Received: 10 June 2021 | Revised: 13 December 2021 | Accepted: 3 January 2022

DOI: 10.1002/for.2851

RESEARCH ARTICLE

WILEY

Uncertainty and forecastability of regional output growth in the UK: Evidence from machine learning

Mehmet Balcilar[1,2] | David Gabauer[3] | Rangan Gupta[4] | Christian Pierdzioch[5]

[1]Department of Economics, Eastern Mediterranean University, Turkish Republic of Northern Cyprus, Via Mersin 10, Famagusta, 99628, Turkey

[2]Department of Economics, OSTIM Technical University, Ankara, 06374, Turkey

[3]Data Analysis Systems, Software Competence Center Hagenberg, Hagenberg, Austria

[4]Department of Economics, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa

[5]Department of Economics, Helmut Schmidt University, Holstenhofweg 85, PO Box 700822, Hamburg, 22008, Germany

Correspondence
Christian Pierdzioch, Department of Economics, Helmut Schmidt University, Holstenhofweg 85, PO Box 700822, 22008 Hamburg, Germany.
Email: macroeconomics@hsu-hh.de

Abstract

Utilizing a machine learning technique known as random forests, we study whether regional output growth uncertainty helps to improve the accuracy of forecasts of regional output growth for 12 regions of the UK using monthly data for the period from 1970 to 2020. We use a stochastic volatility model to measure regional output growth uncertainty. We document the importance of interregional stochastic volatility spillovers and the direction of the transmission mechanism. Given this, our empirical results shed light on the contribution to forecast performance of own uncertainty associated with a particular region, output growth uncertainty of other regions, and output growth uncertainty as measured for London as well. We find that output growth uncertainty significantly improves forecast performance in several cases, where we also document cross-regional heterogeneity in this regard.

KEYWORDS

forecasting, machine learning, regional output growth, uncertainty, UK

1 | INTRODUCTION

Theoretically, the effect of uncertainty on economic activity is generally explained by the "real option theory" (see, e.g., Bernanke, 1983; Dixit & Pindyck, 1994; Pindyck, 1991; and more recently, Bloom, 2009), which suggests that decision making is affected by uncertainty because the latter raises the option value of waiting. In other words, given that the costs associated with wrong investment decisions are high, uncertainty makes firms and, in the case of durable goods, also consumers more cautious. As a result, firms and consumers postpone investment, hiring, and consumption decisions to periods of lower uncertainty (which results in cyclical fluctuations in macroeconomic aggregates). Hence, uncertainty is expected to negatively impact overall output (besides consumption and investment).[1] In the wake of the

[1]It should be noted that there are two potential additional channels through which uncertainty can unfold a positive effect on the economy. The first channel rests on the theory of "growth options." This theory stresses that uncertainty can encourage investment because the upside when uncertainty is resolved can be high, although there is a limited downside (Kraft et al., 2018; Segal et al., 2015). The second channel is called the "Oi–Hartman–Abel" effect and traced back to the works of Oi (1961), Hartman (1972), and Abel (1983). According to this channel, an increase in uncertainty implies an increase in both potential good outcomes and bad outcomes. When firms are able to contract works as an insurance against bad outcomes, an asymmetry emerges because increased risk is looked upon positively. This asymmetry makes firms

investing in large capacity because it will make them able to take advantage of potential positive news. At the same time, if the news is bad, firms will, with low effort, be able to scale back. To sum up, there exist competing effects of uncertainty on economic activity, and the overall effect would be contingent on the relative strength of these effects, though it is generally expected that the real option theory will dominate, and so uncertainty, on balance, has a negative impact on output, as observed during the global financial crisis and the recent outbreak of the COVID-19 pandemic.

"Great Recession" and more recently the COVID-19 pandemic, the large empirical literature (see Al-Thaqeb & Algharabali, 2019; Caggiano et al., 2020; Castelnuovo et al., 2017; Gupta et al., 2018, 2019, 2020, 2021, for detailed reviews) that has emerged involving the impact of uncertainty on output has overwhelmingly confirmed the negative association between these two variables as outlined in theory.

Although the literature dealing with the influence of uncertainty on output primarily relies on in-sample-based structural analyses, more recently, quite a few studies (see e.g., Aye et al., 2019a, 2019b; Balcilar et al., 2016; Gupta et al., forthcoming; Junttila & Vataja, 2018; Karnizova & Li, 2014; Pierdzioch & Gupta, 2020; Salisu et al., 2022; Segnon et al., 2018) have also analyzed the role of uncertainty in forecasting economic activity (output growth and recessions) in out-of-sample analyses. This is an important line of research, because policymakers in general, and central banks in particular, would need accurate predictions of the future path of the economy following periods of heightened uncertainty while making their policy decisions. Understandably, precise forecasting of the macroeconomy is also important for investors. Finally, because in-sample predictability might not translate into forecasting gains, and the ultimate test of any predictive model (in terms of econometric methodologies and the predictors being used) is primarily considered to be in its out-of-sample performance (Campbell, 2008), this area also forms a pertinent question for academic researchers.

Against this backdrop, the objective of this research is to analyze the forecasting ability of uncertainty for output growth in the UK, but from a regional perspective. In particular, we look at 12 regions of the UK (viz., East Midlands, East of England, London, North East, North West, Northern Ireland, Scotland, South East, South West, Wales, West Midlands, Yorkshire, and the Humber) over the quarterly period from February 1970 to February 2020. In the process, we not only study the predictive role of uncertainty associated with a particular region but also incorporate the effect of uncertainty of the other regions, given the widespread evidence of international uncertainty spillovers (see, e.g., Antonakakis et al., 2018, 2019; Christou et al., 2020a; Gabauer & Gupta, 2018; Gupta et al., 2016) and evidence of which we also provide in our particular dataset. We also control for other standard aggregate macroeconomic predictors (inflation rate, financial stress, and interest rate), as well as lagged values of the growth rate of the specific region under investigation and the other regions, which have also been shown to depict interconnectedness (Koop et al., 2020a).

At this stage, we must point out that Junttila and Vataja (2018), Aye et al. (2019b), Gupta et al. (forthcoming), and Salisu et al. (2022) have highlighted the important role played by uncertainty in forecasting alternative measures of the performance of the aggregate real economy of the UK, but our paper makes the first attempt to analyze the forecastability of output growth due to uncertainty at the regional level, based on a newly constructed high-frequency (quarterly) novel dataset of regional gross value added (GVA) by Koop et al. (2020b, 2020c). As highlighted by Mumtaz (2018) and Mumtaz et al. (2018), based on in-sample analyses of state-level data for the USA, the impact of uncertainty is heterogeneous and depends on the underlying conditions of the regions at the time the uncertainty shock originates. Naturally, one cannot generalize the role of uncertainty for the aggregate economy to the various regions comprising the overall country, thus making our regional study of tremendous importance from the policy perspective for determining the nature and size of policy intervention to counteract the negative influence of an uncertainty shock, especially given the well-established heterogeneity involving business-cycle fluctuations and, in general, across regions of the UK (Barrios et al., 2003; Beenstock & Felsenstein, 2008). Note that, although we could have studied the states of the USA, which does indeed have widespread availability of regional data, and could indeed be an area of future research, our decision to look at the UK emanates from the persistent uncertainty witnessed by its regions ever since the Brexit referendum that took place in (23rd) June 2016, besides the impact of the global financial and European sovereign debt crises that took place earlier. Hence, the UK, which has witnessed waves of crises including the current coronavirus episode, forms an interesting case study of the uncertainty-growth nexus.

As far as the econometric approach is concerned, we rely on a machine learning approach known as random forests (Breiman, 2001), which, in turn, has two main advantages in the context of our analysis. First, random forests can accurately and flexibly analyze the links between regional GVA growth and a large number of predictors in a full-fledged data-driven manner. Second, random forests automatically capture potential nonlinear links between output growth and its predictors, including uncertainty, as shown to exist historically for the UK by

Christou et al. (2020b) and Bredin et al. (2021),[2] as well as any interaction effects between the predictors.

We structure the remainder of this research as follows. In Section 2, we briefly describe how a random forest is grown. In Section 3, we describe our data and report our empirical results. Finally, in Section 4, we conclude with final remarks.

2 | RANDOM FORESTS

A random forest consists of a large number of individual regression trees (see Hastie et al., 2009, for a textbook exposition; our notation follows theirs). A regression tree, T, in turn, consists of branches that subdivide the space of predictors, $\mathbf{x}=(x_1,x_2,\ldots)$, of the regional output growth rate (in the following: regional output growth, for short) into l nonoverlapping regions, R_l. These regions are computed by applying a search-and-split algorithm in a recursive top-down fashion.

