01	1.90	2.50	2.49	1.50	1.95	1.82	2.35	1.85	1.55	1.81	1.43	**1.93**
Backcast	0.98	1.05	1.23	1.16	0.86	1.04	1.06	1.17	1.13	0.79	0.92	0.64	**1.00**
MF-VAR model - (with both aggregation constraints but no exogenous predictors)													
Forecast	4.18	2.76	1.88	3.09	3.70	3.08	1.96	1.89	3.35	3.88	2.44	1.93	**2.85**
Nowcast	2.23	2.18	2.40	1.26	1.79	1.59	0.85	0.59	2.38	1.96	0.80	0.74	**1.56**
Backcast	1.00	1.15	1.31	0.22	0.88	0.96	0.44	0.33	1.16	0.78	0.16	0.23	**0.72**

Regional abbreviations: NE - North East England, YM - Yorkshire and Humber, EM - East Midlands, EE - East of England, LON - London, SE - South East England, SW - South West England, WM - West Midlands, NW - North West England, WA - Wales, SCOT - Scotland, NI - Northern Ireland.

Table 2
CRPS (multiplied by 100) for productivity growth, from an AR(1) and five MF-VAR models differing in whether they impose the cross-sectional aggregation constraint (2) on output (Q) and/or labor (L).

	NE	YH	EM	EE	LON	SE	SW	WM	NW	WA	SCOT	NI	Average
AR(1) model													
Forecast	1.44	2.04	2.65	1.31	1.21	2.18	1.66	1.56	1.41	2.36	1.31	1.13	**1.69**
Nowcast	0.76	0.80	0.95	0.78	0.76	0.82	0.79	0.85	0.62	0.77	0.59	0.50	**0.75**
Backcast	0.70	0.71	0.80	0.67	0.71	0.70	0.64	0.72	0.49	0.63	0.43	0.49	**0.64**
MF-VAR model - (with both aggregation constraints)													
Forecast	1.39	1.20	0.96	1.07	1.30	1.22	0.89	0.77	1.10	1.56	0.88	0.81	**1.10**
Nowcast	0.99	0.86	0.80	0.57	0.81	0.72	0.44	0.33	0.87	0.93	0.34	0.39	**0.67**
Backcast	0.39	0.34	0.36	0.14	0.30	0.32	0.20	0.17	0.39	0.30	0.09	0.11	**0.26**
MF-VAR model - (aggregation constraint only in Q)													
Forecast	1.40	1.22	0.95	1.09	1.30	1.28	0.93	0.84	1.15	1.58	0.89	0.81	**1.12**
Nowcast	0.84	0.74	0.76	0.52	0.77	0.76	0.39	0.33	0.76	0.85	0.34	0.37	**0.62**
Backcast	0.38	0.34	0.39	0.13	0.29	0.34	0.20	0.16	0.40	0.30	0.09	0.11	**0.26**
MF-VAR model - (aggregation constraint only in L)													
Forecast	1.07	0.80	0.87	0.64	0.76	0.77	0.68	0.69	0.74	1.04	0.45	0.84	**0.78**
Nowcast	0.75	0.68	0.85	0.69	0.54	0.68	0.66	0.71	0.60	0.66	0.52	0.53	**0.66**
Backcast	0.35	0.32	0.40	0.34	0.30	0.33	0.34	0.37	0.39	0.25	0.26	0.25	**0.33**
MF-VAR model - (No aggregation constraints)													
Forecast	1.08	0.83	0.86	0.63	0.76	0.78	0.69	0.71	0.69	1.01	0.45	0.80	**0.77**
Nowcast	0.90	0.76	0.92	0.78	0.65	0.77	0.71	0.82	0.70	0.72	0.58	0.58	**0.74**
Backcast	0.40	0.32	0.40	0.36	0.31	0.34	0.34	0.38	0.40	0.26	0.26	0.27	**0.34**
MF-VAR model - (with both aggregation constraints but no exogenous predictors)													
Forecast	1.38	1.26	1.04	1.11	1.31	1.25	0.95	0.85	1.07	1.68	0.85	0.83	**1.13**
Nowcast	1.04	0.98	0.99	0.68	0.88	0.84	0.60	0.41	0.95	1.04	0.38	0.46	**0.77**
Backcast	0.43	0.37	0.41	0.16	0.32	0.36	0.23	0.21	0.41	0.32	0.07	0.14	**0.29**

Regional abbreviations: NE - North East England, YM - Yorkshire and Humber, EM - East Midlands, EE - East of England, LON - London, SE - South East England, SW - South West England, WM - West Midlands, NW - North West England, WA - Wales, SCOT - Scotland, NI - Northern Ireland.

Table 3
RMSFE (multiplied by 100) for output growth, from an AR(1) and five MF-VAR models differing in whether they impose the cross-sectional aggregation constraint (2) on output (Q) and/or labor (L).

