

RESEARCH ARTICLE

WILEY

Forecasting and nowcasting emerging market GDP growth rates: The role of latent global economic policy uncertainty and macroeconomic data surprise factors

Oguzhan Cepni¹ | I. Ethem Guney¹  | Norman R. Swanson²

¹Central Bank of the Republic of Turkey, Ankara, Turkey

²Department of Economics, Rutgers University, New Brunswick, New Jersey

Correspondence

I. Ethem Guney, Central Bank of the Republic of Turkey, Anafartalar Mah. Istiklal Cad No:10 06050 Ulus, Altındag, Ankara, Turkey.
Email: Ethem.Guney@tcmb.gov.tr

Abstract

In this paper, we assess the predictive content of latent economic policy uncertainty and data surprise factors for forecasting and nowcasting gross domestic product (GDP) using factor-type econometric models. Our analysis focuses on five emerging market economies: Brazil, Indonesia, Mexico, South Africa, and Turkey; and we carry out a forecasting horse race in which predictions from various different models are compared. These models may (or may not) contain latent uncertainty and surprise factors constructed using both local and global economic datasets. The set of models that we examine in our experiments includes both simple benchmark linear econometric models as well as dynamic factor models that are estimated using a variety of frequentist and Bayesian data shrinkage methods based on the least absolute shrinkage operator (LASSO). We find that the inclusion of our new uncertainty and surprise factors leads to superior predictions of GDP growth, particularly when these latent factors are constructed using Bayesian variants of the LASSO. Overall, our findings point to the importance of spillover effects from global uncertainty and data surprises, when predicting GDP growth in emerging market economies.

KEYWORDS

economic policy uncertainty, emerging markets, factor model, forecasting, lasso, shrinkage

1 | INTRODUCTION

In many countries, initial real gross domestic product (GDP) estimates are released at least 3 weeks after the calendar quarter to which the data pertain. For example, in the euro area and the USA, GDP reporting lags are 3 and 4 weeks, respectively; and in Turkey, first-release GDP data are available only after as many as 10–12 weeks. At the same time, tracking economic activity in real time is crucial to the decision-making process of macroeconomic policymakers. Fortunately, there is now an abundance of both real-time and big datasets available to researchers, allowing for the construction of ever more accurate early

forecasts and nowcasts (i.e., signals) of the current state of the economy. In this context, as pointed out by Giannone, Reichlin, and Small (2008), dynamic factor models (DFMs) have become one of the workhorses for short-term forecasting, and are now widely used in central banks and research institutions for both forecasting and nowcasting. For further discussion, see Artis, Banerjee, and Marcellino (2005) for the UK; Schumacher (2010) for Germany; Liu, Matheson, and Romeu (2012) for Latin America countries; Bessec (2013) for France; Luciani and Ricci (2014) for Norway; Girardi, Gayer, and Reuter (2015) for the Euro Area; Modugno, Soybilgen, and Yazgan (2016) for Turkey; Bragoli (2017) for Japan; Kim and Swanson (2018a) for

the USA; Kim and Swanson (2018b) for Korea; Luciani, Pundit, Ramayandi, and Veronese (2018) for Indonesia; Bragoli and Fosten (2018) for India; and Cepni, Güney, and Swanson (2019) for emerging markets.

In this paper, we contribute to the literature on nowcasting and forecasting real GDP growth in emerging economies by empirically assessing the importance of economic policy uncertainty and data surprises in factor-type econometric forecasting models for five emerging market economies: Brazil, Indonesia, Mexico, South Africa, and Turkey. Our analysis centers around the use of DFMs for constructing GDP predictions, although we also evaluate benchmark linear autoregressive (AR) models. Importantly, our DFMs are specified both with and without uncertainty and data surprise factors constructed using both local and global economic variables. Moreover, in addition to standard econometric estimation methods, we estimate factors using a variety of data shrinkage methods including the standard least absolute shrinkage operator (LASSO), the adaptive LASSO, the Bayesian LASSO, and the Bayesian adaptive LASSO.

As mentioned above, we utilize both local and global datasets when constructing the factors used in our prediction models. Although there are many empirical papers that focus on using only local macroeconomic data in the context of GDP forecasting, the importance of uncertainty regarding policymakers' decisions on international economic policies has received increasing attention since the beginning of 2018. For example, concerns over US–China trade tensions, Brexit negotiations with the EU, Italy's fiscal planning, and how the Federal Reserve Board of the USA will determine the timing and pace of policy normalization all weigh heavily on the global economy. These sorts of international spillovers are particularly important for emerging markets, in particular those with high foreign portfolio ownership and weak macroeconomic balance sheets. For this reason, it is crucial to assess the relevance of uncertainty and data surprises in the context of forecasting emerging market GDP. Needless to say, the impact of uncertainty on economic activity has received considerable attention in the economics literature in recent years (see, e.g., Bloom, 2014). For example, Bontempi, Golinelli, and Squadrani (2016) propose uncertainty indicators based on Google Trends. Their results suggest that online search data can provide early signals of uncertainty, and can be used in macroeconomic forecasting. Baker, Bloom, and Davis (2016) construct an index of economic policy uncertainty (EPU) based on newspaper coverage frequency. They find that economic policy innovations foreshadow declines in investment, output, and employment, using a panel vector autoregressive (VAR) model for 12 major economies. Thorsrud (2018) develops a new coincident index of business cycle activity based on

quarterly GDP and textual information contained in a daily business newspaper.

In addition to the above proxies for economic policy “uncertainty,” a growing strand of the literature uses consensus forecasts to disentangle macroeconomic uncertainty from more “general” uncertainty. In particular, it is argued in this literature that professional forecasters (e.g., those forecasters contributing to the Survey of Professional Forecasters in the USA) closely monitor macroeconomic data and often base their predictions on sophisticated econometric models. Thus departures of their (consensus) predictions from actual realizations can be viewed as data “surprises,” which are themselves measures of macroeconomic uncertainty. There are different proxies for this sort of macroeconomic uncertainty that are proposed in the empirical literature. For example, Rossi and Sekhposyan (2015) construct a macroeconomic uncertainty index based on comparing the realized forecast error of a variable of interest with the sample distribution of the forecast errors of that variable. If the realization is in the tail of the distribution, they conclude that the macroeconomic environment is more uncertain. Carriero, Clark, and Marcellino (2016) develop a model to identify uncertainty by modeling the common component of the volatility of the forecast errors of a large set of macroeconomic and financial variables. Finally, Scotti (2016) proposes a macroeconomic surprise index that exploits the difference between actual releases of data and Bloomberg forecasts to capture economic agents' expectations about the state of the economy. In this paper, we construct global economic policy uncertainty and surprise indices based on a variety of different local and global datasets. More specifically, we incorporate uncertainty into our prediction models in three different ways. First, as our benchmark we utilize only local macroeconomic data. Using this approach, we estimate both DFMs and simple AR models, but do not explicitly include any uncertainty or surprise indexes. Second, we augment our DFMs with “surprise” indices constructed using professional forecasters' expectations. Finally, we additionally augment our DFMs with factors extracted from a wide variety of uncertainty indices of economic policy, trade policy, monetary policy, and migration.

It should be noted that we focus on the prediction of GDP growth in emerging markets (EM) for two main reasons. First, official releases of EM GDP figures are subject to significant publication lags and data revision, as discussed above. Second, again as discussed above, it is likely that the effects of uncertainty on economic activity are particularly significant in environments characterized by large budget deficits, high current account/GDP ratios, and high external funding needs. As a case in point, note that Carriere-Swallow and Cespedes (2013) investigate the effects of an uncertainty shock from the USA on

developed and developing countries. They find that emerging markets suffer much more severe falls in investment and private consumption when there are credit constraints that lead to increases in the uncertainty index. Gauvin, McLoughlin, and Reinhardt (2014) also point out that elevated policy uncertainty in advanced countries may lead to an increase in capital outflows from emerging markets, because of rising global risk aversion. At the same time, it is important to note that low-quality emerging market data present distinct challenges when forecasting GDP growth. However, our dataset contains a very large number of variables, allowing us to mitigate this effect to some degree by “preselection” of key variables for use in our DFMs. Namely, we argue that a small set of indicators from our dataset may be sufficient to identify the most informative uncertainty indexes. Indeed, the idea that a small set of indicators, when chosen appropriately, can improve forecasting performance of factor models is supported by evidence presented in various papers, including Banerjee and Marcellino (2006); Boivin and Ng (2006); Schumacher (2007); Bai and Ng (2008), Banbura and Rünstler (2011), Kim and Swanson (2014, 2018a), Bulligan, Marcellino, and Venditti (2015), and references cited therein. For this reason, we utilize variable selection methods to preselect indicators before the construction of factors. To this end, we apply a number of LASSO methods for data shrinkage, as discussed above. Of course, we also use standard DFM methods where the entire dataset is used in factor construction. In this way, our analysis adds to our understanding not only of GDP forecasting using uncertainty indexes but also to the usefulness of shrinkage methods in DFM modeling.

Our empirical findings can be summarized as follows. First, as expected, there is a substantial reduction in mean square forecast error (MSFE), as more data related to the current quarter become available and are incorporated into our models. Thus the forecast accuracy of our DFM models generally increases by incorporating the latest information. This conclusion, as well as our subsequent findings, are based on point forecast evaluation. As pointed out by a referee, it is important to note that density forecasting may provide further insights into the usefulness of the models and methods analyzed in this paper.¹ Second, uncertainty indexes are quite useful for predicting GDP growth in emerging market economies. More specifically, benchmark AR models, as well as DFMs that utilize only local macroeconomic data yield inferior predictions for Indonesia, South Africa, and Turkey. For these three

countries, both “surprise” and “uncertainty” indexes matter. Third, constructing surprise and uncertainty indexes using LASSO shrinkage methods leads to more accurate forecasts than when factors are constructed without variable preselection. Indeed, across all 10 forecast horizons (including forecasts, nowcasts, and backcasts), and across all five countries, the “globally best” models include factors constructed using targeted variable preselection in 47 of the 50 cases.² Finally, augmenting the DFMs utilized in our regression analysis with lags of the dependent variable (i.e., including an AR component in the DFM) yields forecasting gains relative to models without AR terms only for Indonesia and South Africa. This result is interesting given the preponderance of evidence in the time series forecasting literature concerning the importance of including AR components when forecasting economic variables, and serves to underscore the importance of uncertainty and surprise variables when predicting EM GDP.