Application of this search-and-split algorithm to grow a regression tree requires starting at the top level of the tree, iterating over the various predictors, s, and the all possible splitting points, p, that can be formed using the data on a predictor. For every combination of a predictor and a splitting point, the search-and-split algorithm forms two half planes, $R_1(s,p)=\{x_s|x_s\leq p\}$ and $R_2(s,p)=\{x_s|x_s>p\}$ so as to minimize the standard squared-error loss criterion:

$$\min_{s,p}\left\{\min_{\overline{RG}_1}\sum_{x_s\in R_1(s,p)}\left(RG_i-\overline{RG}_1\right)^2+\min_{\overline{RG}_2}\sum_{x_s\in R_2(s,p)}\left(RG_i-\overline{RG}_2\right)^2\right\}, \tag{1}$$

where the index i denotes those observations on regional output growth, RG, that belong to a half plane, and $\overline{RG}_k=\text{mean}\{RG_i|x_s\in R_k(s,p)\},k=1,2$ denotes the half plane-specific mean of regional output growth. The objective function given in Equation (1), thus, requires (i) searching over all combinations of s and p, and, (ii) for any given combination of s and p, minimizing the half plane-specific squared-error loss by an optimal choice of the half plane-specific means of regional output growth. The solution of this minimization problem gives the top-level optimal splitting predictor and optimal splitting point, and the two $\overline{RG}_k$. The resulting simple regression tree has two terminal nodes.

In order to grow a larger tree, the next step of the search-and-split algorithm requires to carry out the minimization problem in Equation (1) for the two top-level half planes, $R_1(s,p)$ and $R_2(s,p)$, yielding up to two second-level optimal splitting predictors and optimal splitting points, and four second-level region-specific means of regional output growth. Solving the minimization problem over and over again gives an increasingly complex regression tree. Finally, the search-and-split algorithm is terminated when a regression tree has a preset maximum number of terminal nodes or every terminal node has a minimum number of observations. In our empirical research, we cross-validate a technique to determine the optimal minimum number of observations per terminal node (see Section 3.2 for details).

Once the search partition algorithm has stopped, the regression tree sends the predictors of regional output growth from its top level to its leaves along the optimal partitioning points (i.e., the nodes of the tree) and branches. A forecast of regional output growth can then be computed by its region-specific mean. For a regression tree made up of L regions, this forecast is formed as follows ($\mathbf{1}$ denotes the indicator function):

$$T\left(\mathbf{x}_i,\{R_l\}_1^L\right)=\sum_{l=1}^{L}\overline{RG}_l\mathbf{1}(\mathbf{x}_i\in R_l). \tag{2}$$

The search-and-split algorithm can be used in principle to grow an increasingly complex regression tree. However, the resulting complex hierarchical structure of a regression tree gives rise to an overfitting and data-sensitivity problem and, thereby, implies that forecasting performance deteriorates. It is at this stage that a random forest enters the scene. A random forest solves the overfitting problem in two steps. In the first step, a large number of bootstrap samples (sampling with replacement) are drawn from the data. In the second step, a random regression tree is fitted to every bootstrap sample. Such a random regression tree differs from a classic regression tree in that for every splitting step, only a random subset of the predictors is being used. In this way, a random regression tree mitigates the effect of influential predictors on tree building. Moreover, growing a large number of random trees lowers the correlation of forecasts from the individual trees. Finally, averaging the decorrelated forecasts computed by means of the individual random regression trees stabilizes the forecasts of realized output growth.

[2]For a detailed review of the international literature on the nonlinearity between uncertainty and economic activity, the reader is referred to Caggiano et al. (2021). For a recent application of random forests in economics along with a discussion of their advantages, see Bouri et al. (2021). For a detailed analysis of the pros and cons of random forests and other machine learning techniques, see also the comprehensive discussion in the textbook by Hastie et al. (2009).

3 | EMPIRICAL ANALYSIS

3.1 | Data

The annualized GVA growth of the regions (East Midlands, East of England, London, North East, North West, Northern Ireland, Scotland, South East, South West, Wales, West Midlands, Yorkshire, and the Humber) is obtained from the nowcasting project of Koop et al. (2020b, 2020c) associated with the Economic Statistics of the Centre of Excellence.[3] Koop et al. (2020b, 2020c) develop a mixed-frequency vector autoregressive (MF-VAR) model and use it to produce estimates of quarterly regional output growth. Temporal and cross-sectional restrictions are imposed in the model to ensure that the quarterly regional estimates are consistent with the annual regional observations and the observed quarterly UK totals. Koop et al. (2020b, 2020c) use a machine learning method based on the hierarchical Dirichlet–Laplace prior to ensure optimal shrinkage and parsimony in the overparameterized MF-VAR. Because this dataset is available from February 1970 onward, our analysis starts from this period and ends in February 2020, based on data availability at the time of writing of this paper.

It is important to emphasize that uncertainty is a latent variable, and hence, one requires ways to measure it. In this regard, besides the various alternative metrics of uncertainty associated with financial markets (such as the implied volatility indices, realized volatility, idiosyncratic volatility of equity returns, and corporate spreads), there are primarily three broad approaches to quantify uncertainty: (i) The main idea behind the news-based approach is to search newspapers for terms associated with economic and policy uncertainty and, based on the results of this search, to construct indices of uncertainty. (ii) Another approach is to extract uncertainty from stochastic volatility (SV) estimates of various types of small and large-scale structural models analyzed in the macroeconomics and finance literature. (iii) A third approach is to construct measures of uncertainty based on professional forecaster disagreements. As for our metric of uncertainty, motivated by the recent work on the nexus between growth and growth uncertainty by Balcilar and Ozdemir (2020), we use the second approach, because the first and third approaches are not applicable in the context of our analysis due to unavailability of the corresponding regional data. Hence, our measure of regional uncertainty is derived from SV estimates of the regional output growth. Although we could have also relied upon generalized autoregressive conditional heteroskedasticity (GARCH) models, which have a deterministic volatility-generating mechanism, we prefer the SV approach because it does not impose restrictions on conditional moments (as in GARCH models). In addition, SV models have also been shown in earlier literature to produce a better in-sample fit as well as superior out-of-sample forecasts of volatility and, hence, uncertainty (as discussed in detail by Balcilar & Ozdemir, 2020). In particular, building on the work by Kastner and Frühwirth-Schnatter (2014), given observed growth rates for a particular region denoted by $y=(y_1,y_2,...,y_T)'$, we specify the SV model as follows: $y_t=e^{h_t/2}\epsilon_t$, with $h_t=\mu+\psi(h_{t-1}-\mu)+\sigma\nu_t$, where the i.i.d. standard normal innovations, ϵ_t and ν_s, are by assumption independent for $t,s\in\{1,\dots,T\}$. The latent process, $h=(h_0,h_1,...,h_T)$, appearing in the state equation is usually interpreted as the latent time-varying volatility process (i.e., our measure of uncertainty), with the initial state being distributed according to the stationary distribution, that is, $h_0|\mu,\psi,\sigma\sim\mathcal{N}(\mu,\sigma^2/(1-\psi^2))$. Because model reparameterization often helps to improve simulation efficiency in state-space models, that is, a centered parameterization has several disadvantages, we use, like Kastner and Frühwirth-Schnatter (2014), the (fully) noncentered parameterization of the model given by $y_t\sim\mathcal{N}(0,\omega e^{\sigma\tilde{h}_t})$, with $\tilde{h}_t=\psi\tilde{h}_{t-1}+\nu_t$, where $\nu_t\sim\mathcal{N}(0,1)$, where $\omega=e^{\mu}$. The initial value of $\tilde{h}_0|\psi$ is drawn from the stationary distribution of the latent process, that is, $\tilde{h}_0|\psi\sim\mathcal{N}(0,1/(1-\psi^2))$, and note that $\tilde{h}_t=(h_t-\mu)/\sigma$. Figure 1 shows the estimated regional stochastic volatilities.

Our forecasting exercise also includes Consumer Price Index (CPI)-based annualized inflation rate, with the CPI data obtained from the Main Economic Indicators (MEI) Database of the Organisation for Economic Co-operation and Development (OECD). To measure the stance of monetary policy, we consider the official bank rate derived from the Bank of England (BoE) until 1989, and then we use the shadow short rate (SSR) developed by Wu and Xia (2016) from 1990 onwards,[4] given that our period of analysis involves the zero lower bound (ZLB) scenario in the wake of the Great Recession and the global financial crisis, and more recently following the outbreak of the coronavirus in 2020. Given that a range of unconventional monetary policies (such as large-scale asset purchases, a maturity extension program, and efforts of forward guidance in order to manage expectations of a prolonged period of low policy rates) are pursued during the ZLB situations, we would need a

[3]The data are downloadable from https://www.escoe.ac.uk/regionalnowcasting/.

[4]The data are available for download from the website of Professor Jing Cynthia Wu at https://sites.google.com/view/jingcynthiawu/shadow-rates?authuser=0.