	NE	YH	EM	EE	LON	SE	SW	WM	NW	WA	SCOT	NI	Average
AR(1) model													
Forecast	3.10	3.03	4.35	2.63	2.59	3.38	2.51	3.93	2.39	3.11	2.25	1.91	**2.93**
Nowcast	1.96	2.23	2.46	2.29	1.93	2.12	2.14	2.42	1.91	1.95	1.79	1.33	**2.04**
Backcast	1.66	1.91	1.89	1.85	1.73	1.70	1.69	1.87	1.35	1.71	1.32	1.24	**1.66**
MF-VAR model - (with both aggregation constraints)													
Forecast	3.18	1.95	2.01	2.13	2.64	2.21	1.38	1.83	2.34	2.81	1.53	1.07	**2.09**
Nowcast	1.84	1.71	1.87	0.53	1.40	1.65	0.85	0.97	1.96	1.33	0.45	0.71	**1.27**
Backcast	1.07	1.17	1.15	0.24	0.94	0.92	0.45	0.33	1.19	0.81	0.20	0.22	**0.72**
MF-VAR model - (aggregation constraint only in Q)													
Forecast	3.15	1.94	1.99	2.18	2.71	2.33	1.42	1.85	2.33	2.80	1.55	1.08	**2.11**
Nowcast	1.51	1.44	2.02	0.47	1.23	1.58	0.86	0.97	1.63	1.16	0.41	0.67	**1.16**
Backcast	1.04	1.16	1.31	0.22	0.88	1.02	0.45	0.32	1.23	0.82	0.21	0.22	**0.74**
MF-VAR model - (aggregation constraint only in L)													
Forecast	2.79	2.38	2.55	2.38	2.04	2.42	2.23	2.74	2.25	1.95	1.96	2.05	**2.31**
Nowcast	1.64	1.78	2.13	1.89	1.49	1.77	1.79	2.08	1.72	1.27	1.56	1.23	**1.69**
Backcast	0.91	1.05	1.23	1.17	0.87	1.05	1.06	1.15	1.10	0.78	0.93	0.65	**1.00**
MF-VAR model - (No aggregation constraints)													
Forecast	2.76	2.38	2.56	2.41	2.07	2.43	2.24	2.76	2.22	1.92	1.97	1.90	**2.30**
Nowcast	1.75	1.79	2.12	1.91	1.52	1.78	1.79	2.11	1.76	1.29	1.57	1.23	**1.72**
Backcast	0.98	1.05	1.23	1.16	0.86	1.04	1.06	1.17	1.13	0.79	0.92	0.64	**1.00**
MF-VAR model - (with both aggregation constraints but no exogenous predictors)													
Forecast	3.14	2.01	1.90	2.12	2.62	2.23	1.38	1.84	2.19	2.74	1.47	1.12	**2.06**
Nowcast	1.61	1.67	2.10	0.51	1.26	1.66	0.84	0.95	1.79	1.20	0.38	0.76	**1.23**
Backcast	1.00	1.15	1.31	0.22	0.88	0.96	0.44	0.33	1.16	0.78	0.16	0.23	**0.72**

Regional abbreviations: NE - North East England, YM - Yorkshire and Humber, EM - East Midlands, EE - East of England, LON - London, SE - South East England, SW - South West England, WM - West Midlands, NW - North West England, WA - Wales, SCOT - Scotland, NI - Northern Ireland.

the cross-region average of RMSE (CRPS) is 0.72 (0.26) when they are imposed but rises to 1.00 (0.34) when they are not imposed.

It is worth stressing that the MF-VAR without the aggregation constraints allows newly released UK information to update the regional backcasts, since UK quantities are included as dependent variables in the VAR. However, the MF-VAR with aggregation constraints has this property with the additional link between the UK and its regions provided by these constraints. This additional link helps improve the backcasts. This same pattern holds to a lesser extent with the nowcasts. For the density forecasts, like the point forecasts, there is little or no benefit to imposing the aggregation constraint. That is, the MF-VAR with no aggregation constraints imposed is forecasting better than other alternatives.

Having established the usefulness of the aggregation constraints, at least in improving nowcasts and backcasts, we assess whether including exogenous regional-level predictors is similarly useful. The results are quite similar if we compare the MF-VAR with aggregation constraints to this model with these exogenous predictors added. In other words, these predictors are adding only very small improvements to our forecasts, nowcasts, and backcasts.

5.3. Forecasts, nowcasts, and backcasts of output and employment growth

The previous subsection discussed results for regional productivity (Q/L) growth. But, of course, our MF-VARs also produce results for Q and L growth individually. By examining these, we can better understand where the improvements in the productivity estimates provided by the MF-VAR come from. Tables 3 through 6 present results for these variables individually using the same models as in the previous subsection. Note that regional L data are released more quickly than regional Q data, and at the time we make the backcast, the initial release for L has already occurred. For this reason, we do not provide a backcast for L.