Summarizing, we find strong evidence that using targeted predictors in combination with global uncertainty indices as well as global surprise indices based on expectations of professional forecasters leads to more precise GDP predictions for emerging market economies. These findings suggest that when assessing macroeconomic conditions policymakers and central bankers may increasingly need to take into account the level of economic policy uncertainty originating from other countries, as well as related macroeconomic forecasts formulated by professional forecasters.

The rest of the paper is structured as follows. In Section 2 we present the main features of the datasets used in our empirical investigation. In Section 3 we outline the econometric methodology used in the papers. In Section 4 we discuss our empirical findings. Concluding remarks are gathered in Section 5. Finally, the Supporting Information Appendices include supplementary tables as well as additional details describing the datasets analyzed in the sequel.

2 | DATA

We analyze a relatively large set of economic indicators consists of 97, 87, 116, 109, and 102 economic variables for Brazil, Indonesia, Mexico, South Africa, and Turkey, respectively. The dataset is composed of both “hard indicators” and survey data. Among the hard indicators, we have both supply-side variables, such as industrial production

¹Along these lines, for example, Abbate and Marcellino (2018) find that interval and density forecasts of exchange rates are relevant to the modeling and decision-making process of economic agents interested in constructing predictions G10 currencies.

²A globally best model is one that yields the lowest MSFE across all five data selection methods, including: (i) use all data in DFM construction; (ii) use the standard LASSO for dataset reduction; (iii) use the adaptive LASSO for dataset reduction; (iv) use the Bayesian LASSO for dataset reduction; and (v) use the Bayesian adaptive LASSO for dataset reduction.

indexes, and demand-side variables, such as electricity consumption. Among the survey variables, we have the Market PMI survey, one of the most watched indicators of the business cycle. Given the sensitivity of EM economies to external conditions, we also include current account balance and volume indices of exports and imports, as well as real effective exchange rates. The dataset can be divided into six categories: *Housing and Order Variables*—house price index, real estate units sold and new orders; *Labor Market Variables*—employment and unemployment; *Prices*—producers' prices and consumer prices; *Financial Variables*—Treasury bond yields, credit default swaps, exchange rates, and stock prices; *Money, Credit and Quantity Aggregates*—money supply, commercial bank loans, time and sight deposits; *Real Economic Activity*—PMI survey, industrial production, retail sales, vehicle production, and capacity utilization. In general, the survey variables and nominal indicators are released during the reference month (i.e., the calendar month to which the data pertains), whereas real and labor variables are announced with a publication lag of 1–3 months.

In addition to the above collection of macroeconomic and financial indicators, which are used in our baseline DFM model (this DFM model is called Specification 1 in Section 3.3), we examine an uncertainty dataset that contains a wide variety of uncertainty indices pertaining to economic policy, trade policy, monetary policy, and migration. These variables are largely the same as those constructed by Baker et al. (2016) for major economies and are called economic policy uncertainty (EPU) indices. Their data construction approach is based on computing the proportion of newspaper articles referring to a specific type of uncertainty over a given period. In particular, the EPU indices reflect the frequency of articles that include terms related to three categories—that is, economy (E), policy (P) and uncertainty (U).³ In total, we utilize 45 “uncertainty” indices from various fully and less developed countries when constructing our “uncertainty” factors for use in our DFM forecasting models (see Specifications 2 and 4 in Section 3.3).

We now turn to a discussion of our “surprise” indices. Understanding how economic data evolve, relative to consensus expectations, is important for gauging potential shifts in macroeconomic sentiment, for a given country. For this purpose, we collect a large set of surprise indices across key regions and countries. Our dataset consists of Citi and Goldman Sachs surprise indices, which are constructed based on similar methodologies. These indices are designed to summarize the degree of surprise inher-

ent in a data release, relative to the associated Bloomberg median forecast of the variable in question. Multiplying the so-called relevance score for a variable by the surprise score allows us to track whether the economic data in a country are “outperforming” or “underperforming” consensus expectations. Since units of measurement vary across variables, the surprise scores are normalized by dividing by sample standard deviation of the corresponding variables.⁴ In total, we have 74 surprise indices across different regions and countries, including the USA, UK, euro area, Japan, Asia Pacific region, and Latin America. These indices are then used to construct our “surprise” factors for use in our DFM forecasting models (see Specifications 3 and 4 in Section 3.3).

All datasets were collected from Bloomberg, and are for the period January 2003 to June 2018. The complete list of Bloomberg tickers is provided in Tables A8–A14 in the Supporting Information Appendix.

Many of the macroeconomic variables that we analyze, and which are likely to contain seasonality, are seasonally adjusted by the reporting agencies. In the case of seasonally adjusted data not directly available for download, we construct year-on-year (YoY) growth rates or yearly differences, and do not implement any other variety of seasonality filter. One reason for this is that, as suggested by Luciani et al. (2018), the use of standard seasonality filters results in data deformation, which in turn may induce estimation bias. It is also worth pointing out that the use of YoY growth rates is consistent with the target variable that we forecast, which is the year-on-year GDP growth rate. Finally, as noted in Luciani et al. (2018), the use of YoY GDP growth rates makes our results comparable with YoY forecasts made by institutional forecasters, such as the IMF and various other data-reporting agencies.⁵ For further discussion of the importance of seasonality and issues associated with seasonal adjustment, including the fact that use of YoY growth rates may induce spurious correlation, see Ghysels, Osborn, and Sargent (2001), and Swanson and Urbach (2015), and Uhlig (2009).

⁴Note that relevance scores are defined as the absolute value of the contemporaneous correlation between each variable and real GDP growth.

⁵To incorporate quarterly YoY GDP data into this setting, we construct a partially observable monthly YoY GDP series and link it to the monthly variables, imposing restrictions on the loadings for the quarterly variables which constitute a slight modification of the Mariano and Murasawa (2003) methodology. For further details refer to the Supporting Information Appendix for this paper, as well as to Giannone, Agrippino, and Modugno (2013).

³More details on the EPU indices can be found at <http://www.policyuncertainty.com/index.html>.

3 | ECONOMETRIC METHODOLOGY

3.1 | Dynamic factor models

In our analysis, we separately extract potentially useful forecasting information from our three datasets (i.e., our macroeconomic indicator dataset, surprise dataset, and uncertainty index dataset). To do this, we employ the widely used DFM of Giannone et al. (2008). In this framework, the dynamics of individual variables is represented as the sum of a component that is common to all variables in the economy and an orthogonal idiosyncratic residual. Formally, the DFM can be written as a system with two types of equation: a measurement equation (Equation 1), which links the observed variables to the unobserved common factor to be estimated, and the transition equation (Equation 2), which describes the dynamics of the common factors and the residuals of the measurement equation. Once Equations 1 and 2 are written in state-space form, the Kalman filter and smoother are applied to extract common factors and generate forecasts for all of the variables in the model.

More specifically, we consider a panel of observable economic variables, X_{it} , where i indicates the cross-section unit, $i = 1, \dots, N$, and t denotes the time index, $t = 1, \dots, T$. Each variable in the dataset can be decomposed into common part and idiosyncratic components, where the common component captures comovement in the data, and is driven by a small number of shocks. The DFM model can be written as

$$X_t = \Lambda F_t + \xi_t, \quad \xi_t \sim N(0, \Sigma_\varepsilon), \quad (1)$$

$$F_t = \sum_{i=1}^k \Psi_i F_{t-i} + u_t, \quad u_t \sim N(0, Q), \quad (2)$$

where F_t is an $r \times 1$ vector of unobserved common factors with zero mean and unit variance, Λ is a corresponding $N \times r$ factor loading matrix, and the idiosyncratic disturbance, ξ_t , is uncorrelated with F_t at all leads and lags, and has a diagonal covariance matrix Σ_ε . It is assumed that the common factors, F_t , follow a stationary VAR(p) process driven by the common shocks, $u_t \sim N(0, Q)$, and that the Ψ_i are $r \times r$ matrices of autoregressive coefficients. Also, the common shocks (u_t) and idiosyncratic components (ξ_t) are orthogonal. To handle missing observations at the end of the sample due to the nonsynchronous flow of data, we characterize the variance of the idiosyncratic component as an extremely large variance. In this way, we ensure that the Kalman filter will put no weight on missing observations in the extraction of the common factors. Finally, in order to construct forecasts of our quarterly GDP target series, say y_t , in our monthly DFM framework, we express each quarterly variable in terms of a partially observed monthly counterpart following the approach of Mariano

and Murasawa (2003). Put differently, the general form of the forecasting model can be described as follows:

$$y_t = \mu + \beta' F_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \quad (3)$$

where ε is a stochastic disturbance term. In order to select the optimal number of factors (r), one may use various methods suggested in the literature. A nonexhaustive list of possible methods is discussed in the following papers: Bai and Ng (2002); Onatski (2009), Alessi, Barigozzi, and Capasso (2010), and Ahn and Horenstein (2013). Although the Bai and Ng (2002) criterion is frequently used in the empirical literature, we find that it generally chooses too many factors, resulting in deterioration in forecast accuracy. Hence we adopt the Onatski (2009) approach, which is based on testing the null of $r - 1$ common factors against the alternative of r common factors. Optimal lag length is selected using the Schwarz information criterion (SIC).⁶ In summary, we find that simple model specifications, with one or two factors and one or two lags, often yield best out-of-sample forecast performance. Specification of a parsimonious one- or two-factor model is also consistent with the literature on factor models, which has shown that heavily parametrized models with many factors usually lead to poor forecasting performance (see Bragoli, 2017; Forni, Hallin, Lippi, & Reichlin, 2000; Stock & Watson, 2002).