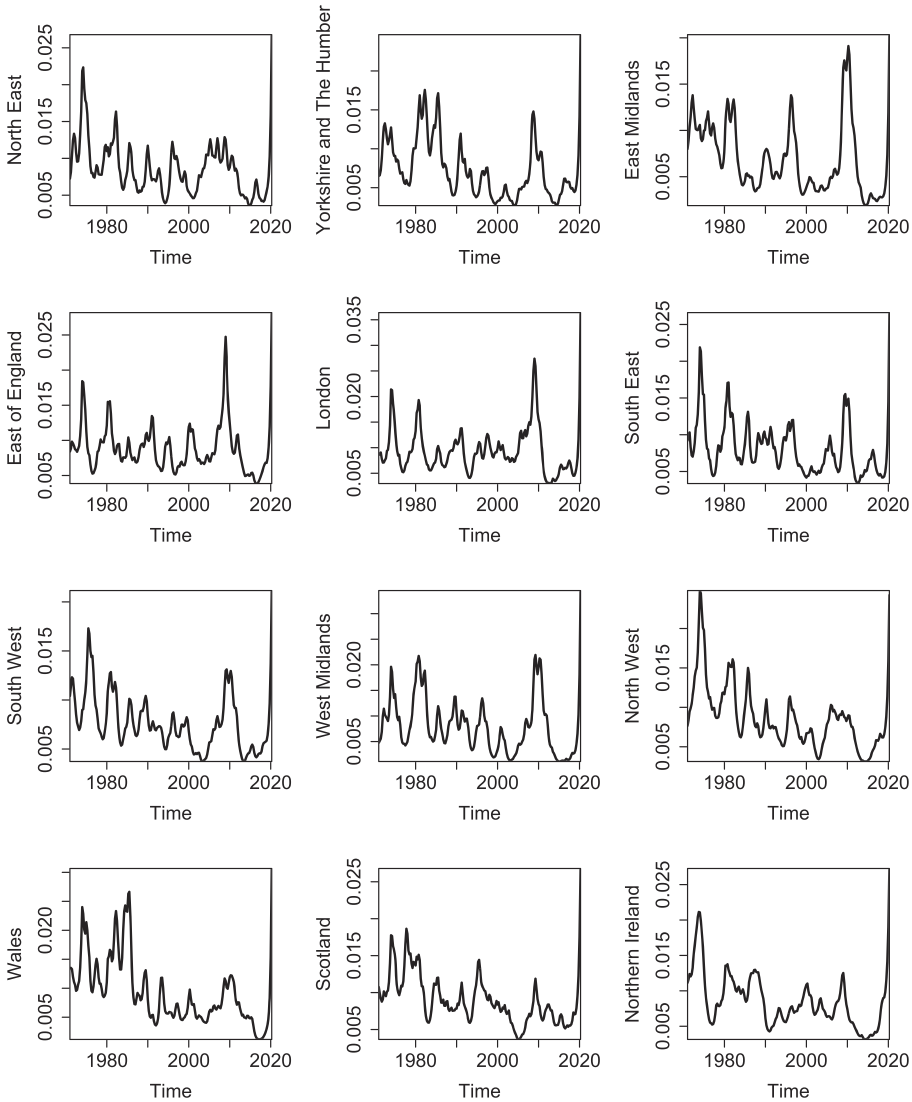

FIGURE 1 Regional stochastic volatilities

uniform and coherent measure of the monetary policy stance. Thus, we use the SSR, which measures the nominal interest rate that would prevail in the absence of its effective lower bound.[5] Finally, we incorporate the information of the Financial Stress Index (FSI) derived from the Statistical Data Warehouse of the European Central Bank.[6] The index includes six market-based financial stress measures that capture returns and (realized) volatility of three financial market segments, that is, equity, bond, and foreign exchange. In addition, when aggregating the subindices, the FSI takes the comovement across market segments into account. The reader is referred to Duprey et al. (2017) for further details. Note that data that are available at higher monthly frequency are converted to quarterly values by taking 3-month averages.

[5]The SSR is based on models of the term structure, which essentially removes the effect that the option to invest in physical currency (at an interest rate of zero) has on yield curves, resulting in a hypothetical "shadow yield curve" that would exist if physical currency was not available. The process allows one to answer the question: "What policy rate would generate the observed yield curve if the policy rate could be taken negative?" The "shadow policy rate" generated in this manner, therefore, provides a measure of the monetary policy stance after the actual policy rate reaches zero. The main advantage of the SSR is that it is not constrained by the ZLB and thus allows us to combine the data from the ZLB period with that of the non-ZLB era and use it as the common metric of monetary policy stance across the conventional and unconventional monetary policy episodes.

[6]The data can be downloaded from https://sdw.ecb.europa.eu/quickview.do;jsessionid=D122B96CF06237259EFEBFB2ADCA10F0SERIES_KEY=383.CLIFS.M.GB._Z.4F.EC.CLIFS_CI.IDX.

3.2 | Empirical results

We carry out our empirical analysis by using the statistical computing program R (R Core Team, 2021), where we make use of the add-on package "grf" (Tibshirani et al., 2020). Our results are based on estimates of random forests for rolling-estimation windows of length 40, 60, and 80 quarters (i.e., 10, 15, and 20 years). Although shifting the rolling-estimation windows across the dataset, we optimize, by means of cross validation, the number of predictors randomly selected for splitting, the minimum node size of a tree, and the parameter that governs the maximum imbalance of a node, where we use 2000 regression trees to grow a random forest. We study three forecast horizons: one, two, and four quarters, where the target variable in case $h > 1$ is the arithmetic average of the regional output growth rates under scrutiny over the respective forecast horizon.

We estimate random forests for four different models. Model 1 features, in addition to the inflation rate, the monetary policy-related interest rate, and the FSI as proxies of monetary and financial conditions, as predictors only the own lagged regional output growth of a region along with the regional output growth of all other regions, given the evidence of spillovers of regional growth as shown by Koop et al. (2020a). Model 2 features the predictors of Model 1 plus the own SV of a region, capturing the associated uncertainty of that region. Comparing Models 1 and 2 sheds light on whether today's regional output growth uncertainty helps to improve the accuracy of forecasts of subsequent regional output growth. Model 3 features all predictors of Model 2 and, in addition, the regional stochastic volatilities of all other regions.

To motivate the formulation of Model 3, we would like to formally highlight the importance of interregional SV spillovers. In this regard, we utilize a full-fledged time-varying version of the spillover approaches of Diebold and Yilmaz (2009, 2012, 2014), as proposed based on a time-varying parameter–vector autoregressive (TVP-VAR) model by Antonakakis et al. (2020). This framework is based on the generalized forecast-error variance decomposition for a VAR, but the biggest drawback of the generalized spillover method is that it provides misleading information when it comes to aggregate spillover as the associated index is bounded between 0% and 100%, and so when a shock is introduced to the individual variable, it brings most of the variation in other factors than the factor to which shock was introduced. In light of this, we also use the joint spillover method by following Lastrapes and Wiesen (2021) capable in gauging the system-wide spillovers, as developed in a TVP-VAR context by Balcilar et al. (2020).

Both approaches provide qualitatively similar results, illustrating the robustness of our findings with respect to the spillover analysis. Figure 2 represents the dynamic total connectedness, which describes the average amount of shock spillover one series has to all others in the network. We see that the Antonakakis et al. (2020) results are constantly smaller in magnitude than those of

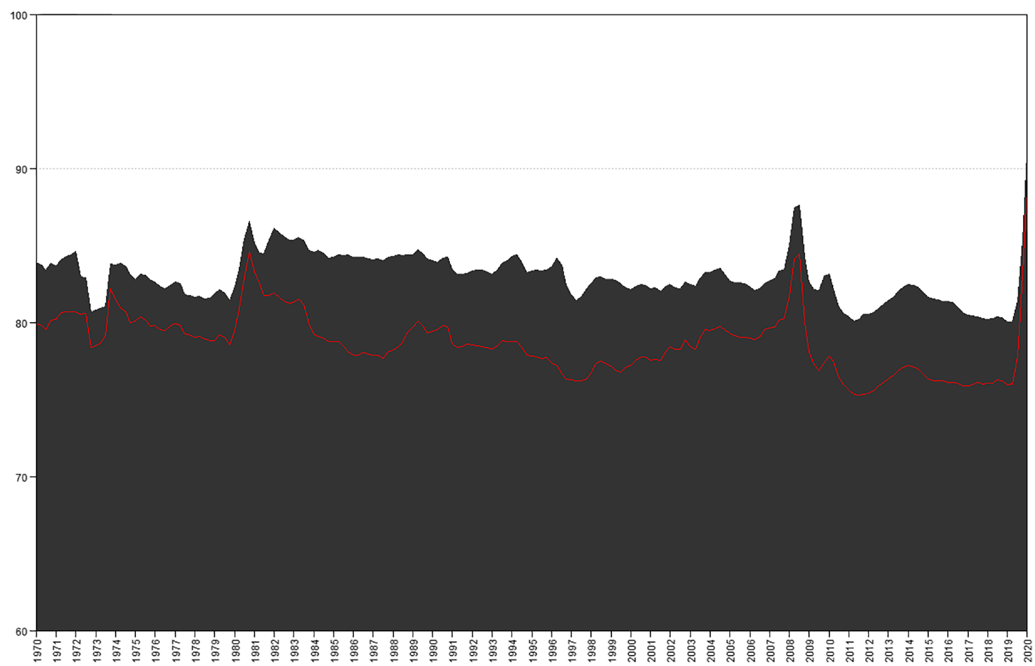

FIGURE 2 Dynamic total connectedness. Black area illustrates Balcilar et al. (2020), whereas red line demonstrates Antonakakis et al. (2020) results based upon a 20-quarter-ahead forecast horizon. Both approaches are based on a TVP-VAR with a lag length of as suggested by the Bayesian information criteria