The results for output growth exhibit a similar pattern to those for productivity growth. The MF-VARs offer substantial improvements over the AR(1) benchmark, particularly for the nowcasts and backcasts. In addition, the benefits of imposing the cross-sectional aggregation constraints can be seen. However, for the growth in hours worked, results are weaker. Our MF-VARs are not beating the AR(1) benchmark, and there is little benefit to imposing the aggregation constraints. This may well reflect greater persistence of changes in hours worked than output growth. Or, the time series properties of growth in hours worked relative to output growth make an AR(1) model harder to beat (as we see from the results' tables). We are finding evidence that the improvement in the quality of the estimates of productivity growth produced by the MF-VAR is mostly due to improvements in the quality of the output growth estimates.

5.4. Calibration of the forecasts, nowcasts, and backcasts

The results above indicated that the MF-VAR performed well relative to a simple benchmark and

Table 4
CRPS (multiplied by 100) for output growth, from an AR(1) and five MF-VAR models differing in whether they impose the cross-sectional aggregation constraint (2) on output (Q) and/or labor (L).

	NE	YH	EM	EE	LON	SE	SW	WM	NW	WA	SCOT	NI	Average
AR(1) model													
Forecast	1.61	1.41	2.07	1.40	1.67	1.68	1.36	1.83	1.08	1.37	0.95	0.91	**1.45**
Nowcast	0.85	0.89	1.06	0.92	0.88	0.91	0.88	1.01	0.76	0.81	0.68	0.58	**0.85**
Backcast	0.70	0.71	0.80	0.67	0.71	0.70	0.64	0.72	0.49	0.63	0.43	0.49	**0.64**
MF-VAR model - (with both aggregation constraints)													
Forecast	1.22	0.86	0.95	0.89	1.03	1.00	0.66	0.75	0.91	1.15	0.60	0.55	**0.88**
Nowcast	0.74	0.58	0.80	0.30	0.58	0.59	0.39	0.43	0.61	0.57	0.19	0.34	**0.51**
Backcast	0.39	0.34	0.36	0.14	0.30	0.32	0.20	0.17	0.39	0.30	0.09	0.11	**0.26**
MF-VAR model - (aggregation constraint only in Q)													
Forecast	1.20	0.85	0.94	0.90	1.05	1.04	0.68	0.77	0.93	1.14	0.62	0.55	**0.89**
Nowcast	0.65	0.52	0.83	0.28	0.54	0.57	0.40	0.44	0.55	0.52	0.19	0.32	**0.48**
Backcast	0.38	0.34	0.39	0.13	0.29	0.34	0.20	0.16	0.40	0.30	0.09	0.11	**0.26**
MF-VAR model - (aggregation constraint only in L)													
Forecast	1.43	1.02	1.19	1.11	1.10	1.09	0.98	1.12	1.12	0.97	0.76	0.99	**1.07**
Nowcast	0.81	0.71	0.88	0.78	0.71	0.74	0.70	0.81	0.72	0.59	0.54	0.54	**0.71**
Backcast	0.35	0.32	0.40	0.34	0.30	0.33	0.34	0.37	0.39	0.25	0.26	0.25	**0.33**
MF-VAR model - (No aggregation constraints)													
Forecast	1.43	1.03	1.19	1.12	1.11	1.08	0.98	1.13	1.08	0.94	0.77	0.94	**1.07**
Nowcast	0.87	0.71	0.88	0.80	0.73	0.73	0.71	0.82	0.75	0.59	0.55	0.55	**0.72**
Backcast	0.40	0.32	0.40	0.36	0.31	0.34	0.34	0.38	0.40	0.26	0.26	0.27	**0.34**
MF-VAR model - (with both aggregation constraints but no exogenous predictors)													
Forecast	1.30	0.90	0.94	0.91	1.04	0.99	0.68	0.77	0.87	1.10	0.58	0.57	**0.89**
Nowcast	0.82	0.64	0.93	0.35	0.63	0.66	0.49	0.52	0.63	0.59	0.17	0.42	**0.57**
Backcast	0.43	0.37	0.41	0.16	0.32	0.36	0.23	0.21	0.41	0.32	0.07	0.14	**0.29**

Regional abbreviations: NE - North East England, YM - Yorkshire and Humber, EM - East Midlands, EE - East of England, LON - London, SE - South East England, SW - South West England, WM - West Midlands, NW - North West England, WA - Wales, SCOT - Scotland, NI - Northern Ireland.

Table 5
RMSFE (multiplied by 100) for growth in hours worked, from an AR(1) and five MF-VAR models differing in whether they impose the cross-sectional aggregation constraint (2) on output (Q) and/or labor (L).