3.2 | Identifying targeted predictors: LASSO-based approaches

Since DFMs do not explicitly incorporate knowledge of the target variable being forecast, factor extraction is called “untargeted.” Needless to say, targeted forecasting, in which factors with information specific to the target variable are constructed and used in subsequent model specification, may yield superior predictions, relative to predictions constructed using factors based on untargeted methods. These issues are discussed at length in Boivin and Ng (2006); Bai and Ng (2008); Schumacher (2010), Caggiano, Kapetanios, and Labhard (2011), Medeiros and Vasconcelos (2016), Kapetanios, Marcellino, and Papailias (2016), Kim and Swanson (2014, 2018b), and many others. In the sequel, we utilize both untargeted and targeted methods to construct factors. For untargeted forecasting, we simply use the DFM model outlined in the previous section. For targeted forecasting, we note that one of the most commonly used variable selection and parameter shrinkage methods is the LASSO. This method can be thought of as a type of penalized regression, related to classical ridge regression. However, in the case of ridge regression, an ℓ_2 -norm penalty is imposed, while under the LASSO,

⁶The list of selected r and p pairs is (2, 2), (1, 2), (1, 3), (3, 3), and (1, 2) for Brazil, Indonesia, Mexico, South Africa, and Turkey, respectively.

an ℓ_1 -norm penalty function is added to the usual least squares minimization problem. Interestingly, the ℓ_1 -norm penalty function induces shrinkage to 0 of coefficients associated with “irrelevant” variables.⁷ For our purposes, since we are interested in selecting a targeted subset of the original variables in our dataset when constructing shrinkage type factors, we use the LASSO. In particular, we analyze factors constructed using the standard LASSO, the adaptive LASSO, the Bayesian LASSO, and the Bayesian adaptive LASSO. When carrying out variable selection, we apply LASSO shrinkage to the full dataset. As suggested by a referee, recursive selection prior to reestimation of the model at each point in time may lead to a variable selection instability. This in turn would complicate the interpretation of forecasts’ revisions, which is crucial when reporting forecasts to policymakers. However, it should be stressed that the issue of recursive estimation is important, and that future research in this area is warranted. In order to illustrate why this issue is important, we carried out an experiment in which we repeated our variable selection, but using recursive estimation, with a new set of variables selected at each point in time. The percentage of variables that are reselected more than one half of the time under recursive is given below, for each country.

Percentage of variables that are reselected more than half in case of iterative variable selection

Country	Ratio
Brazil	0.37
Indonesia	0.60
Mexico	0.37
South Africa	0.54
Turkey	0.50

Inspection of the entries in this table indicates the importance of recursive variable selection, as the percentage of variables that are reselected more than one half of the time is as low as 37% (for Brazil and Mexico).

Turning again to our empirical setup, we consider a panel of observable economic variables, $X_{i,t}$, where i indicates the cross-section unit, $i = 1, \dots, N$, and t denotes the time index, $i = 1, \dots, T$, as discussed above. Following the notation of Hastie, Tibshirani, and Friedman (2009), we consider the problem of selecting a subset of X , where X is a $T \times N$ matrix to be used for forecasting quarterly GDP growth, say Y , for $i = 1, \dots, T$.

⁷As an example of the coefficients referred to in the previous statement, consider coefficients that are defined to be the weights in a linear combination of variables forming a latent factor, when principal component analysis (PCA) is used to construct F_t . When using the LASSO, various of these factor loadings may be identically zero. On the other hand, the classical ridge regression penalty function results in all coefficients being nonzero, for each latent factor, under PCA.

3.2.1 | Least absolute shrinkage operator

We implement the LASSO, as developed in Tibshirani (1996). This shrinkage operator performs both variable selection and regularization, using a regularization parameter, λ . The idea is to impose an ℓ_1 -norm penalty on the regression coefficients, thus allowing for cases where $N > T$. This penalty also results in (possible) shrinkage of coefficients (called $\hat{\beta}^{\text{LASSO}}$) to zero, as discussed above. The LASSO estimator is defined as

$$\hat{\beta}^{\text{LASSO}} = \underset{\beta}{\operatorname{argmin}} \|Y - X\beta\|_2 + \lambda \sum_{j=1}^N |\beta_j|, \quad (4)$$

where λ is a tuning (regularization) parameter that controls the strength of the ℓ_1 -norm penalty. We choose the tuning parameter λ via cross-validation, which is a data-driven method that is designed to maximize expected out-of-sample forecasting performance. Since the objective function in the LASSO is not differentiable, numerical optimization must be used when constructing $\hat{\beta}^{\text{LASSO}}$. For example, an efficient iterative algorithm called the “shooting algorithm” is proposed in Fu (1998). One of the limitations of the LASSO approach is that the sample size bounds the number of selected variables. For example, if $N > T$, the LASSO yields at most T nonzero coefficients (see ; Swanson, 2016), for further discussion). The variables associated with these nonzero coefficients constitute our set of targeted predictors when using the LASSO method.

3.2.2 | Adaptive LASSO (AdaLASSO)

Although the LASSO can perform automatic variable selection because of the ℓ_1 -norm penalty, it may yield biased estimates when coefficients are large. Fan and Li (2001) conjecture that LASSO does not have oracle properties and may yield an inconsistent set of selected variables in some cases. In order to address these issues, Zou (2006) introduces a new version of the LASSO, called the adaptive LASSO, where adaptive weights are used for penalizing different coefficients in the ℓ_1 -norm penalty. They show that adaptive LASSO enjoys oracle properties. Moreover, Medeiros and Mendes (2016) show that the adaptive LASSO estimator maintains its consistency, under very general conditions, and performs well even when the number of variables increases faster than the number of observations.

The adaptive LASSO estimator can be written as follows:

$$\hat{\beta}^{\text{AdaLASSO}} = \underset{\beta}{\operatorname{argmin}} \|Y - X\beta\|_2 + \lambda \sum_{j=1}^N w_j |\beta_j|, \quad (5)$$

where $w_j = |\hat{\beta}_j^*|^{-\tau}$ represents different weights on the penalization of each variable, $|\hat{\beta}_j^*|$ is an initial estimator, such as the ordinary least squares (OLS) or ridge estimator

(which is estimated in a first step), and $\tau > 0$ controls the difference in weights. Since the adaptive LASSO is still an ℓ_1 -norm penalization method, the algorithms for solving the LASSO can be employed for constructing adaptive LASSO estimates.

3.2.3 | Bayesian LASSO (BLASSO)

Tibshirani (1996) suggested that LASSO estimates can be interpreted as the posterior mode of the Bayes estimate, under the Laplace priors. While the LASSO attributes a value of exactly zero to regression coefficients of irrelevant variables, redundant Bayesian LASSO coefficient modes are not precisely zero. Instead, the BLASSO provides a posterior distribution of coefficients. In our BLASSO estimator, we implement a conditional Laplace prior of the form

$$\pi(\beta|\sigma^2) = \prod_{j=1}^p \frac{\lambda}{2\sqrt{\sigma^2}} e^{\frac{-\lambda|\beta_j|}{\sqrt{\sigma^2}}}. \quad (6)$$

Park and Casella (2008) show that conditioning on error variance, σ^2 , ensures a unimodal full posterior. Otherwise, expensive simulation methods lead to slow convergence of the Gibbs sampler and result in less meaningful point estimates. Since the Laplace distribution can be represented as a scale mixture of normal densities with an exponential mixing density, Park and Casella (2008) propose the following hierarchical BLASSO model:

$$\begin{aligned} y|X, \beta, \sigma^2 &\sim N_n(X\beta, \sigma^2 I_n) \\ \beta|\sigma^2, \{\tau_1^2, \dots, \tau_p^2\} &\sim N_p(0_p, \sigma^2 D_\tau) \\ D_\tau &= \text{diag}\{\tau_1^2, \dots, \tau_p^2\}, \end{aligned} \quad (7)$$

with the following priors on σ^2 and $\tau = \{\tau_1^2, \dots, \tau_p^2\}$:

$$\begin{aligned} \sigma^2, \tau_1^2, \dots, \tau_p^2 &\sim \pi(\sigma^2) d\sigma^2 \\ &\times \prod_{j=1}^p \frac{\lambda^2}{2} e^{\frac{-\lambda^2 \tau_j^2}{2}} d\tau_j^2, \quad \sigma^2, \tau_1^2, \dots, \tau_p^2 > 0, \end{aligned} \quad (8)$$

where D_τ is the prior covariance matrix and λ is a rate parameter of the exponential distribution. Korobilis (2013) compares the forecasting performance of hierarchical Bayesian shrinkage and factor models, and finds that Bayesian shrinkage serves as a valuable addition to existing methods, in the presence of many predictor variables. Our approach is to use this BLASSO method and store the posterior distributions of all coefficients. Thereafter we calculate the 95% confidence intervals, and set a coefficient equal to zero if its interval includes zero. The remaining variables constitute our set of targeted predictors, when using the Bayesian LASSO.

3.2.4 | Bayesian adaptive LASSO (BaLASSO)

As discussed in Section 3.2.2, the AdaLASSO uses weighted shrinkage for consistent estimation of regression coefficients, while retaining the attractive convexity property of the LASSO. However, the AdaLASSO requires consistent and informative initial estimates of the regression coefficients, which are generally not available when the number of regressors is larger than the number of observations. Since BaLASSO does not require any informative initial estimates of the regression coefficients, this motivates us to replace Equation 8 in the Bayesian LASSO section to allow for a more adaptive penalty, as follows:

$$\sigma^2, \tau_1^2, \dots, \tau_p^2 \sim \pi(\sigma^2) d\sigma^2 \prod_{j=1}^p \frac{\lambda_j^2}{2} e^{\frac{-\lambda_j^2 \tau_j^2}{2}} d\tau_j^2. \quad (9)$$

Similar to the BLASSO, we select variables such that their corresponding 95% confidence intervals do not include zero.

3.3 | Factor-augmented prediction models

To evaluate the forecasting performance of DFMs based on different factor specification types, we run a recursive pseudo out-of-sample forecasting exercise over the period July 2008 to June 2018. For each reference quarter, we produce a sequence of 10 forecasts, starting with the forecast based on the information available in the first month of the two previous quarters, and stopping on the first of the month of the subsequent quarter, before GDP is released. Put differently, we construct three monthly forecasts ($h = 1, 2$), three monthly nowcasts ($h = 0$), and one monthly backcast ($h = -1$) for a quarterly forecast of GDP.

We apply the DFM to extract the leading common factors in our large set of uncertainty and surprise indices. We called these new factors our “global economic policy uncertainty” (GEPU) and “global macro-surprise” (GMS) factors. Earlier, we referred to these variables as our “uncertainty” and “surprise” variables, respectively. These factors can be interpreted as measures of common variation in economic policy uncertainty (uncertainty) and macroeconomic data uncertainty (surprise) across countries. These variables are individually “added” to our DFM model. In particular, noting that “Local” refers to factors that are constructed using only “own” country variables, we construct predictions using the following specifications, where Specification 1 is the model used in Giannone et al. (2008), and Specifications 2–4 are extensions that incorporate our new uncertainty and surprise factors.