Balcilar et al. (2020). Besides the fact that the high degree of dynamic total connectedness highlights the importance of uncertainty shock spillovers when it comes to regional UK output growth, it further points out significant economic events that had a substantial effect on its dynamic behavior such as the mid-1970s recessions that was marked by the 1973 oil crisis, and stagflation, as well as the early 1980s recession characterized by the transition from a manufacturing to a services economy and a period of considerable spending cuts. More recent dynamics cover the time of the global financial crisis, the European sovereign debt crisis, and the coronavirus pandemic that has spread over to the European continent in the beginning of 2020.

But even more to the point is the direction of the transmission mechanism as it lays out a more in-depth analysis of the regional shock propagation. Figure 3 depicts the relative strength of each region in a time-varying behavior underlining the significant and permanent effect the global financial crisis of 2009 had on most UK regional dynamics. In particular, regions such as London, East of England, North West, and Scotland decreased in its net transmission power until the end of the sample period. Furthermore, similar but less severe adjustments can be

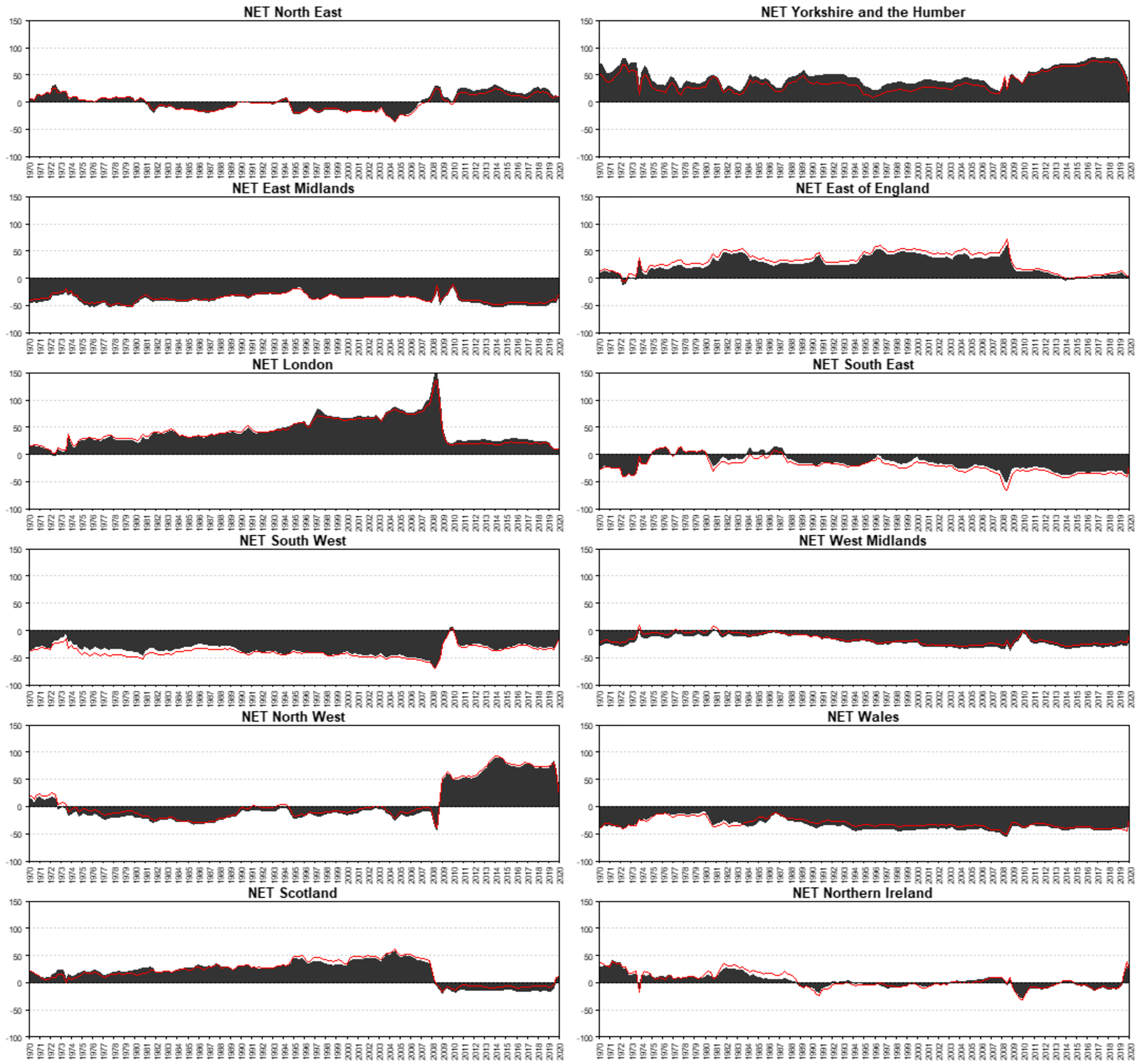

FIGURE 3 Net total directional connectedness measures. Black area illustrates Balcilar et al. (2020), whereas red line demonstrates Antonakakis et al. (2020) results based upon a 20-quarter-ahead forecast horizon. Both approaches are based on a TVP-VAR with a lag length of as suggested by the Bayesian information criteria

observed during the coronavirus pandemic. In general, our results reveal that Yorkshire and the Humber, East of England, and London have been permanent transmitters of shocks, whereas East Midlands, West Midlands, South West, and Wales have been permanent receivers of shocks. Two notable evolutions are that the North West has become an essential transmitter after 2009, whereas Scotland has become a receiver of shocks. It should also be mentioned that our findings indicate the importance of economic weight London has in the evolution of UK's regional uncertainty by its persistent net transmission characteristic, the unprecedented magnitude in its transmitting power prior the global financial crisis of 2009, and its still continuing—even though not as significant—role afterwards. Thus, this analysis shows that the UK regional SV spillovers are strong, as its dynamics explain between 75% and 90% of the evolution of uncertainty.

Going back to our Model 3, given the evidence of output growth volatility spillovers across regions, upon comparing Models 2 and 3, we can assess whether regional uncertainty spillovers onto other regions help *t*0 predict regional output growth of a particular region. Finally, Model 4 features the own SV of a region plus the SV of output growth estimated for London, given its importance as a transmitter of uncertainty shocks. When we compare Models 2 and 4, we can study the contribution of the capital city to forecasting regional output growth over and above the own uncertainty of a region.

Turning next to our out-of-sample forecasting analysis, Table 1 summarizes results for root-mean-square forecast-error (RMSFE) ratios. An RMSFE ratio larger than unity implies that the alternative model outperforms out of sample in terms of the RMSFE of the corresponding baseline model. The first bloc of results obtains when the baseline model features only regional output growth as predictors (Model 1), whereas the alternative model features, in addition, the own SV of a region (Model 2). We observe in general RMSFE ratios that exceed unity for Yorkshire and the Humber, East of England, Scotland, and Northern Ireland. Results for North East and Wales are mixed, and for East Midlands, London, South East, South West, West Midlands, and North West, we observe RMSFE ratios smaller than unity or hovering around unity for several combinations of the length of the rolling-estimation window and forecast horizon. On balance, the results suggest that taking into account regional output growth uncertainty over and above regional output growth and monetary and financial conditions helps to improve the accuracy of forecasts of regional output growth in several cases, where the results certainly display a certain degree of cross-regional heterogeneity.

The second bloc of results in Table 1 compares Model 2 and Model 3. This comparison sheds light on the contribution of uncertainty that originates in other regions for the accuracy of output growth forecasts. We observe in the majority of cases RMSFE ratios smaller than unity when we study the short 40-quarter rolling-estimation window. For the two longer rolling-estimation windows; in contrast, we observe several RMSFE ratios that exceed unity, especially for Yorkshire and the Humber, West Midlands, Wales, Scotland, and to a somewhat lesser extent for North West and Northern Ireland, with evidence that regional spillover effects help to improve the accuracy of output growth forecasts being weak for North East, East of England, and London. Hence, it appears that accounting for output growth uncertainty that has its origins in other regions implies that Model 3 for several regions and model configurations has a better forecast performance than Model 2 in terms of the RMSFE criterion.