	NE	YH	EM	EE	LON	SE	SW	WM	NW	WA	SCOT	NI	Average
AR(1) model													
Forecast	2.10	3.22	2.39	2.91	2.28	2.98	2.88	2.36	3.77	2.68	2.34	2.51	**2.70**
Nowcast	0.97	1.01	1.10	1.20	0.80	0.93	0.94	1.06	1.07	0.96	1.05	1.09	**1.01**
MF-VAR model - (with both aggregation constraints)													
Forecast	1.92	1.85	1.79	1.90	1.59	1.84	1.75	1.89	2.00	1.69	1.82	1.87	**1.82**
Nowcast	0.94	1.05	1.16	1.35	0.77	0.98	0.97	1.09	1.05	0.94	1.10	1.21	**1.05**
MF-VAR model - (aggregation constraint only in Q)													
Forecast	1.93	1.87	1.80	1.91	1.56	1.82	1.77	1.89	2.00	1.71	1.83	1.87	**1.83**
Nowcast	0.95	1.02	1.06	1.18	0.76	0.95	0.93	1.04	1.03	0.92	1.03	1.07	**0.99**
MF-VAR model - (aggregation constraint only in L)													
Forecast	2.06	1.80	1.85	1.99	1.70	1.88	1.78	1.90	2.03	1.71	1.85	1.95	**1.87**
Nowcast	1.12	1.03	1.18	1.33	0.79	0.98	0.96	1.10	1.06	0.96	1.11	1.23	**1.07**
MF-VAR model - (No aggregation constraints													
Forecast	2.09	1.86	1.87	2.02	1.74	1.93	1.85	1.92	2.08	1.74	1.89	1.97	**1.91**
Nowcast	1.35	1.23	1.44	1.62	0.92	1.23	1.14	1.33	1.24	1.10	1.29	1.35	**1.27**
MF-VAR model - (with both aggregation constraints but no exogenous predictors)													
Forecast	1.94	1.87	1.83	1.97	1.74	1.94	1.77	1.96	2.04	1.76	1.81	1.89	**1.88**
Nowcast	0.93	1.07	1.15	1.39	0.87	1.04	0.98	1.12	1.08	0.98	1.09	1.17	**1.07**

Regional abbreviations: NE - North East England, YM - Yorkshire and Humber, EM - East Midlands, EE - East of England, LON - London, SE - South East England, SW - South West England, WM - West Midlands, NW - North West England, WA - Wales, SCOT - Scotland, NI - Northern Ireland.

established that the cross-sectional aggregation constraints, particularly the one associated with output growth, led to substantial improvements in nowcast performance. In this subsection, we use probability integral transforms (PITs) to investigate the calibration of the best-performing model: the MF-VAR, with both cross-sectional aggregation constraints imposed. Figs. 2 through 4 plot the cumulative distribution functions of the PITs for our backcasts, nowcasts, and forecasts, respectively. If the forecasts were perfectly calibrated, we would see the empirical

Table 6
CRPS (multiplied by 100) for growth in hours worked, from an AR(1) and five MF-VAR models differing in whether they impose the cross-sectional aggregation constraint (2) on output (Q) and/or labor (L).

	NE	YH	EM	EE	LON	SE	SW	WM	NW	WA	SCOT	NI	Average
AR(1) model													
Forecast	1.06	2.10	1.53	1.33	1.19	1.83	1.74	1.29	2.17	1.70	1.49	1.43	**1.57**
Nowcast	0.39	0.41	0.41	0.43	0.33	0.38	0.36	0.38	0.41	0.41	0.40	0.41	**0.39**
MF-VAR model - (with both aggregation constraints)													
Forecast	0.87	0.90	0.80	0.77	0.66	0.82	0.74	0.78	0.91	0.92	0.82	0.84	**0.82**
Nowcast	0.42	0.46	0.46	0.51	0.29	0.41	0.37	0.41	0.43	0.47	0.46	0.51	**0.43**
MF-VAR model - (aggregation constraint only in Q)													
Forecast	0.88	0.92	0.79	0.77	0.63	0.82	0.75	0.79	0.91	0.94	0.82	0.83	**0.82**
Nowcast	0.43	0.43	0.39	0.42	0.28	0.38	0.33	0.36	0.40	0.45	0.41	0.42	**0.39**
MF-VAR model - (aggregation constraint only in L)													
Forecast	0.92	0.88	0.82	0.80	0.70	0.85	0.75	0.79	0.94	0.94	0.83	0.86	**0.84**
Nowcast	0.52	0.46	0.47	0.50	0.29	0.41	0.37	0.42	0.45	0.48	0.47	0.51	**0.45**
MF-VAR model - (No aggregation constraints)													
Forecast	0.94	0.92	0.81	0.81	0.69	0.86	0.78	0.80	0.95	0.96	0.85	0.86	**0.85**
Nowcast	0.60	0.54	0.55	0.58	0.36	0.50	0.44	0.50	0.53	0.55	0.53	0.54	**0.52**
MF-VAR model - (with both aggregation constraints but no exogenous predictors)													
Forecast	0.89	0.90	0.82	0.80	0.72	0.90	0.76	0.83	0.94	0.98	0.83	0.85	**0.85**
Nowcast	0.45	0.52	0.50	0.60	0.42	0.50	0.43	0.48	0.50	0.53	0.51	0.55	**0.50**