- *Specification 1: Local factor model*

$$y_{t+h} = \mu + \beta' F_t^{\text{Local}} + \varepsilon_{t+h}$$

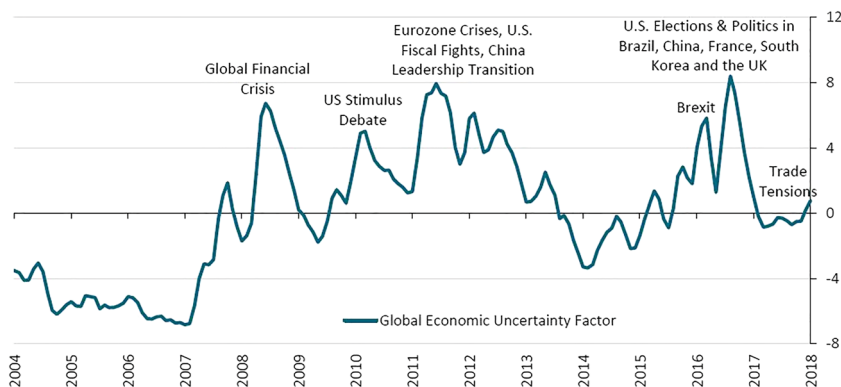


FIGURE 1 Global economic uncertainty factor [Colour figure can be viewed at wileyonlinelibrary.com]

- *Specification 2: Uncertainty factor model*

$$y_{t+h} = \mu + \beta' F_t^{\text{Local}} + \vartheta' F_t^{\text{GEPu}} + \varepsilon_{t+h}$$

- *Specification 3: Macro-surprise factor model*

$$y_{t+h} = \mu + \beta' F_t^{\text{Local}} + \theta' F_t^{\text{GMS}} + \varepsilon_{t+h}$$

- *Specification 4: Global factor model*

$$y_{t+h} = \mu + \beta' F_t^{\text{Local}} + \theta' F_t^{\text{GEPu}} + \delta' F_t^{\text{GMS}} + \varepsilon_{t+h}.$$

An additional set of models is also used to construct predictions, in which each of the above DFM models is augmented with lags of the dependent variable.⁸ Finally, we also construct forecasts using a straw-man AR model, with lags selected using the SIC.⁹ We assess the precision of the different sequences of forecasts using mean square forecast error (MSFE), which is measured as the average of the squared differences between predicted and actual GDP values for the 10-year period from July 2008 to June 2018. In order to assess the statistical significance of differences in MSFE across specification types, pairwise Diebold–Mariano (DM: 1995) tests of equal predictive accuracy are implemented. In this test, the null hypothesis is that of equal predictive accuracy. For a discussion of the test, which is normally distributed under nonnestedness, and has a nonstandard distribution if models are nested, see Kim and Swanson (2018a, 2018b). When carrying out DM tests, the benchmark against which each of

our Specification types 1–8 models are compared with our straw-man AR model.

4 | EMPIRICAL RESULTS

4.1 | Global economic policy uncertainty and macroeconomic surprise factors

Before describing our prediction results, it is of interest to investigate how global economic policy uncertainty and macroeconomic surprise factors are related to macroeconomic fluctuations across our EM economies. As seen in Figure 1, the global uncertainty factor captures the crucial events and spikes near elections in the USA, effects likely associated with Brexit, global financial crises, and major uncertainty surrounding fiscal policy in the USA. Given the pivotal role of uncertainty on world trade volume, spillovers from elevated global uncertainty to emerging markets likely affect economic activity through a number of channels. For example, the uncertainty about whether the USA will place tariffs on steel or other goods that are imported from China is likely to undermine export-oriented investments that are allocated to these sectors. This in turn puts pressure on Chinese GDP growth. Also, in the era of Brexit, we have seen that many firms have stopped hiring and are restricting production until the fallout from Brexit becomes more clear. This has a negative impact on GDP growth in countries with high exports to the UK. These sorts of linkages are precisely what we are trying to capture with our uncertainty and surprise factors.

As shown in Figure 2, the global macroeconomic surprise factor tracks the world GDP growth quite closely and captures the sharp contraction in economic activity over the crisis period around 2008. While we might indeed expect markets to move in response to the “surprise” factor, it is noteworthy that the global macroeconomic surprise factor fails to track world GDP growth rates since the beginning of 2017. One reason might be that over this period, post US election, economic agents’ expectations were more pessimistic about the economy than warranted, given elevated global uncertainty in this period.

⁸To this end, we augment each model specification by adding lagged values of GDP growth, where the number of lags is selected using the SIC. Namely, we estimate the following additional models:

- *Specification 5—Local-AR:* $y_{t+h} = \mu + \mathcal{L}Y_t + \beta' F_t^{\text{Local}} + \varepsilon_{t+h}$
- *Specification 6—Uncertainty-AR:* $y_{t+h} = \mu + \mathcal{L}Y_t + \beta' F_t^{\text{Local}} + \vartheta' F_t^{\text{GEPu}} + \varepsilon_{t+h}$
- *Specification 7—Surprise-AR:* $y_{t+h} = \mu + \mathcal{L}Y_t + \beta' F_t^{\text{Local}} + \theta' F_t^{\text{GMS}} + \varepsilon_{t+h}$
- *Specification 8—Global-AR:* $y_{t+h} = \mu + \mathcal{L}Y_t + \beta' F_t^{\text{Local}} + \theta' F_t^{\text{GEPu}} + \delta' F_t^{\text{GMS}} + \varepsilon_{t+h}.$

where $Y_t = (y_t, y_{t-1}, \dots, y_{t-p})$, with p selected using the SIC, and where \mathcal{L} is a conformably defined vectors of coefficients.

⁹Summarizing, we estimate Specifications 1–8 using least squares, after first extracting local, uncertainty, surprise and global factors from the set of targeted predictors selected using the shrinkage approaches discussed in Sections 3.2.1, 3.2.2, 3.2.3, 3.2.4.

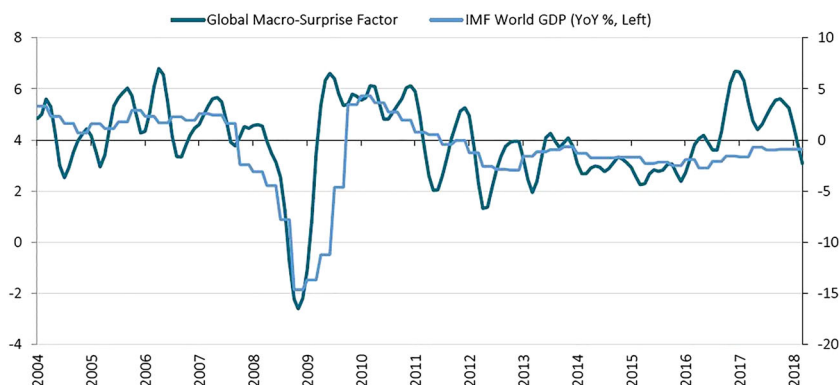


FIGURE 2 World GDP growth rates plotted against global macro-surprise factor [Colour figure can be viewed at wileyonlinelibrary.com]

Correlation coefficients between local and global factors

Brazil	Indonesia	Mexico	S. Africa	Turkey	Surprise	Uncertainty	
Brazil	1	0.52	0.34	0.67	0.38	0.48	-0.27
Indonesia	0.52	1	0.18	0.32	-0.02	0.04	0.14
Mexico	0.34	0.18	1	0.63	0.70	0.40	-0.17
S. Africa	0.67	0.32	0.63	1	0.51	0.31	-0.57
Turkey	0.38	-0.02	0.70	0.51	1	0.51	-0.20
Surprise	0.43	0.04	0.40	0.31	0.51	1	-0.25
Uncertainty	-0.27	0.14	-0.17	-0.57	-0.20	-0.25	1

The table below reports correlation coefficients between the global economic uncertainty factor (called “Uncertainty”), the global macroeconomic surprise factor (called “Surprise”) and country-specific (local) factors. The correlation coefficients between “Uncertainty” and country-specific factors range from -0.17 to -0.57 , showing that the global economic uncertainty factor and economic activity is negatively correlated, except for Indonesia. A reason for this might be that an increase in uncertainty regarding economic policy is likely to trigger recessionary effects, including delays in investment and increases in unemployment. Furthermore, Carriere-Swallow and Cespedes (2013) show that the impact of exogenous uncertainty shocks on emerging economies is to increase the burden of foreign currency denominated debt, due to credit constraints. This may also have an adverse effect on economic growth. On the other hand, the positive relation between “Surprise” and country-specific factors shows that, in general, the macroeconomic data and market expectations move synchronously, as expected. Finally, note that country-specific factors are positively correlated, across countries, probably because of trade linkages.

In order to provide insight into the evolution of country-specific factors (local factors), Figure A1 in the Supporting Information Appendix plots GDP growth against estimated common factors. As is evident from inspection of the plots in this figure, local common factors track GDP growth quite well, and also capture the GDP dynamics in these countries during the global financial crisis.

4.2 | Forecasting experiment results

Table 1 (panels A–E) summarizes the results of our prediction experiments, for each of Brazil, Indonesia, Mexico, South Africa, and Turkey, respectively. As discussed above, there are eight main DFM varieties in our experiment, including Specifications 1–4 (called “Local,” “Uncertainty,” “Surprise,” and “Global,” respectively, in the table), as well as Specifications 1–4 with AR terms (called “Local-AR,” “Uncertainty-AR,” “Surprise-AR,” and “Global-AR,” respectively). Additionally, each of these eight DFM models is estimated using five different types of shrinkage, including “All Sample,” in which all country-specific variables are used in the DFM specification, “LASSO,” “AdaLASSO,” “Bayesian LASSO,” and “Bayesian AdaLASSO.” For complete details, refer to Section 3. In each of the panels, entries in the first row correspond to MSFEs associated with forecasts constructed using our straw-man AR model. All other entries in the panels are relative MSFEs, where the numeraire is MSFE of the AR model. Thus entries that are less than unity indicate point MSFEs that are lower than that of the AR model. For each of two quarterly h -step-ahead forecast horizons (under the headers “ $h = 1$ ” and “ $h = 2$ ”), MSFEs from three monthly forecasts (denoted as months “1,” “2,” and “3”) are reported. Results are also reported for three monthly nowcasts, under the header “ $h = 0$,” and one monthly backcast (under the header “ $h = -1$ ”). For each country, entries denoted in bold indicate the MSFE-“best” model, across all specification types, for a given forecast horizon and shrinkage method. Additionally, for each

TABLE 1 MSFEs based on the use of different-dimension reduction and shrinkage methods with added global diffusion indexes