The third bloc of results in Table 1 sheds light on the role of uncertainty as measured for London. The RMSFE ratios show that accounting for the "London effect" leads to more accurate forecasts for the following regions, especially when we consider the two longer rolling-estimation windows: East Midlands, South East, South West, West Midlands, North West, and Wales. The "London effect" is either small or even deteriorates the accuracy of forecasts when we consider North East, Yorkshire and the Humber, East of England, Scotland, and Northern Ireland.

We use the test proposed by Clark and West (2007) of equal mean-squared prediction errors to shed light on the statistical significance of differences in forecast performance across the various models. The null hypothesis is that the alternative model has the same out-of-sample forecasting performance as the baseline model. The alternative hypothesis is that the alternative model performs better than the baseline model. Table 2 summarizes the results (p-values).[7]

We find relatively strong evidence (in terms of the significance of the test results, using roughly a 5% threshold) that adding own regional uncertainty to Model 1 improves forecast accuracy for Yorkshire and the

[7]Comparing forecasts is complicated by the complex structure of random forests. We, therefore, use alternative techniques (RMSFE ratios in addition to formal tests) to compare forecasts. Moreover, as an additional analysis, we decomposed the regional output volatility into common and idiosyncratic components using the nonparametric and model-free two-step general dynamic factor approach of Barigozzi and Hallin (2016) to check whether such a decomposition adds value to the forecasting exercise. Based on Table A1, we find that such a decomposition gives largely insignificant results. This result is possibly an indication of the strong evidence of overall output growth volatility spillovers and interconnectedness across the regions, whereby distinguishing between common and idiosyncratic volatilities corresponding to common and local factors that drive these respective volatilities results in loss of information that tends to add value to the forecasting analysis.

TABLE 1 Comparing models by means of root-mean-square forecast-error ratios

Model combination	Region	Window length = 40			Window length = 60			Window length = 80		
		$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$
Base vs. Base + own SV	North East	0.9991	1.0028	1.0099	0.9985	1.0098	0.9849	1.0151	0.9900	1.0140
Base vs. Base + own SV	Yorkshire and the Humber	1.0162	1.0070	1.0188	1.0195	1.0509	1.0258	1.0747	1.0839	1.0594
Base vs. Base + own SV	East Midlands	1.0062	1.0032	0.9819	1.0078	0.9666	1.0035	0.9934	0.9823	1.0054
Base vs. Base + own SV	East of England	1.0286	1.0283	1.0211	1.0570	1.0287	1.0227	1.0683	1.0213	0.9956
Base vs. Base + own SV	London	1.0077	0.9957	1.0045	1.0348	0.9948	0.9936	0.9985	1.0072	0.9844
Base vs. Base + own SV	South East	0.9768	0.9950	0.9929	0.9698	0.9863	0.9862	0.9923	0.9806	0.9695
Base vs. Base + own SV	South West	0.9796	0.9980	0.9965	1.0007	0.9935	0.9914	1.0070	1.0032	0.9817
Base vs. Base + own SV	West Midlands	0.9956	1.0059	0.9976	0.9891	0.9749	0.9962	1.0387	0.9958	0.9894
Base vs. Base + own SV	North West	0.9767	1.0008	0.9898	0.9904	0.9887	0.9826	1.0183	1.0058	1.0131
Base vs. Base + own SV	Wales	0.9959	0.9832	1.0190	0.9874	1.0006	0.9962	1.0105	1.0263	1.0004
Base vs. Base + own SV	Scotland	1.0206	1.0430	1.0474	1.0128	1.0140	1.0230	1.0200	1.0285	1.0223
Base vs. Base + own SV	Northern Ireland	1.0077	1.0067	1.0243	1.0322	1.0076	0.9967	1.0326	1.0254	1.0053
Base + own SV vs. Base + all SVs	North East	0.9530	0.9329	0.9471	0.9847	0.9627	0.9856	1.0247	0.9779	0.9660
Base + own SV vs. Base + all SVs	Yorkshire and the Humber	0.9514	0.9391	0.9370	1.0375	1.0110	1.0009	1.0641	1.0049	1.0068
Base + own SV vs. Base + all SVs	East Midlands	0.9812	0.9678	0.9779	0.9904	0.9495	0.9703	1.0671	1.0233	1.0018
Base + own SV vs. Base + all SVs	East of England	0.9596	0.9125	0.9446	1.0070	0.9588	0.9795	0.9979	0.9626	0.9899
Base + own SV vs. Base + all SVs	London	0.9613	0.9316	0.9472	0.9998	0.9728	0.9574	1.0555	0.9750	0.9758
Base + own SV vs. Base + all SVs	South East	0.9360	0.9220	0.9256	1.0276	0.9849	0.9920	1.0203	0.9940	1.0141
Base + own SV vs. Base + all SVs	South West	0.9801	0.9455	0.9648	1.0303	0.9939	1.0046	1.0317	0.9746	0.9900
Base + own SV vs. Base + all SVs	West Midlands	0.9374	0.9075	0.9423	1.0933	1.0313	1.0113	1.0445	0.9965	1.0233
Base + own SV vs. Base + all SVs	North West	0.9801	0.9642	0.9536	1.0687	1.0072	1.0181	1.0671	0.9945	0.9756
Base + own SV vs. Base + all SVs	Wales	0.9718	0.9782	0.9522	1.1280	1.0497	1.0588	1.1719	1.0523	1.0605
Base + own SV vs. Base + all SVs	Scotland	0.9964	0.9800	0.9999	1.0400	1.0327	1.0089	1.0272	1.0163	1.0105
Base + own SV vs. Base + all SVs	Northern Ireland	1.0095	1.0064	1.0095	1.0422	0.9890	0.9868	1.0720	1.0097	1.0080
Base + own SV vs. Base + own SV + London SV	North East	0.9802	0.9563	0.9711	1.0242	1.0033	0.9972	1.0157	0.9940	1.0027
Base + own SV vs. Base + own SV + London SV	Yorkshire and the Humber	0.9938	0.9664	0.9489	1.0288	1.0299	0.9988	1.0009	0.9986	0.9925
Base + own SV vs. Base + own SV + London SV	East Midlands	1.0193	1.0179	1.0420	1.0328	1.0511	1.0309	1.0557	1.0153	1.0083

(Continues)

TABLE 1 (Continued)

Model combination	Region	Window length = 40			Window length = 60			Window length = 80		
		$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$
Base + own SV vs. Base + own SV + London SV	East of England	0.9956	1.0046	0.9976	0.9563	0.9736	0.9907	0.9682	0.9826	1.0044
Base + own SV vs. Base + own SV + London SV	London	—	—	—	—	—	—	—	—	—
Base + own SV vs. Base + own SV + London SV	South East	1.0200	1.0076	1.0137	1.0563	1.0401	1.0212	1.0294	1.0347	1.0286
Base + own SV vs. Base + own SV + London SV	South West	1.0399	1.0237	1.0446	1.0553	1.0323	1.0308	1.0490	1.0254	1.0429
Base + own SV vs. Base + own SV + London SV	West Midlands	1.0019	1.0001	1.0390	1.0205	1.0362	1.0311	1.0051	1.0304	1.0172
Base + own SV vs. Base + own SV + London SV	North West	0.9972	0.9825	0.9846	1.0385	1.0169	1.0335	1.0396	1.0030	0.9927
Base + own SV vs. Base + own SV + London SV	Wales	1.0126	1.0014	0.9906	1.0689	1.0276	1.0211	1.0594	0.9880	1.0039
Base + own SV vs. Base + own SV + London SV	Scotland	0.9792	0.9486	0.9642	1.0120	1.0000	0.9949	0.9890	1.0002	0.9945
Base + own SV vs. Base + own SV + London SV	Northern Ireland	0.9971	1.0097	0.9783	1.0071	0.9798	0.9996	1.0171	0.9892	0.9860

Note: This table reports RMSFE ratios, computed for out-of-sample forecasts. The column entitled "Model combination" gives the baseline and the alternative model. A ratio larger than unity indicates that the alternative model outperforms the corresponding baseline model. Estimation is by a rolling window. The parameter h denotes the forecast horizon (in months). The random forests are built using 2000 trees.