Regional abbreviations: NE - North East England, YM - Yorkshire and Humber, EM - East Midlands, EE - East of England, LON - London, SE - South East England, SW - South West England, WM - West Midlands, NW - North West England, WA - Wales, SCOT - Scotland, NI - Northern Ireland.

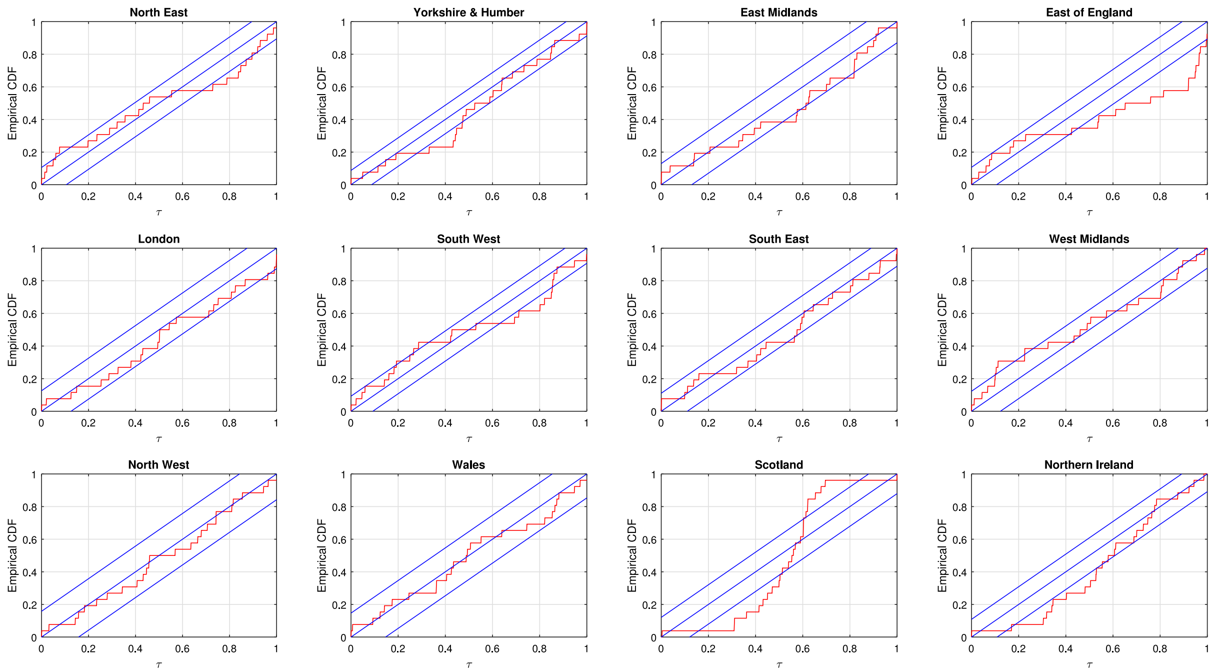

Fig. 2. Empirical cumulative distribution function of the PITs for the backcasts and Rossi and Sekhposyan (2019) 90% bands for the test of correct specification.

distribution function for the PITs on the 45-degree line. We have only 25 observations in our backcast evaluation period and 26 in our nowcast and forecast evaluation period. So, with such a small sample size, we can expect deviations from the 45-degree line even from a well-calibrated model. Following Rossi and Sekhposyan (2019), we plot 90% confidence bands around the 45-degree line to account for sample uncertainty.[21] With the small-sample qualification in mind, visually, the PITs plots indicate that calibration is quite good for most of

21 These bands should be interpreted as providing general guidance since they are derived assuming a rolling window of estimation while we use an expanding estimation window.