All sample	Forecast ($h = 2$)			Forecast ($h = 1$)			Nowcast ($h = 0$)			Backcast ($h = -1$)
	1	2	3	1	2	3	1	2	3	
Panel A: Brazil										
AR	4.43	4.43	4.20	3.85	3.85	3.52	2.99	2.99	2.62	2.62
Local	0.96	0.88	0.85	0.77	0.67*	0.62*	0.50*	0.39*	0.33*	0.43**
Uncertainty	0.94	0.85	0.83	0.72*	0.60*	0.57*	0.44*	0.30*	0.30*	0.46**
Surprise	0.96	0.89	0.85	0.78	0.68	0.60*	0.51*	0.39*	0.28*	0.39**
Global	0.92	0.84	0.83	0.72*	0.60*	0.57*	0.45*	0.30*	0.28*	0.44**
Local-AR	0.96	0.89	0.86	0.78	0.68	0.63*	0.51*	0.39*	0.31*	0.38**
Uncertainty-AR	0.93	0.85	0.84	0.72*	0.60*	0.57*	0.44*	0.29*	0.27*	0.43**
Surprise-AR	0.96	0.89	0.86	0.78	0.68	0.60*	0.51*	0.38*	0.27*	0.38**
Global-AR	0.91	0.84	0.84	0.71*	0.60*	0.57*	0.44*	0.29*	0.27*	0.41**
LASSO										
Local	0.90	0.81	0.76	0.72	0.64	0.53**	0.49*	0.42*	0.32*	0.44**
Uncertainty	0.98	0.81	0.81	0.72*	0.58*	0.56**	0.44*	0.32*	0.33*	0.51**
Surprise	0.92	0.82	0.83	0.74	0.66	0.54**	0.49*	0.41*	0.33*	0.41**
Global	0.92	0.80	0.87	0.70*	0.59*	0.57**	0.44*	0.32*	0.34*	0.53*
Local-AR	0.95	0.85	0.85	0.77	0.68	0.66*	0.50*	0.41*	0.35*	0.34**
Uncertainty-AR	1.00	0.83	0.85	0.74*	0.59*	0.57**	0.43*	0.28*	0.25*	0.42**
Surprise-AR	0.96	0.84	0.84	0.75*	0.67*	0.65*	0.48*	0.39*	0.35*	0.35**
Global-AR	0.95	0.82	0.85	0.71*	0.59*	0.57*	0.42*	0.29*	0.25*	0.39**
AdaLASSO										
Local	0.91	0.86	0.85	0.75	0.67	0.63*	0.51*	0.40*	0.33*	0.38**
Uncertainty	0.87	0.82	0.83	0.71*	0.61*	0.59*	0.46*	0.33*	0.31*	0.40**
Surprise	0.91	0.86	0.85	0.77	0.68	0.62*	0.53*	0.40*	0.30*	0.35**
Global	0.85	0.81	0.83	0.69*	0.60*	0.58*	0.46*	0.32*	0.29*	0.37**
Local-AR	0.93	0.87	0.87	0.77	0.68	0.65*	0.52*	0.39*	0.31*	0.34** _{GB}
Uncertainty-AR	0.89	0.83	0.85	0.72*	0.61*	0.59*	0.46*	0.31*	0.27*	0.37**
Surprise-AR	0.93	0.87	0.86	0.77	0.68	0.63*	0.52*	0.40*	0.29*	0.34**
Global-AR	0.87	0.83	0.84	0.70*	0.60*	0.58*	0.45*	0.31*	0.27*	0.39*
Bayesian LASSO										
Local	0.80	0.75	0.76	0.62* _{GB}	0.54*	0.53**	0.28* _{GB}	0.14* _{GB}	0.13*	0.47*
Uncertainty	0.81	0.76	0.77	0.64*	0.56*	0.54**	0.31*	0.16*	0.12* _{GB}	0.45*
Surprise	0.79	0.74	0.75	0.63*	0.54* _{GB}	0.53**	0.30*	0.16*	0.14*	0.46*
Global	0.80	0.75	0.76	0.64*	0.56*	0.54*	0.31*	0.17*	0.13*	0.43**
Local-AR	0.88	0.82	0.82	0.70*	0.61*	0.58*	0.39*	0.24*	0.14*	0.36**
Uncertainty-AR	0.89	0.83	0.83	0.71	0.62*	0.59*	0.41*	0.26*	0.14*	0.35**
Surprise-AR	0.88	0.82	0.82	0.70*	0.61*	0.58*	0.40*	0.24*	0.14*	0.36**
Global-AR	0.88	0.82	0.82	0.71*	0.62*	0.58*	0.40*	0.25*	0.13*	0.35**
Bayesian AdaLASSO										
Local	0.79	0.73 _{GB}	0.71* _{GB}	0.66*	0.58*	0.53**	0.48*	0.41*	0.37*	0.48**
Uncertainty	0.80	0.74	0.72	0.67*	0.59*	0.52** _{GB}	0.48*	0.42*	0.35*	0.47**
Surprise	0.80	0.79	0.75	0.70*	0.65*	0.56**	0.53*	0.48*	0.37*	0.48**
Global	0.76 _{GB}	0.79	0.74	0.65**	0.62	0.53*	0.49*	0.44*	0.33*	0.45**
Local-AR	0.86	0.79	0.79	0.73*	0.65*	0.61**	0.52*	0.43*	0.34*	0.35**
Uncertainty-AR	0.83	0.76	0.75	0.70**	0.60*	0.54**	0.49*	0.39*	0.26*	0.36**
Surprise-AR	0.85	0.80	0.78	0.73*	0.66*	0.59**	0.52*	0.46*	0.31*	0.38**
Global-AR	0.81	0.79	0.77	0.68**	0.62*	0.54*	0.49*	0.40*	0.26*	0.41**

(Continues)

country, entries denoted in bold and subscripted “GB” (for “globally best”) indicate the MSFE-best models across all specification types and shrinkage methods, for a given forecast horizon.¹⁰

¹⁰As mentioned in the Introduction, density forecasting may provide further insights into our analysis. Although this topic is left to future

research, there are many papers espousing the usefulness of such forecasts, and policy analyses used at central banks; for example, routinely include point and density forecasts. As noted by Rossi (2014), “A natural application of forecast densities is the analysis and evaluation of macroeconomic risk. For example, before every Federal Open Market Committee meeting, the Federal Reserve Bank of New York complements its best

TABLE 1 (Continued)

All sample	Forecast ($h = 2$)			Forecast ($h = 1$)			Nowcast ($h = 0$)			Backcast ($h = -1$)
	1	2	3	1	2	3	1	2	3	
Panel B: Indonesia										
AR	1.07	1.07	1.01	1.05	1.05	0.97	0.80	0.80	0.73	0.73
Local	0.83	0.83 _{GB}	0.92	0.72*	0.75*	0.82*	0.76*	0.79	0.88	1.16
Uncertainty	1.35	1.56	1.33	1.25	1.39	1.32	1.21	1.30	1.31	1.40
Surprise	1.13	1.18	1.28	0.98	1.01	1.11	0.94	0.93	1.02	1.07
Global	1.13	1.19	1.24	1.09	1.13	1.23	1.11	1.13	1.23	1.32
Local-AR	0.86	0.86	0.97	0.76*	0.76*	0.85	0.75*	0.71*	0.84	0.94
Uncertainty-AR	1.38	1.55	1.33	1.24	1.33	1.27	1.12	1.17	1.16	1.25
Surprise-AR	1.26	1.21	1.30	1.10	1.00	1.11	1.01	0.89	0.97	1.01
Global-AR	1.28	1.21	1.26	1.21	1.13	1.23	1.17	1.09	1.18	1.30
LASSO										
Local	0.95	0.90	0.93	0.87*	0.87	0.82*	0.89*	0.86	0.88	1.03
Uncertainty	1.60	1.63	1.29	1.43	1.45	1.30	1.36	1.38	1.31	1.45
Surprise	1.26	1.17	1.23	1.15	1.08	1.06	1.13	1.08	0.99	1.11
Global	1.33	1.23	1.19	1.28	1.19	1.17	1.30	1.25	1.18	1.36
Local-AR	0.90	0.87	0.93	0.81*	0.83*	0.83	0.81*	0.79*	0.86	0.98
Uncertainty-AR	1.63	1.70	1.12	1.41	1.44	1.15	1.26	1.27	1.12	1.34
Surprise-AR	1.31	1.22	1.20	1.16	1.09	1.06	1.11	1.00	0.98	1.07
Global-AR	1.36	1.25	1.09	1.29	1.21	1.13	1.29	1.21	1.16	1.38
AdaLASSO										
Local	0.89	0.89	0.92	0.77	0.74	0.76*	0.73*	0.65*	0.63*	0.96
Uncertainty	0.90	1.01	0.96	0.83	0.91	0.85*	0.91	0.92	0.83*	1.09
Surprise	0.93	1.01	1.00	0.83	0.86	0.84*	0.79	0.74*	0.68*	0.96
Global	0.90	0.94	0.89	0.82	0.85	0.79*	0.88	0.85	0.73*	1.05
Local-AR	0.95	0.97	0.97	0.79	0.80	0.75	0.65*	0.61*	0.55*	1.00
Uncertainty-AR	0.82 _{GB}	0.89	0.84 _{GB}	0.67* _{GB}	0.68* _{GB}	0.71*	0.72*	0.72	0.67*	1.11
Surprise-AR	1.04	1.04	1.04	0.86	0.85	0.81	0.71*	0.70*	0.63*	0.95
Global-AR	0.88	0.91	0.85	0.71*	0.70*	0.73*	0.74*	0.76	0.70*	1.08
Bayesian LASSO										
Local	1.32	1.22	1.24	0.92	0.78	0.68*	0.55*	0.56*	0.39*	0.79
Uncertainty	1.25	1.20	1.21	0.85	0.76	0.69*	0.48*	0.57*	0.41*	0.86
Surprise	1.40	1.31	1.30	0.99	0.86	0.78	0.54*	0.54*	0.37*	0.73
Global	1.36	1.28	1.28	0.94	0.83	0.80	0.46*	0.51*	0.39*	0.78
Local-AR	1.21	1.13	1.06	0.85	0.77	0.67* _{GB}	0.43*	0.38*	0.29* _{GB}	0.49
Uncertainty-AR	1.23	1.16	1.17	0.82	0.77	0.75	0.35* _{GB}	0.36* _{GB}	0.32*	0.56
Surprise-AR	1.16	1.09	1.12	0.88	0.82	0.77*	0.53*	0.46*	0.42*	0.43 _{GB}
Global-AR	1.21	1.16	1.20	0.89	0.84	0.82	0.50*	0.46*	0.44*	0.50
Bayesian AdaLASSO										
Local	0.95	0.94	1.03	0.83*	0.80	0.87	0.81*	0.73*	0.82	0.99
Uncertainty	1.04	0.97	1.13	1.05	0.97	1.06	1.20	1.16	1.17	1.43
Surprise	1.15	1.11	1.21	1.02	0.97	1.04	0.99	0.88	0.92	1.08
Global	1.28	0.99	1.12	1.19	0.95	1.04	1.24	1.10	1.09	1.33
Local-AR	1.00	1.00	1.10	0.83	0.83	0.92	0.71*	0.63*	0.75	0.88
Uncertainty-AR	1.09	0.98	0.99	1.03	0.88	0.90	1.09	1.02	0.97	1.36
Surprise-AR	1.22	1.14	1.18	1.04	0.96	1.00	0.91	0.77*	0.83	1.00
Global-AR	1.42	1.16	1.08	1.26	1.01	0.96	1.24	1.09	0.97	1.36