TABLE 2 Comparing models by means of the Clark–West test

Model combination	Region	Window length = 40			Window length = 60			Window length = 80		
		$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$
Base vs. Base + own SV	North East	0.1449	0.1968	0.0934	0.3079	0.0002	0.5351	0.0184	0.4173	0.1025
Base vs. Base + own SV	Yorkshire and the Humber	0.1167	0.1667	0.1439	0.0072	0.0875	0.0347	0.0126	0.0116	0.0658
Base vs. Base + own SV	East Midlands	0.0888	0.1867	0.8780	0.2424	0.7637	0.3380	0.4868	0.7249	0.2125
Base vs. Base + own SV	East of England	0.0236	0.0806	0.0232	0.0501	0.0068	0.0369	0.0188	0.0122	0.4277
Base vs. Base + own SV	London	0.1352	0.2039	0.1180	0.0522	0.3222	0.4287	0.3726	0.1410	0.6536
Base vs. Base + own SV	South East	0.8660	0.3319	0.3498	0.9002	0.7089	0.6230	0.6152	0.8144	0.8677
Base vs. Base + own SV	South West	0.8836	0.3027	0.3720	0.1916	0.3903	0.2501	0.1068	0.1344	0.6089
Base vs. Base + own SV	West Midlands	0.3562	0.2243	0.2762	0.5375	0.6829	0.3409	0.0363	0.3117	0.2963
Base vs. Base + own SV	North West	0.9370	0.2238	0.4862	0.3702	0.5129	0.7474	0.0236	0.0660	0.0145
Base vs. Base + own SV	Wales	0.3756	0.8652	0.0068	0.6842	0.2012	0.3997	0.1373	0.0249	0.2988
Base vs. Base + own SV	Scotland	0.0262	0.0354	0.0401	0.0610	0.1072	0.0382	0.0603	0.0360	0.0646
Base vs. Base + own SV	Northern Ireland	0.1075	0.1204	0.0094	0.0139	0.1169	0.2788	0.0463	0.1370	0.1579
Base + own SV vs. Base + all SVs	North East	0.6158	0.7253	0.5957	0.2274	0.3542	0.1214	0.1762	0.2465	0.4851
Base + own SV vs. Base + all SVs	Yorkshire and the Humber	0.6158	0.7030	0.6765	0.0544	0.0830	0.1812	0.1000	0.1804	0.1190

TABLE 2 (Continued)

Model combination	Region	Window length = 40			Window length = 60			Window length = 80		
		$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$	$h=1$	$h=2$	$h=4$
Base + own SV vs. Base + all SVs	East Midlands	0.4621	0.6262	0.4597	0.3120	0.5921	0.5761	0.1109	0.0065	0.1819
Base + own SV vs. Base + all SVs	East of England	0.6829	0.9244	0.7158	0.1386	0.8201	0.6254	0.3141	0.7973	0.4323
Base + own SV vs. Base + all SVs	London	0.8059	0.9038	0.8476	0.3023	0.6002	0.9218	0.1564	0.8152	0.8368
Base + own SV vs. Base + all SVs	South East	0.9491	0.9465	0.9267	0.1452	0.5365	0.4301	0.0474	0.1085	0.1222
Base + own SV vs. Base + all SVs	South West	0.4345	0.7151	0.5767	0.0037	0.2118	0.0644	0.1560	0.6090	0.2292
Base + own SV vs. Base + all SVs	West Midlands	0.8332	0.8708	0.6431	0.0516	0.0041	0.1335	0.1481	0.2203	0.0059
Base + own SV vs. Base + all SVs	North West	0.4127	0.5203	0.5926	0.0426	0.0475	0.0106	0.1224	0.1226	0.4704
Base + own SV vs. Base + all SVs	Wales	0.1989	0.1749	0.3591	0.0312	0.0086	0.0150	0.0618	0.0031	0.0144
Base + own SV vs. Base + all SVs	Scotland	0.2496	0.5762	0.1380	0.0231	0.0128	0.0887	0.0134	0.0373	0.0778
Base + own SV vs. Base + all SVs	Northern Ireland	0.0446	0.0811	0.0702	0.0323	0.1245	0.1223	0.0655	0.0989	0.0601
Base + own SV vs. Base + own SV + London SV	North East	0.4086	0.5615	0.5312	0.0571	0.0926	0.1102	0.1179	0.2030	0.1227
Base + own SV vs. Base + own SV + London SV	Yorkshire and the Humber	0.3238	0.6059	0.7084	0.0350	0.0350	0.2158	0.1422	0.1272	0.1795
Base + own SV vs. Base + own SV + London SV	East Midlands	0.0303	0.0511	0.0152	0.0601	0.0526	0.0454	0.0518	0.0734	0.1613
Base + own SV vs. Base + own SV + London SV	East of England	0.4066	0.0513	0.2776	0.9097	0.7805	0.7449	0.9325	0.9262	0.2283
Base + own SV vs. Base + own SV + London SV	London	—	—	—	—	—	—	—	—	—
Base + own SV vs. Base + own SV + London SV	South East	0.0111	0.1439	0.0914	0.0113	0.0045	0.0544	0.0092	0.0133	0.0219
Base + own SV vs. Base + own SV + London SV	South West	0.0199	0.0822	0.0169	0.0080	0.0475	0.0144	0.0041	0.0480	0.0138
Base + own SV vs. Base + own SV + London SV	West Midlands	0.2695	0.2871	0.0683	0.0489	0.0445	0.1227	0.1254	0.0963	0.1347
Base + own SV vs. Base + own SV + London SV	North West	0.2328	0.3481	0.3225	0.0120	0.0169	0.0021	0.1026	0.1523	0.3882
Base + own SV vs. Base + own SV + London SV	Wales	0.1143	0.2426	0.4898	0.0426	0.0330	0.1330	0.1457	0.4752	0.2282
Base + own SV vs. Base + own SV + London SV	Scotland	0.6258	0.9812	0.8416	0.0610	0.2234	0.3270	0.4343	0.1869	0.3191
Base + own SV vs. Base + own SV + London SV	Northern Ireland	0.1151	0.0241	0.3786	0.1497	0.5799	0.1117	0.1057	0.3982	0.5123

Note: This table reports the results (p-values) of the Clark–West test of equal mean-squared prediction errors. The null hypothesis is that the alternative model has the same out-of-sample forecasting performance as the baseline model. The alternative hypothesis is that the alternative model performs better than the baseline model. Results are based on Newey–West robust standard errors. Estimation is by a rolling window. The parameter h denotes the forecast horizon (in months). The random forests are built using 2000 trees.

Humber, East of England, and Scotland. The test results are occasionally significant for North East, North West, Wales, and Northern Ireland. Hence, there is some evidence that output growth uncertainty matters for forecasting regional output growth, though our results clearly show that it is important to differentiate between regions in this regard. As far as a comparison of Models 2 and 3 is concerned, we find that the model that includes the other regions stochastic volatilities as predictors often produces significantly better forecasting results for Wales, Scotland, and Northern Ireland than the model that dismisses uncertainty originating in other regions, for the former mainly for the two longer rolling-estimation windows. There is also some, albeit weaker, evidence that regional uncertainty spillover effects matter in some model configurations for West Midlands and North West. Finally, we find strong evidence that accounting for the "London effect" significantly improves forecast accuracy in the case of East Midlands, South East, and South West. We also find a few significant test results for West Midlands and North West, and Yorkshire and the Humber and Wales.

As a further extension, and as a robustness check, we replicated the analysis given in Table 2 for forecasts of the regional output growth rate h periods ahead, given data when a forecast has to be made, rather than its arithmetic average over the forecast horizon. The results (not reported to safe journal space, but available from the authors upon request) in some cases strengthen the evidence of a role of uncertainty. Specifically, evidence of predictive value of own regional output uncertainty strengthens for North East and London, and North West, while including the regional output uncertainty of all regions gives significant results for all regions at the two longer forecast horizons, that is, for $h = 2$ and 4 for the intermediate window length and in the overwhelming majority of regions for the long window. Finally, evidence of the "London effect" strengthens for North East, West Midlands, North West, and Wales. In sum, these results further back our conclusion that uncertainty matters for forecasting regional output growth and that it is important to carefully take into account regional heterogeneity in this regard.

4 | CONCLUDING REMARKS

We have used random forests and an SV model to study the out-of-sample predictive value of regional output growth uncertainty for regional output growth in 12 regions of the UK over the sample period from 1970 to 2020, where we have accounted for a region's own uncertainty, the uncertainty of other regions, and uncertainty as measured for London, given evidence of regional volatility connectedness. We have reported evidence that uncertainty helps to improve forecast accuracy and that spillover effects of uncertainty onto other regions as well as the "London effect" are beneficial in this regard too. The results, however, turned out to display a substantial extent of cross-regional heterogeneity, one interpretation of which is that the relative importance of the different channels, described in Section 1, through which uncertainty may affect output differs across regions.

From the perspective of policymaking, our results highlight primarily two issues: First, due to the evidence of volatility spillovers of output growth across regions, policymakers need to take into account the growth uncertainty of other regions beyond its own when making predictions about the future path of growth of a specific region and accordingly deciding on policy choices to mitigate the adverse effect of uncertainty. Second, given the underlying heterogeneity, understandably, policy decisions, both in terms of the type of intervention and its associated strength, cannot be uniform at the aggregate UK level, but need to be conducted in a region-specific manner. As far as academics are concerned, we confirm that the general in-sample results in the international literature suggesting a predictive impact of uncertainty on economic growth, as well as findings of out-of-sample gains due to the information contained in uncertainty for output growth of the aggregate UK, tend to hold at the regional level for the UK too, based on robust forecasting models that account for many predictors, nonlinearity, and interactions. This is an important finding, recalling that out-of-sample forecasting tests are more robust test of predictability compared with in-sample versions of the same. Finally, just like policy authorities, investors, when making investment decisions (like setting up a production plant) in a particular region, must be cognizant of the fact that uncertainty across the regions of the UK is connected and spillover, and, hence, a decision to invest in a particular region should not only be based on the corresponding uncertainty levels of that region, but also uncertainty emanating from other regions, and in particular that of London, should be taken into account. This information should assist in, for example, labor-hiring decisions too, as uncertainty also has a direct effect on unemployment (Gupta et al., forthcoming).