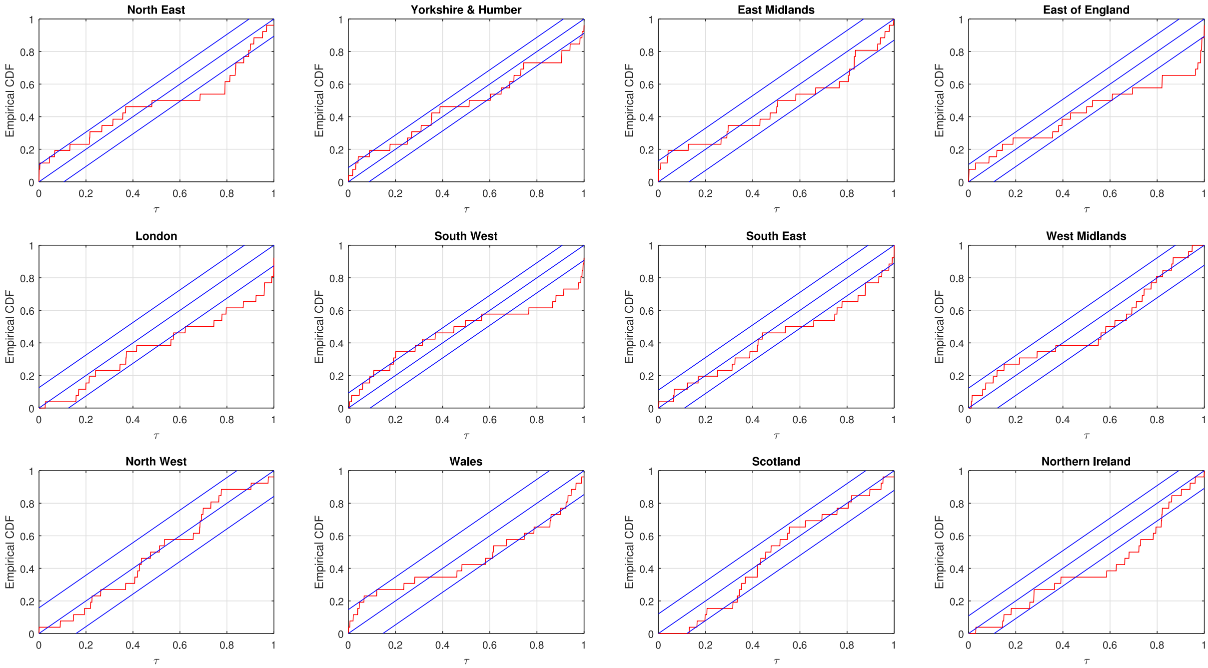

Fig. 3. Empirical cumulative distribution function of the PITs for the nowcasts and Rossi and Sekhposyan (2019) 90% bands for the test of correct specification.

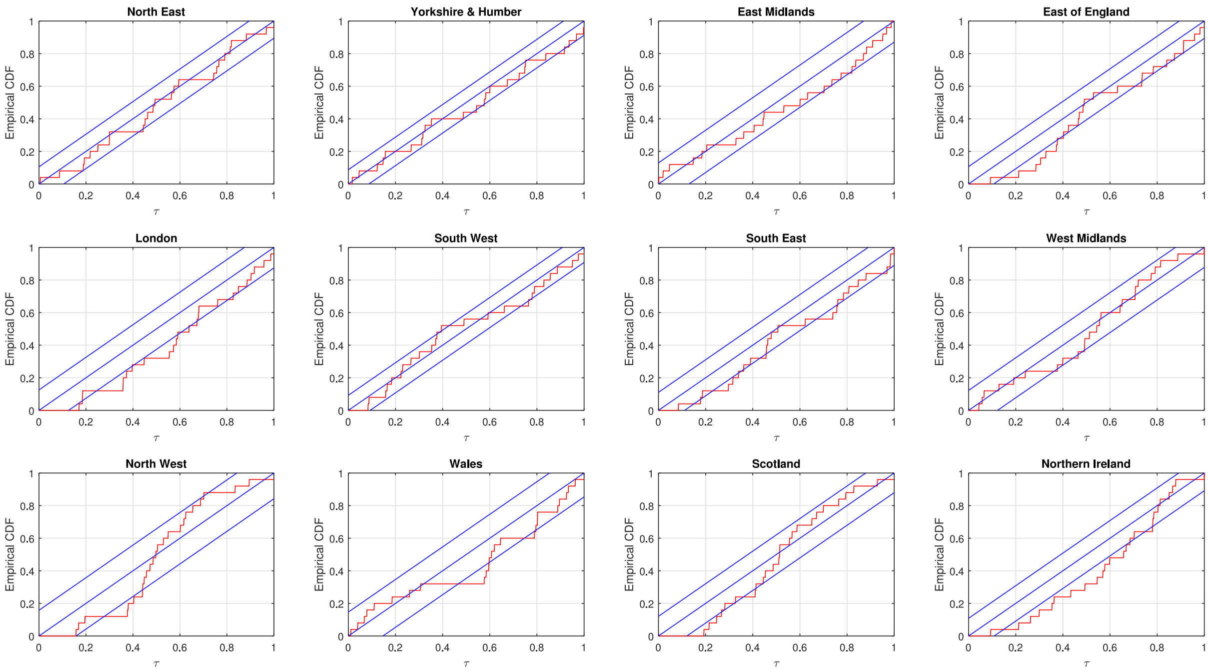

Fig. 4. Empirical cumulative distribution function of the PITs for the forecasts and Rossi and Sekhposyan (2019) 90% bands for the test of correct specification.