(Continues)

The results in Table 1 reveal various interesting insights. First, there is a substantial reduction in MSFE as one moves from left to right along each row in the panels (i.e., guess of the future path of key macroeconomic variables (the so called ‘modal forecast’) with an assessment of the risk around it (Alessi, Ghysels, Onorante, Peach, & Potter, 2014).” For further discussion of predictive density construction and evaluation, see Corradi and Swanson (2006).

as more data related to the current quarter become available). This increase in forecast accuracy is as expected, and indicates that the DFM models are able to correctly revise GDP prediction by effectively exploiting the flow of monthly data releases (for further discussion see ; Banbura & Rünstler, 2011; Giannone et al., 2008; Li & Chen, 2014). This result is consistent with other emerging-market

TABLE 1 (Continued)

All sample	Forecast ($h = 2$)			Forecast ($h = 1$)			Nowcast ($h = 0$)			Backcast ($h = -1$)
	1	2	3	1	2	3	1	2	3	
Panel C: Mexico										
AR	4.38	4.38	3.98	3.72	3.72	3.26	2.73	2.73	2.25	2.25
Local	0.63	0.53	0.52	0.45*	0.38**	0.44**	0.35**	0.33**	0.60*	0.66
Uncertainty	0.70	0.63	0.54	0.50*	0.45*	0.48**	0.37**	0.36**	0.62*	0.72
Surprise	0.65	0.53	0.57	0.48*	0.39**	0.52**	0.36**	0.33**	0.68	0.78
Global	0.70	0.61	0.57	0.50*	0.43*	0.52**	0.37**	0.34**	0.68	0.79
Local-AR	0.72	0.59	0.55	0.49*	0.40**	0.45**	0.34**	0.32**	0.62*	0.70
Uncertainty-AR	0.75	0.65	0.55	0.51*	0.45*	0.48**	0.36**	0.35**	0.63*	0.74
Surprise-AR	0.67	0.53	0.51	0.47*	0.37**	0.47**	0.33**	0.30**	0.65	0.79
Global-AR	0.70	0.58	0.52	0.49*	0.40*	0.49**	0.35**	0.32**	0.65	0.79
LASSO										
Local	0.68	0.58	0.53	0.47*	0.39**	0.46**	0.32**	0.28**	0.60*	0.71
Uncertainty	0.73	0.69	0.55	0.48*	0.43**	0.48**	0.31**	0.29**	0.61*	0.77
Surprise	0.69	0.55	0.53	0.45*	0.37**	0.48**	0.30**	0.26**	0.64	0.77
Global	0.71	0.65	0.56	0.46*	0.37**	0.48**	0.29**	0.26**	0.64	0.80
Local-AR	0.78	0.64	0.55	0.52*	0.42*	0.46**	0.31**	0.27**	0.60*	0.74
Uncertainty-AR	0.81	0.73	0.57	0.52*	0.45*	0.50**	0.31**	0.28**	0.60*	0.79
Surprise-AR	0.77	0.61	0.52	0.50*	0.39*	0.46**	0.30**	0.25**	0.61*	0.79
Global-AR	0.78	0.70	0.56	0.49*	0.39*	0.48**	0.29**	0.25**	0.62*	0.81
AdaLASSO										
Local	0.53_{GB}	0.48_{GB}	0.48*_{GB}	0.41*_{GB}	0.40*	0.42**_{GB}	0.34**	0.33**	0.52**	0.64*_{GB}
Uncertainty	0.63	0.56	0.61	0.45*	0.44*	0.53**	0.34**	0.32**	0.57*	0.79
Surprise	0.55	0.52	0.52	0.42*	0.42*	0.49**	0.34**	0.34**	0.62*	0.78
Global	0.61	0.54	0.62	0.45*	0.42*	0.56**	0.33**	0.31**	0.64*	0.87
Local-AR	0.64	0.56	0.54	0.46*	0.41*	0.44**	0.31**	0.29**	0.52*	0.67
Uncertainty-AR	0.71	0.64	0.67	0.49*	0.47*	0.58**	0.32**	0.32**	0.56*	0.79
Surprise-AR	0.63	0.57	0.52	0.44*	0.43*	0.46**	0.31**	0.32**	0.58*	0.81
Global-AR	0.67	0.59	0.64	0.46*	0.43*	0.56**	0.31**	0.31**	0.60*	0.87
Bayesian LASSO										
Local	0.91	0.74	0.82	0.47*	0.41*	0.57*	0.19**_{GB}	0.17**	0.18**	0.78
Uncertainty	0.90	0.80	0.83	0.47*	0.43*	0.58*	0.20**	0.21**	0.18**	0.78
Surprise	0.91	0.75	0.83	0.47*	0.41*	0.58*	0.19**	0.16**_{GB}	0.18**_{GB}	0.77
Global	0.90	0.80	0.84	0.47*	0.43*	0.59*	0.21**	0.22**	0.18**	0.77
Local-AR	1.18	0.87	0.99	0.66	0.52*	0.72	0.23**	0.26**	0.31**	0.70
Uncertainty-AR	1.25	0.97	1.03	0.71	0.56*	0.75	0.24**	0.20**	0.33**	0.71
Surprise-AR	1.17	0.88	0.99	0.65	0.52*	0.72	0.23**	0.25**	0.32**	0.70
Global-AR	1.23	0.96	1.04	0.70	0.55*	0.76	0.24**	0.22**	0.34**	0.70
Bayesian AdaLASSO										
Local	0.61	0.53	0.48	0.45*	0.40*	0.43**	0.35**	0.33**	0.58*	0.69
Uncertainty	0.66	0.56	0.50	0.47*	0.41*	0.47**	0.35**	0.34**	0.61*	0.78
Surprise	0.62	0.51	0.48	0.42*	0.36**	0.44**	0.32**	0.31**	0.62*	0.74
Global	0.65	0.52	0.51	0.44*	0.36**_{GB}	0.47**	0.33**	0.31**	0.63*	0.80
Local-AR	0.73	0.62	0.53	0.52*	0.44*	0.46**	0.35**	0.32**	0.59*	0.74
Uncertainty-AR	0.76	0.62	0.54	0.52*	0.44*	0.49**	0.35**	0.33**	0.60*	0.80
Surprise-AR	0.74	0.60	0.52	0.49*	0.41*	0.46**	0.33**	0.31**	0.62*	0.78
Global-AR	0.75	0.60	0.54	0.49*	0.38*	0.49**	0.33**	0.30**	0.62*	0.81

(Continues)

studies in which it is found that DFM-type forecasts and nowcasts are often more accurate than a variety of other benchmark models; see Bragoli, Metelli, and Modugno (2015) for Brazil, Caruso (2015) for Mexico, Dahlhaus, Guénette, and Vasishtha (2017) for BRIC and Mexico, and Luciani et al. (2018) for Indonesia. For example, Dahlhaus et al. adopt judgmental variable selection methods based

on use of the Bloomberg relevance index and expert opinion, and find that DFMs are very flexible when used for modeling emerging markets that are sometimes characterized by more volatile data and structural change than that observed in more advanced economies. Our results add to those discussed in the above papers by analyzing the usefulness of data-driven statistical learning methods.

TABLE 1 (Continued)