In future research, it is interesting to apply the methodology we use in our empirical research to study the output growth uncertainty nexus at the regional level for other countries (such as the USA). Another promising avenue for future research is to use alternative machine learning techniques to study the output growth uncertainty nexus. Such a comparison can also be used to trace out which machine learning technique performs best when applied to regional output growth data. In the process, this will allow researchers to improve upon a possible limitation of our current work, which is associated with the usage of only one, though appropriate for our context (in modeling multiple predictors, and nonlinearity as well as interaction among them), specific machine learning approach, namely, random forests.

ACKNOWLEDGMENTS

We would like to thank two anonymous referees and the editor, Professor Derek W. Bunn, for many helpful comments. Furthermore, David Gabauer would like to acknowledge that this research has been partly funded by BMK, BMDW and the Province of Upper Austria in the frame of the COMET Programme managed by FFG. However, any remaining errors are solely ours. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

David Gabauer https://orcid.org/0000-0003-2465-5741
Rangan Gupta https://orcid.org/0000-0001-5002-3428

REFERENCES

Abel, A. B. (1983). Optimal investment under uncertainty. *The American Economic Review*, *73*(1), 228–233.

Al-Thaqeb, S. A., & Algharabali, B. G. (2019). Economic policy uncertainty: A literature review. *The Journal of Economic Asymmetries*, *20*, e00133.

Antonakakis, N., Chatziantoniou, I., & Gabauer, D. (2020). Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions. *Journal of Risk and Financial Management*, *13*(4), 84.

Antonakakis, N., Gabauer, D., & Gupta, R. (2019). Greek economic policy uncertainty: Does it matter for the European Union? *Physica A: Statistical Mechanics and its Applications*, *535*, 122280.

Antonakakis, N., Gabauer, D., Gupta, R., & Plakandaras, V. (2018). Dynamic connectedness of uncertainty across developed economies: A time-varying approach. *Economics Letters*, *166*, 63–75.

Aye, G. C., Christou, C., Gil-Alana, L. A., & Gupta, R. (2019a). Forecasting the probability of recessions in South Africa: The role of decomposed term spread and economic policy uncertainty. *Journal of International Development*, *31*, 101–116.

Aye, G. C., Gupta, R., Lau, C. K. M., & Sheng, X. (2019b). Is there a role for uncertainty in forecasting output growth in OECD countries? Evidence from a time-varying parameter-panel vector autoregressive model. *Applied Economics*, *51*, 3624–3631.

Balcilar, M., Gabauer, D., & Zaghum, U. (2020). Crude oil futures contracts and commodity markets: New evidence from a TVP-VAR extended joint connectedness approach. ResearchGate Preprint: https://doi.org/10.13140/RG.2.2.25876.32643

Balcilar, M., Gupta, R., & Segnon, M. (2016). The role of economic policy uncertainty in predicting US recessions: A mixed-frequency Markov-switching vector autoregressive approach. *Economics: The Open-Access Open-Assessment E-Journal*, *10*, 1–20.

Balcilar, M., & Ozdemir, Z. A. (2020). A re-examination of growth and growth uncertainty relationship in a stochastic volatility in the mean model with time-varying parameters. *Empirica*, *47*(3), 611–641.

Barigozzi, M., & Hallin, M. (2016). Generalized dynamic factor models and volatilities: Recovering the market volatility shocks. *Econometrics Journal*, *19*(1), 33–60.

Barrios, S., Brülhart, M., Elliott, R. J. R., & Sensier, M. (2003). A tale of two cycles: Co-fluctuations between UK regions and the Euro Zone. *The Manchester Scool*, *71*(3), 265–292.

Beenstock, M., & Felsenstein, D. (2008). Regional heterogeneity, conditional convergence and regional inequality. *Regional Studies*, *42*(4), 475–488.

Bernanke, B. (1983). Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics*, *98*, 85–106.

Bloom, N. A. (2009). The impact of uncertainty shocks. *Econometrica*, *77*(3), 623–685.

Bouri, E., Gkillas, K., Gupta, R., & Pierdzioch, C. (2021). Forecasting realized volatility of Bitcoin: The role of the trade war. *Computational Economics*, *57*, 29–53.

Bredin, D., Fountas, S., & Savva, C. (2021). Is British output growth related to its uncertainty? Evidence using eight centuries of data. *Scottish Journal of Political Economy*, *68*(3), 345–364. https://doi.org/10.1111/sjpe.12270

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Caggiano, G., Castelnuovo, E., & Kima, R. (2020). The global effects of Covid-19-induced uncertainty. *Economics Letters*, *194*, 109392.

Caggiano, G., Castelnuvo, E., & Pellegrino, G. (2021). Uncertainty shocks and the great recession: Nonlinearities matter. *Economics Letters*, *198*, 109669.

Campbell, J. Y. (2008). Viewpoint: Estimating the equity premium. *Canadian Journal of Economics*, *41*, 1–21.

Castelnuovo, E., Lim, G. C., & Pellegrino, G. (2017). A short review of the recent literature on uncertainty. *Australian Economic Review*, *50*(1), 68–78.

Christou, C., Gabauer, D., & Gupta, R. (2020b). Time-varying impact of uncertainty shocks on macroeconomic variables of the United Kingdom: Evidence from over 150 years of monthly data. *Finance Research Letters*, *37*, 101363.

Christou, C., Gözgör, G., Gupta, R., & Lau, C. K. M. (2020a). Are uncertainties across the world convergent? *Economics Bulletin*, *40*(1), 855–862.

Clark, T. D., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*, 291–311.

Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*, *119*(534), 158–171.

Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, *28*(1), 57–66.

Diebold, F. X., & Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, *182*(1), 119–134.

Dixit, A. K., & Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton, NJ: Princeton University Press.

Duprey, T., Klaus, B., & Peltonen, T. (2017). Dating systemic financial stress episodes in the EU countries. *Journal of Financial Stability*, *32*, 30–56.

Gabauer, D., & Gupta, R. (2018). On the transmission mechanism of country-specific and international economic uncertainty spillovers: Evidence from a TVP-VAR connectedness decomposition approach. *Economics Letters*, *171*, 63–71.

Gupta, R., Lau, C. K. M., & Wohar, M. E. (2019). The impact of U.S. uncertainty on the Euro area in good and bad times: Evidence from a quantile structural vector autoregressive model. *Empirica*, *46*, 353–68.

Gupta, R., Ma, J., Risse, M., & Wohar, M. E. (2018). Common business cycles and volatilities in US states and MSAs: The role of economic uncertainty. *Journal of Macroeconomics*, *5*, 317–37.

Gupta, R., Olasehinde-Williams, G. A., & Wohar, M. E. (2020). The impact of US uncertainty shocks on a panel of advanced and emerging market economies: The role of exchange rate, trade and financial channels. *The Journal of International Trade & Economic Development*, *29*(6), 711–721.

Gupta, R., Pierdzioch, C., & Risse, M. (2016). On international uncertainty links: BART-based empirical evidence for Canada. *Economics Letters*, *143*, 24–27.

Gupta, R., Pierdzioch, C., & Salisu, A. A. (forthcoming). Oil-price uncertainty and the U.K. unemployment rate: A forecasting experiment with random forests using 150 years of data. *Resources Policy*.

Gupta, R., Sheng, X., Balcilar, M., & Ji, Q. (2021). Time-varying impact of pandemics on global output growth. *Finance Research Letters*, *41*, 101823. https://doi.org/10.1016/j.frl.2020.101823

Hartman, R. (1972). The effects of price and cost uncertainty on investment. *Journal of Economic Theory*, *5*(2), 258–266.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference and prediction*. (2nd ed.). New York, NY: Springer.

Junttila, J., & Vataja, J. (2018). Economic policy uncertainty effects for forecasting future real economic activity. *Economic Systems*, *42*, 569–583.

Karnizova, L., & Li, J. C. (2014). Economic policy uncertainty, financial markets and probability of US recessions. *Economics Letters*, *125*, 261–265.

Kastner, G., & Frühwirth-Schnatter, S. (2014). Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis*, *76*, 408–423.

Koop, G., McIntyre, S., & Mitchell, J. (2020c). UK regional nowcasting using a mixed frequency vector autoregressive model. *Journal of the Royal Statistical Society, Series A*, *183*(1), 91–119.