the regions. However, for several regions, we see that the empirical distribution function is somewhat flatter than the 45-degree line, particularly for the backcasts and nowcasts, indicating that we are underestimating uncertainty. It is well-established that VB methods can often underestimate posterior and, thus, predictive variances; e.g., see Giordano et al. (2018) for a general discussion and Gefang et al. (2022) for an investigation of this issue in VARs. This could be partially explaining some of these findings. However, this feature of the PITs plots does not happen universally, so it is hard to extract a general message. For example, looking at the backcasts for Scotland (see Fig. 2), we see a tendency for our model to overestimate uncertainty. The data for Scotland are constructed on a slightly different basis to those for the English regions, which might help explain these apparent differences.

6. Conclusions

This paper has taken on the challenge of nowcasting regional productivity growth in the UK at the quarterly frequency. Productivity is not something that is directly measured but is the ratio of a measure of output (Q) to a measure of labor (L), and it is Q and L that are directly measured. The characteristics of UK data for these two variables (i.e., that they change in frequency over time and that this change occurs at different times and that they have different release schedules and are measured subject to various errors) necessitates the development of a high-dimensional mixed frequency econometric model that accommodates these features. In this paper, we have developed such a model and derived a VB algorithm that allows for estimating and forecasting with the model.

We have explored the contribution that different features of our modeling approach make in a pseudo real–time empirical exercise, comparing backcast, nowcast, and forecast accuracy across five different versions of our model, each reflecting different combinations of our hierarchical (cross-sectional) aggregation constraints, as well as the contribution of additional regional-level predictors. We demonstrated the importance that hierarchical aggregation constraints, imposed in stochastic form, play in producing more accurate estimates of sub-national output, hours worked, and productivity growth, when the national aggregate was known (as it is in our 'nowcasts' and 'backcasts') but not as clearly when the aggregate is also being forecast.

For sub-national policymakers seeking to close the information gap that exists between national-level understanding of economic fluctuations and changes in the local economy, the methods set out in this paper provide a good way of producing timely regional growth estimates consistent with the national–level aggregate. Given our focus in this paper on producing more timely sub-national estimates of productivity growth, and exploring the role of national–level data in improving these estimates, we have not explored many features that might improve the national–level forecasting and which might be expected, via the cross-sectional aggregation constraint, to improve our forecast of the national aggregate. We leave this to future work, but note that the methods set out in this paper should help translate improvements in forecasting national growth into improved regional estimates via the aggregation constraint in the same way as we saw happen in our empirical exercise when the aggregate was known.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2022.04.002.

References

Athanasopoulos, George, Gamakumara, Puwasala, Panagiotelis, Anastasios, Hyndman, Rob J., & Affan, Mohamed (2020). Hierarchical forecasting. In Peter Fuleky (Ed.), *Advanced studies in theoretical and applied econometrics*, *Macroeconomic forecasting in the era of big data* (1st). (pp. 689–719). Springer, http://dx.doi.org/10.1007/978-3-030-31150-6_21.

Brave, Scott A., Butters, R. Andrew, & Justiniano, Alejandro (2019). Forecasting economic activity with mixed frequency BVARs. *International Journal of Forecasting*, *35*(4), 1692–1707. http://dx.doi.org/10.1016/j.ijforecast.2019.02.010.

Di Fonzo, Tommaso (1990). The estimation of M disaggregate time series when contemporaneous and temporal aggregates are known. *The Review of Economics and Statistics*, *72*(1), 178–182. http://dx.doi.org/10.2307/2109758.

Diebold, Francis X., & Mariano, Robert S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *20*(1), 134–144. http://dx.doi.org/10.1198/073500102753410444.

Eckert, Florian, Hyndman, Rob J., & Panagiotelis, Anastasios (2021). Forecasting Swiss exports using Bayesian forecast reconciliation. *European Journal of Operational Research*, *291*(2), 693–710. http://dx.doi.org/10.1016/j.ejor.2020.09.046.

Eraker, Bjørn, Chiu, Ching Wai (Jeremy), Foerster, Andrew T., Kim, Tae Bong, & Seoane, Hernán D. (2015). Bayesian mixed frequency VARs. *Journal of Financial Econometrics*, *13*(3), 698–721. http://dx.doi.org/10.1093/jjfinec/nbu027.

Frale, Cecilia, Marcellino, Massimiliano, Mazzi, Gian Luigi, & Proietti, Tommaso (2011). EUROMIND: A monthly indicator of the euro area economic conditions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *174*(2), 439–470. http://dx.doi.org/10.1111/j.1467-985X.2010.00675.x.