All sample	Forecast ($h = 2$)			Forecast ($h = 1$)			Nowcast ($h = 0$)			Backcast ($h = -1$)
	1	2	3	1	2	3	1	2	3	
Panel D: South Africa										
AR	2.55	2.55	2.38	2.10	2.10	1.90	1.44	1.44	1.30	1.30
Local	0.94	0.91	1.10	0.93	0.77	0.90	0.85	0.77	1.01	1.04
Uncertainty	1.04	0.94	0.94	0.86	0.74	0.74*	0.76*	0.78	0.84	0.88
Surprise	0.91	0.82	1.18	0.91	0.76	0.97	0.85	0.73*	0.99	1.03
Global	0.78	0.76	0.99	0.79*	0.71*	0.80	0.72*	0.72**	0.80	0.80
Local-AR	1.28	1.13	1.27	1.04	1.00	1.04	0.89	0.88	1.09	1.14
Uncertainty-AR	1.07	1.14	1.12	0.95	0.98	1.05	0.78*	0.89	1.08	0.98
Surprise-AR	1.25	1.06	1.47	1.01	0.94	1.21	0.88	0.82	1.09	1.12
Global-AR	1.09	1.04	1.05	0.88	0.94	0.93	0.73**	0.88	0.99	0.91
LASSO										
Local	0.82	0.85	1.10	0.81**	0.72**	1.04	0.70**	0.67**	0.86	0.77
Uncertainty	0.92	0.85	1.02	0.82**	0.69**	0.95	0.69**	0.65**	0.80	0.96
Surprise	0.75	0.80	1.18	0.79**	0.68**	1.18	0.68**	0.65	0.93	0.77
Global	0.92	0.87	1.06	0.89	0.75**	1.06	0.80	0.73**	0.82	1.01
Local-AR	0.87	0.86	1.18	0.83**	0.72**	1.16	0.71**	0.64**	0.90	0.68
Uncertainty-AR	0.85	0.80	1.07	0.82*	0.68*	1.01	0.72**	0.67**	0.83	0.75
Surprise-AR	0.77	0.87	1.28	0.76**	0.72**	1.35	0.65**	0.66**	0.99	0.67* _{GB}
Global-AR	0.90	0.85	1.18	0.84	0.72**	1.18	0.81	0.67**	0.88	0.88
AdaLASSO										
Local	0.99	0.99	1.09	0.87*	0.89*	1.09	0.86*	0.89	1.50	0.87
Uncertainty	0.90	0.92	0.83	0.79*	0.82*	0.70*	0.78**	0.86*	0.71*	0.84
Surprise	0.95	0.98	0.93	0.85*	0.89	0.78*	0.85*	0.84**	0.82	0.85
Global	0.91	0.97	0.99	0.80*	0.89	1.03	0.80*	0.89	1.47	0.86
Local-AR	1.20	1.08	1.07	0.99	1.02	0.95	0.88**	1.00	0.94	0.80
Uncertainty-AR	1.10	0.98	0.98	0.90	0.92	0.85	0.82**	0.90	0.94	0.70
Surprise-AR	1.29	1.24	1.12	1.15	1.18	1.21	1.09	1.24	1.88	0.91
Global-AR	1.03	1.02	0.95	0.91	1.03	0.77	0.81**	0.94	0.73*	0.70
Bayesian LASSO										
Local	0.97	0.71	0.79	0.95	0.68**	0.79*	0.81**	0.89*	1.05	1.16
Uncertainty	0.87	0.78	0.79	0.85*	0.71**	0.82*	0.83**	0.92*	1.09	1.17
Surprise	0.97	0.69* _{GB}	0.76	0.89	0.66**	0.77*	0.79**	0.85*	1.01	1.16
Global	0.82	0.75	0.77	0.79*	0.69**	0.77*	0.83*	0.92*	1.05	1.18
Local-AR	0.80	0.74	0.68	0.61**	0.58**	0.53**	0.52**	0.57**	0.70**	0.91
Uncertainty-AR	0.79	0.75	0.68	0.62**	0.57**	0.53**	0.53**	0.61**	0.73**	0.94
Surprise-AR	0.74	0.72	0.68* _{GB}	0.58** _{GB}	0.56**	0.52** _{GB}	0.50** _{GB}	0.56** _{GB}	0.67**	0.88
Global-AR	0.74 _{GB}	0.75	0.69*	0.60**	0.55** _{GB}	0.53**	0.51**	0.61**	0.70**	0.92
Bayesian AdaLASSO										
Local	0.92	0.81	0.91	0.79*	0.65**	0.81	0.73*	0.81	0.77*	0.85
Uncertainty	1.11	0.97	0.85	0.98	0.82	0.77	0.91	1.00	0.64* _{GB}	0.92
Surprise	0.94	0.75	1.04	0.79*	0.55**	0.97	0.71*	0.77*	0.80	0.82
Global	1.12	0.95	1.01	0.96	0.81	0.97	0.95	0.83	0.74*	1.12
Local-AR	1.12	1.04	1.15	0.95	0.93	1.04	0.78*	1.05	0.86	0.93
Uncertainty-AR	1.29	1.17	1.04	1.18	1.06	1.01	1.03	1.25	0.87	1.10
Surprise-AR	1.13	1.00	1.33	0.93	0.88	1.29	0.73*	1.09	0.98	0.87
Global-AR	1.10	1.20	1.16	0.98	1.09	1.16	0.94	0.95	0.94	1.08

(Continues)

Second, virtually all of the entries in Table 1 are less than unity, indicating that the DFM model generally produces smaller point MSFEs than our straw-man or benchmark AR model. Additionally, noting that starred entries indicate rejection of the null of equal predictive accuracy (when comparing a given model against the AR benchmark), it is apparent upon inspection of the tabulated

results that many of these MSFEs are significantly lower than that of the AR model.¹¹

¹¹ Following the suggestion of a referee, we also calculated mean forecast errors (forecast bias) for each period. In general, mean forecast errors are negative and close to zero, implying that there is little evidence of bias in our results. The reason for negative mean forecast errors is that models

TABLE 1 (Continued)

All sample	Forecast ($h = 2$)			Forecast ($h = 1$)			Nowcast ($h = 0$)			Backcast ($h = -1$)
	1	2	3	1	2	3	1	2	3	
Panel E: Turkey										
AR	7.21	7.21	6.58	6.63	6.63	5.94	5.48	5.48	4.69	4.69
Local	0.68	0.65	0.72	0.62**	0.56**	0.60	0.56	0.55	0.66**	0.67
Uncertainty	0.69	0.69	0.75	0.60	0.55	0.58	0.56	0.55	0.66**	0.72
Surprise	0.68 _{GB}	0.64 _{GB}	0.69	0.60**	0.54**	0.56*	0.53	0.53	0.64**	0.74
Global	0.68	0.67	0.73	0.60**	0.55*	0.57	0.55**	0.55	0.64**	0.71
Local-AR	0.71	0.68	0.73	0.62	0.58	0.58	0.51	0.50	0.52	0.58*
Uncertainty-AR	0.71	0.71	0.75	0.60	0.56	0.56*	0.51*	0.50	0.52	0.56*
Surprise-AR	0.73	0.69	0.73	0.62**	0.58**	0.56*	0.50*	0.49	0.48	0.54*
Global-AR	0.72	0.71	0.74	0.62*	0.59*	0.55*	0.51*	0.50	0.48	0.51*
LASSO										
Local	0.73	0.66	0.70	0.64*	0.57**	0.58	0.52*	0.51	0.58*	0.61*
Uncertainty	0.71	0.70	0.73	0.61	0.57	0.58*	0.54	0.53	0.58**	0.66
Surprise	0.73	0.65	0.68* _{GB}	0.63**	0.56**	0.56	0.50*	0.51	0.57	0.71
Global	0.70	0.67	0.71	0.59*	0.56*	0.55*	0.53	0.55	0.57*	0.65
Local-AR	0.79	0.72	0.73	0.66**	0.61**	0.59	0.49*	0.48	0.47	0.56*
Uncertainty-AR	0.75	0.72	0.75	0.63*	0.59	0.57*	0.49*	0.48	0.48	0.54*
Surprise-AR	0.78	0.71	0.73	0.65*	0.60**	0.58	0.48**	0.49	0.45	0.54*
Global-AR	0.74	0.70	0.74	0.62*	0.59	0.56*	0.50	0.50	0.47	0.51*
AdaLASSO										
Local	0.74	0.70	0.81	0.67	0.62	0.70	0.60	0.58	0.70**	0.72
Uncertainty	0.75	0.79	0.82	0.64	0.62	0.64	0.63	0.61	0.69**	0.76
Surprise	0.71	0.68	0.75	0.64	0.60	0.64	0.56	0.56	0.67**	0.78
Global	0.72	0.76	0.79	0.60	0.60	0.61	0.62	0.62	0.68**	0.74
Local-AR	0.75	0.75	0.80	0.66	0.64	0.66	0.54*	0.52	0.51	0.56*
Uncertainty-AR	0.76	0.77	0.79	0.62	0.61	0.61	0.55	0.52	0.54	0.58*
Surprise-AR	0.76	0.75	0.79	0.66	0.63	0.64	0.53*	0.52	0.48*	0.53*
Global-AR	0.73	0.75	0.77	0.60*	0.59*	0.58*	0.55	0.54	0.51*	0.53*
Bayesian LASSO										
Local	0.76	0.73	0.78	0.62*	0.58	0.63	0.40**	0.39*	0.40	0.46 _{GB}
Uncertainty	0.74	0.72	0.78	0.58*	0.52	0.59	0.37**	0.36*	0.37*	0.61*
Surprise	0.75	0.72	0.76	0.61*	0.58	0.61	0.39**	0.38**	0.38*	0.65*
Global	0.74	0.72	0.78	0.57*	0.52	0.59	0.36** _{GB}	0.35*	0.36*	0.62*
Local-AR	0.94	0.82	0.84	0.69	0.61	0.62	0.38**	0.36*	0.32**	0.55*
Uncertainty-AR	0.93	0.82	0.83	0.67	0.59	0.61	0.37**	0.35*	0.30**	0.54*
Surprise-AR	0.96	0.83	0.83	0.69	0.62	0.62	0.38**	0.36*	0.31**	0.55*
Global-AR	0.94	0.82	0.83	0.67	0.59	0.60	0.36**	0.35* _{GB}	0.30** _{GB}	0.53*
Bayesian AdaLASSO										
Local	0.76	0.67	0.72	0.62*	0.55*	0.58	0.46**	0.45	0.50	0.53*
Uncertainty	0.73	0.70	0.74	0.54	0.50	0.54**	0.42**	0.43	0.48	0.65
Surprise	0.76	0.66	0.69	0.62*	0.54*	0.56	0.44**	0.44	0.49	0.72
Global	0.72	0.69	0.71	0.54* _{GB}	0.50* _{GB}	0.52**	0.42**	0.44	0.48	0.65
Local-AR	0.85	0.71	0.73	0.64	0.56*	0.57*	0.42**	0.42*	0.43	0.61
Uncertainty-AR	0.81	0.72	0.75	0.58	0.52	0.54**	0.39**	0.39*	0.42*	0.58*
Surprise-AR	0.84	0.69	0.72	0.63	0.55*	0.55*	0.41**	0.41*	0.42*	0.61*
Global-AR	0.78	0.70	0.72	0.57	0.51	0.52** _{GB}	0.39**	0.40	0.41*	0.56*

Note. Entries in this table are MSFEs associated with various forecasting models for which latent factors are selected using the shrinkage methods given in bold in the first column of the table. The forecasting models, depicted as Local, Uncertainty, Surprise, Global, Local-AR, Uncertainty-AR, Surprise-AR, and Global-AR, correspond to Specifications 1–8, respectively (see Section 3.3). Models that yield the smallest MSFE are denoted in bold, for each estimation method and forecast horizon; and models that are denoted in bold with a “GB” subscript denote models that are “globally” MSFE-“best”, across all models and estimation methods, for a given forecast horizon. Entries in the first row of the table are point MSFEs based our benchmark AR(SIC) model, while the rest of the entries in the table are relative MSFEs (i.e., relative to the AR(SIC) benchmark model). Thus a value of less than unity indicates that a particular model and estimation method is more accurate than that based on the AR(SIC) benchmark, for a particular forecast horizon. Quarterly forecast horizons are denoted by $h = -1, 0, 1$, or 2 ; and monthly forecasts within each of these quarters are denoted by month 1, 2, or 3. Entries followed by asterisks (**5% level; *10% level) are significantly superior to the AR(SIC) benchmark model, based on application of the Diebold and Mariano (1995) predictive accuracy test. All models are listed in the first column of the table, and Local, Uncertainty, Surprise, and Global correspond to Specifications 1–4 from Section 3.3, respectively; while the models appended with “-AR” are the same as Specifications 1–4, but with additional lagged dependent variables added as regressors. Entries in each panel show results of mentioned specification types where factors are extracted based on the targeted predictors selected by using the corresponding shrinkage method. For complete details, refer to Section 3.