Koop, G., McIntyre, S., Mitchell, J., & Poon, A. (2020a). Regional output growth in the United Kingdom: More timely and higher frequency estimates, 1970-2017. *Journal of Applied Econometrics*, *35*(2), 176–197.

Koop, G., McIntyre, S., Mitchell, J., & Poon, A. (2020b). Reconciled estimates and nowcasts of regional output in the UK. *National Institute Economic Review*, *253*, R44–R59.

Kraft, H., Schwartz, E. S., & Weiss, F. (2018). Growth options and firm valuation. *European Financial Management*, *24*(2), 209–238.

Lastrapes, W. D., & Wiesen, T. F. P. (2021). The joint spillover index. *Economic Modelling*, *94*, 681–691.

Mumtaz, H. (2018). Does uncertainty affect real activity? Evidence from state-level data. *Economics Letters*, *167*, 127–130.

Mumtaz, H., Sunder-Plassmann, L., & Theophilopoulou, A. (2018). The state-level impact of uncertainty shocks. *Journal of Money, Credit and Banking*, *50*(8), 1879–1899.

Oi, W. Y. (1961). The desirability of price instability under perfect competition. *Econometrica*, *29*(1), 58–64.

Pierdzioch, C., & Gupta, R. (2020). Uncertainty and forecasts of U.S. recessions. *Studies in Nonlinear Dynamics & Econometrics*, *24*(4), 1–20.

Pindyck, R. S. (1991). Irreversibility, Uncertainty, and Investment. *Journal of Economic Literature*, *24*, 1110–1148.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Salisu, A. A., Gupta, R., Karmakar, S., & Das, S. (2022). Forecasting output growth of advanced economies over eight centuries: The role of gold market volatility as a proxy of global uncertainty. *Resources Policy*, *75*, 102527.

Segal, G., Shaliastovich, I., & Yaron, A. (2015). Good and bad uncertainty: Macroeconomic and financial market implications. *Journal of Financial Economics*, *117*(2), 369–97.

Segnon, M. K., Gupta, R., Bekiros, S., & Wohar, M. E. (2018). Forecasting US GNP growth: The role of uncertainty. *Journal of Forecasting*, *37*(5), 541–559.

Tibshirani, J., Athey, S., & Wager, S. (2020). grf: Generalized random forests. R package version 1.1.0. https://CRAN.R-project.org/package=grf

Wu, J. C., & Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit, and Banking*, *48*(2-3), 253–291.

AUTHOR BIOGRAPHIES

Mehmet Balcilar is a Professor and Chair of the Department of Economics at the Eastern Mediterranean University in Famagusta, Turkey. He earned his PhD degree in economics from Wayne State University, his MA degree in economics from Cleveland State University, and his bachelor's degree in economics from Dokuz Eylül University. Macroeconomics, monetary economics, forecasting, financial markets, nonlinear time series analysis, and risk modeling are his primary academic interests.

David Gabauer received his BSc and MSc degrees in economics from Johannes Kepler University (JKU), Austria, and PhD degree in social and economic sciences from the Vienna University of Economics and Business, Austria. He is currently a senior data scientist at the Software Competence Center Hagenberg (SCCH), Austria, where he is involved in several applied and fundamental industrial research projects related to quality control and predictive maintenance.

He has published more than 30 articles in widely renowned academic journals. His research interests include machine learning, deep learning, computer vision, pattern recognition, international finance, and time series econometrics.

Rangan Gupta is currently a Professor at the Department of Economics, University of Pretoria, South Africa. He completed his PhD in May 2005 from the Department of Economics, University of Connecticut, USA. He then joined the Department of Economics, University of Pretoria, as a senior lecturer in August 2005. His academic interests are mainly monetary theory and policy and time series econometrics. He has published widely in international journals in these areas of research.

Christian Pierdzioch is a Professor of Economics at the Department of Economics, Helmut Schmidt University, Hamburg, Germany. His research interests include macroeconomics, monetary economics, forecasting, financial econometrics, and sports economics.

How to cite this article: Balcilar, M., Gabauer, D., Gupta, R., & Pierdzioch, C. (2022). Uncertainty and forecastability of regional output growth in the UK: Evidence from machine learning. *Journal of Forecasting*, *41*(6), 1049–1064. https://doi.org/10.1002/for.2851

APPENDIX A

TABLE A1 Comparing models by means of the Clark–West test (common and idiosyncratic SV)

		Window length = 40			Window length = 60			Window length = 80		
Model combination	**Region**	$h = 1$	$h = 2$	$h = 4$	$h = 1$	$h = 2$	$h = 4$	$h = 1$	$h = 2$	$h = 4$
Base + own common SV vs. Base + all common SVs	North East	0.9260	0.9482	0.9452	0.0962	0.4193	0.5402	0.2633	0.5425	0.0767
Base + own common SV vs. Base + all common SVs	Yorkshire and The Humber	0.8923	0.0924	0.4338	0.8078	0.3016	0.6088	0.7728	0.4453	0.6389
Base + own common SV vs. Base + all common SVs	East Midlands	0.2932	0.9079	0.9233	0.5802	0.1168	0.7282	0.7778	0.7748	0.8121
Base + own common SV vs. Base + all common SVs	East of England	0.9075	0.7013	0.9319	0.7421	0.8695	0.2382	0.9032	0.5933	0.4280
Base + own common SV vs. Base + all common SVs	London	0.5024	0.2672	0.1434	0.0629	0.6327	0.8038	0.1545	0.8796	0.9477
Base + own common SV vs. Base + all common SVs	South East	0.8536	0.5019	0.5417	0.8736	0.8631	0.8900	0.1674	0.7104	0.4691
Base + own common SV vs. Base + all common SVs	South West	0.9214	0.3154	0.8325	0.9311	0.1154	0.8067	0.4269	0.0646	0.1084
Base + own common SV vs. Base + all common SVs	West Midlands	0.8177	0.9235	0.7920	0.0965	0.6426	0.2488	0.8897	0.5346	0.5057
Base + own common SV vs. Base + all common SVs	North West	0.3446	0.2558	0.3144	0.8078	0.8798	0.9448	0.3044	0.5803	0.7886
Base + own common SV vs. Base + all common SVs	Wales	0.5409	0.7728	0.2971	0.4314	0.4208	0.3647	0.7098	0.0649	0.3825
Base + own common SV vs. Base + all common SVs	Scotland	0.2266	0.2451	0.0356	0.8405	0.1505	0.3147	0.1266	0.8088	0.4292
Base + own common SV vs. Base + all common SVs	Northern Ireland	0.8664	0.6145	0.0945	0.6944	0.4794	0.3336	0.7931	0.8105	0.5142

(Continues)

TABLE A1 (Continued)

		Window length = 40			Window length = 60			Window length = 80		
Model combination	**Region**	$h = 1$	$h = 2$	$h = 4$	$h = 1$	$h = 2$	$h = 4$	$h = 1$	$h = 2$	$h = 4$
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	North East	0.7320	0.8375	0.3176	0.5591	0.3079	0.1087	0.0784	0.5314	0.8685
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	Yorkshire and the Humber	0.2516	0.4289	0.4058	0.7261	0.8606	0.8297	0.0788	0.5971	0.5312
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	East Midlands	0.7034	0.8749	0.7483	0.8928	0.4053	0.7786	0.5180	0.4500	0.4468
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	East of England	0.0441	0.0367	0.0567	0.8560	0.2677	0.3496	0.3037	0.0813	0.2857
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	London	0.8726	0.9574	0.9728	0.9381	0.9202	0.9245	0.9822	0.9505	0.8420
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	South East	0.9246	0.6381	0.9321	0.8385	0.7433	0.8884	0.9296	0.8834	0.3958
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	South West	0.8742	0.9440	0.2947	0.8570	0.8862	0.8578	0.3765	0.0545	0.5296
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	West Midlands	0.8773	0.8649	0.9402	0.8663	0.8692	0.8285	0.8374	0.8989	0.8014
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	North West	0.9539	0.9366	0.1254	0.1051	0.5899	0.5347	0.4192	0.3207	0.6165
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	Wales	0.5805	0.9397	0.5879	0.7806	0.8720	0.9154	0.3568	0.1093	0.1410
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	Scotland	0.7093	0.6803	0.0752	0.4434	0.8332	0.3895	0.0170	0.8506	0.2728
Base + own idiosyncratic SV vs. Base + all idiosyncratic SVs	Northern Ireland	0.0261	0.2371	0.2887	0.0666	0.1016	0.1555	0.0075	0.0807	0.0123

Note: This table reports the results (p-values) of the Clark–West test of equal mean-squared prediction errors. The null hypothesis is that the extended model has the same out-of-sample forecasting performance as the baseline model. The alternative hypothesis is that the full model performs better than the baseline model. Results are based on Newey–West robust standard errors. Estimation is by a rolling window. The parameter h denotes the forecast horizon (in months). The random forests are built using 2000 trees.