Gefang, Deborah (2014). Bayesian doubly adaptive elastic-net lasso for VAR shrinkage. *International Journal of Forecasting*, *30*(1), 1–11. http://dx.doi.org/10.1016/j.ijforecast.2013.04.004.

Gefang, Deborah, Koop, Gary, & Poon, Aubrey (2022). Forecasting using variational Bayesian inference in large vector autoregressions with hierarchical shrinkage. *International Journal of Forecasting*, http://dx.doi.org/10.1016/j.ijforecast.2021.11.012.

George, Edward I., Sun, Dongchu, & Ni, Shawn (2008). Bayesian stochastic search for VAR model restrictions. *Journal of Econometrics*, *142*(1), 553–580. http://dx.doi.org/10.1016/j.jeconom.2007.08.017.

Giacomini, Raffaella, & White, Halbert (2006). Tests of conditional predictive ability. *Econometrica*, *74*(6), 1545–1578. http://dx.doi.org/10.1111/j.1468-0262.2006.00718.x.

Giordano, Ryan, Broderick, Tamara, & Jordan, Michael I. (2018). Covariances, robustness and variational Bayes. *Journal of Machine Learning Research*, *19*(1), 1981–2029.

Guerrero, Victor, & Nieto, Fabio (1999). Temporal and contemporaneous disaggregation of multiple time series. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*, *8*, 459–489. http://dx.doi.org/10.1007/BF02595880.

Kastner, Gregor, & Huber, Florian (2020). Sparse Bayesian vector autoregressions in huge dimensions. *Journal of Forecasting*, *39*(7), 1142–1165. http://dx.doi.org/10.1002/for.2680.

Koop, Gary (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, *28*(2), 177–203. http://dx.doi.org/10.1002/jae.1270.

Koop, Gary, McIntyre, Stuart, & Mitchell, James (2020). UK regional nowcasting using a mixed frequency vector auto-regressive model with entropic tilting. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *183*(1), 91–119. http://dx.doi.org/10.1111/rssa.12491.

Koop, Gary, McIntyre, Stuart, Mitchell, James, & Poon, Aubrey (2020a). Reconciled estimates and nowcasts of regional output in the UK. *National Institute Economic Review*, *253*, 44–59. http://dx.doi.org/10.1017/nie.2020.29.

Koop, Gary, McIntyre, Stuart, Mitchell, James, & Poon, Aubrey (2020b). Regional output growth in the United Kingdom: More timely and higher frequency estimates from 1970. *Journal of Applied Econometrics*, *35*(2), 176–197. http://dx.doi.org/10.1002/jae.2748.

Korobilis, Dimitris (2013). VAR Forecasting using Bayesian variable selection. *Journal of Applied Econometrics*, *28*(2), 204–230. http://dx.doi.org/10.1002/jae.1271.

Mariano, Roberto S., & Murasawa, Yasutomo (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, *18*(4), 427–443. http://dx.doi.org/10.1002/jae.695.

McCracken, Michael W., Owyang, Michael T., & Sekhposyan, Tatevik (2021). Real-time forecasting and scenario analysis using a large mixed-frequency Bayesian VAR. *International Journal of Central Banking*, *17*, 327–367, https://www.ijcb.org/journal/ijcb21q5a8.htm.

Panagiotelis, Anastasios, Athanasopoulos, George, Gamakumara, Puwasala, & Hyndman, Rob J. (2021). Forecast reconciliation: A geometric view with new insights on bias correction. *International Journal of Forecasting*, *37*(1), 343–359. http://dx.doi.org/10.1016/j.ijforecast.2020.

Proietti, Tommaso (2011a). Estimation of common factors under cross-sectional and temporal aggregation constraints. *International Statistical Review*, *79*(3), 455–476. http://dx.doi.org/10.1111/j.1751-5823.2011.00152.x.

Proietti, Tommaso (2011b). Multivariate temporal disaggregation with cross-sectional constraints. *Journal of Applied Statistics*, *38*(7), 1455–1466. http://dx.doi.org/10.1080/02664763.2010.505952.

Rossi, Barbara, & Sekhposyan, Tatevik (2019). Alternative tests for correct specification of conditional predictive densities. *Journal of Econometrics*, *208*(2), 638–657. http://dx.doi.org/10.1016/j.jeconom.2018.07.

Schorfheide, Frank, & Song, Dongho (2015). Real-time forecasting with a mixed-frequency VAR. *Journal of Business & Economic Statistics*, *33*(3), 366–380. http://dx.doi.org/10.1080/07350015.2014.954707.

Zou, Hui (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, *101*(476), 1418–1429. http://dx.doi.org/10.1198/016214506000000735.