Third, recall that there are 10 forecast horizons and five shrinkage methods, so there are a total of 50 specification types for each country. Our results indicate that there are notable decreases in MSFE for a number of countries when “uncertainty” and “surprise” factors are extracted from our uncertainty and surprise index datasets are included in the DFM.¹² Thus latent factors that capture uncertainty and macro data surprise appear to contain significant marginal predictive content for country-specific growth prediction. For example, note that in Table 1 (panel A; i.e., the case of Brazil) the inclusion of factors extracted from uncertainty and/or surprise index datasets results in the globally MSFE-best models for 4 of 10 forecast horizons, across all shrinkage methods. For Indonesia (see panel B), we see that, for 7 of 10 forecast horizons, the inclusion of uncertainty and/or surprise factors results in the globally MSFE-best models. For Mexico, South Africa, and Turkey, the number of analogous “wins” for our global factors are 3 of 10, 10 of 10, and 9 of 10, respectively. Overall, therefore, globally MSFE-best models include our global uncertainty and surprise factors in 33 of 50 cases, across all five countries. In summary, there is strong evidence that global factors play a key role in GDP growth for Indonesia, South Africa, and Turkey, whereas the evidence is mixed for Brazil and Mexico. This result is consistent with an argument that the low level of domestic savings in Turkey and South Africa results in heavy reliance on capital inflows to finance new investment projects. This in turn is reflected in the importance of global factors for these countries, since global swings in economic policy uncertainty are a key driver of capital flows. On the other hand, Indonesia is a leading exporter of various commodities and is heavily dependent on exports to sustain economic growth. This makes growth in Indonesia vulnerable to global demand shocks and to increases in trade protectionism.

Fourth, if one digs more deeply into the findings concerning Brazil, a further pattern emerges. In particular, note that for Brazil 3 of the 4 “wins” discussed above occur when forecasting, whereas only 1 of 4 “wins” occurs when nowcasting or backcasting. Thus, for this country, global information appears to lose its predictive content, relative to local information, as the calendar date of

available monthly data becomes closer to (or surpasses) the calendar date of the GDP value being predicted. This result may be explained by the fact that the main driver of the Brazil GDP growth is private consumption and construction. For Mexico the story is different, as local information remains the most useful, regardless of forecast horizon. One reason for this might be that our global factors do not adequately account for shocks that are important to Mexico. This may in turn be due, in part, to the fact that Mexico has approximately the 15th largest economy in the world, in nominal terms, and exhibits a surprising level of macroeconomic stability. Moreover, Mexico has a largely export-oriented economy with manufactured products accounting for approximately 90% of all exports.

In order to determine whether the above findings are apparent upon visual inspection of DFM predictions, we have plotted nowcasts against actual GDP growth for the five countries (see Figure A2 in the Supporting Information Appendix). In that figure, the “Actual GDP” plot is of actual GDP growth rates, “AR” corresponds to nowcasts made with our AR model using all available data (i.e., no shrinkage), and “Local” corresponds to DFM nowcasts made using all available data. Finally, for each country, we also include a “Local,” “Surprise,” or “Global” plot, corresponding to nowcasts made with the MSFE-best DFM model that includes “uncertainty” and/or “surprise” factors. Examination of these plots indicates that specification types based on uncertainty and macroeconomic data surprises tend to predict turning points relatively well, and outperform the “Local” factor model in volatile episodes, particularly for Brazil, South Africa, and Turkey. However, broad rankings of the variety available by summarizing the data in Table 1 cannot be made by inspecting the plots in Figure A2 in the Supporting Information Appendix. For this reason, we provide an alternative summary of the information given in Table 1. This summary is given in Table 2 which shows that the number of model “wins” is tabulated by model specification type, across all forecast horizons. Results in this table corroborate the findings above concerning the usefulness of our global factors for predicting GDP growth in EM countries.

Summarizing, our findings with respect to the importance of “uncertainty” factors are consistent with business cycle synchronization studies that focus on the growing financial and trade integration of world economies, and note that this integration is likely to result in stronger spillovers of shocks across economies. Additionally, our findings with regard to the importance of “surprise” factors underscore the significance of expectations of economic agents, which in turn reflect the importance of the ever greater amounts of “soft” news that is available to agents,

tend to under-predict the quick recovery in growth rates of emerging markets after the global financial crisis. This finding is consistent with our finding that MSFEs decrease appreciably as more data related to current quarter become available. These results are summarized in Tables A3 to A7 of the Supporting Information Appendix.

¹²Recall that entries labeled “Uncertainty” in the tables correspond to DFMs with “uncertainty” factors (i.e., Specification 2), entries labeled “Surprise” correspond to DFMs with “surprise” factors (i.e., Specification 3), and entries labeled “Global” correspond to DFMs with “uncertainty” and “surprise” factors (i.e., Specification 4).

	Brazil	Indonesia	Mexico	S. Africa	Turkey
AR	0	3	0	1	0
Local	10	11	25	1	2
Uncertainty	3	0	2	15	1
Surprise	5	0	7	11	12
Global	12	0	1	8	13
Local-AR	3	24	3	1	0
Uncertainty-AR	5	10	0	1	2
Surprise-AR	1	2	8	10	6
Global-AR	11	0	4	2	14

Note. See note to Table 1. Fifty cases are summarized in this table. Namely, there are five shrinkage types and 10 prediction horizons in Table 1. Across these 50 cases, MSFE best models are tabulated, by specification types, where the forecasting models depicted as Local, Uncertainty, Surprise, Global, Local-AR, Uncertainty-AR, Surprise-AR, and Global-AR correspond to Specifications 1–8, respectively (see Section 3.3). Finally, AR is the benchmark autoregressive model.

and that may not be reflected in the timely release of macroeconomic data.¹³

We now turn to a discussion of the usefulness of data shrinkage in our experiments. Even casual inspection of the findings in Table 1 indicates that preselection of indicators using shrinkage delivers models with more accurate forecasts than when DFMs are estimated without preselection. In particular, when comparing the “globally MSFE-best” models in Table 1, which we have defined to be the MSFE-best models for each country across all specification and shrinkage methods, we see that 47 of the 50 “globally best” models utilize shrinkage.¹⁴ This suggests that the selection of relevant predictors from a large dataset mitigates data noisiness, and model (and coefficient estimator) imprecision due to multicollinearity, as might be expected. This result corroborates that of Boivin and Ng (2006), where it is suggested that the possibility of correlated errors increase as more series from the same “category” are included in a dataset, and creates a situation where more data might not be desirable. Among the different shrinkage methods that we utilize in our experiment, the Bayesian methods (Bayesian LASSO and Bayesian AdaLASSO) perform surprisingly well, as they attain the top rank in 36 out of 50 cases. This evidence strongly supports the use of Bayesian shrinkage methods for variable selection.¹⁵ These results are summarized in Tables 2 and 3.

¹³For a summary of a small experiment in which we use only real variables in a nowcasting experiment, see Tables A1–A2 in the Supporting Information Appendix. Our results are consistent with findings in Bragoli and Fosten (2018), Banbura, Giannone, Modugno, and Reichlin (2013), and Andreou, Ghysels, and Kourtellis (2013) concerning the importance of carefully managing the number of real and nominal variables, and of imposing sparsity in forecasting exercises.

¹⁴Recall that there are 10 forecast horizons and five countries, so that we have total of 50 cases.

¹⁵See De Mol, Giannone, and Reichlin (2008), where it is found that a few appropriately selected variables often capture the bulk of the covariation in large macroeconomic datasets that are characterized by collinearity,

TABLE 2 The number of MSFE-“wins” by specification type, across all shrinkage methods and prediction horizons

5 | CONCLUDING REMARKS

DFMs are widely used in the forecasting literature. In this paper, we add to this literature by utilizing these models to predict emerging market (EM) GDP growth rates. Moreover, we augment the standard DFM type models that we examine with global “uncertainty” and “surprise” factors that are meant to capture the deepening interdependencies among all countries in the world. To do this, we construct three new datasets—one focusing on country-specific data, and two focusing on worldwide uncertainty, sentiment, and so-called “surprise” data—from which global latent factors are extracted. These factors are constructed using both standard estimation methods as well as via implementation of a number of LASSO-type shrinkage methods. We find that our new global economic “uncertainty” and macroeconomic data “surprise” factors are indeed useful, in the sense that they contain substantial marginal predictive content for real GDP growth in numerous EM economies, as shown through a series of real-time forecasting experiments. Moreover, the data shrinkage methods employed in our experiments are found to significantly improve the predictive content of the latent factors used in our DFMs.

This paper is meant as a starting point, as many questions remain unanswered. For example, it will be of interest to investigate the regime-dependent impact of uncertainty shocks on growth by distinguishing effects associated with phases of the business cycle. Also, it remains to analyze the impacts of possible changes in information rigidities in consensus forecasts (such as those used in the construction of our global “surprise” factors) when the economy moves from one phase of the business cycle to

and use of these small groups of variables often yields comparable forecasting performance, relative to the case where all variables are used when constructing predictions.

TABLE 3 Summary of the “globally best” models across all shrinkage methods and specification types, for each prediction horizon

	Forecast ($h = 2$)			Forecast ($h = 1$)			Nowcast ($h = 0$)			Backcast ($h = -1$)	
	1	2	3	1	2	3	1	2	3	1	1
	Brazil	BaLASSO-4	BaLASSO-1	BaLASSO-1	BLASSO-1	BLASSO-2	BaLASSO-2	BLASSO-1	BLASSO-2	BLASSO-2	AdaLASSO-5
Indonesia	AdaLASSO-6	ALL-1	AdaLASSO-6	AdaLASSO-6	AdaLASSO-6	AdaLASSO-6	BLASSO-5	BLASSO-6	BLASSO-5	BLASSO-5	BLASSO-7
Mexico	AdaLASSO-1	AdaLASSO-1	AdaLASSO-1	AdaLASSO-1	BaLASSO-1	BaLASSO-1	AdaLASSO-1	BLASSO-1	BLASSO-3	BLASSO-3	AdaLASSO-1
S. Africa	BLASSO-8	BLASSO-3	BLASSO-7	BLASSO-7	BLASSO-8	BLASSO-8	BLASSO-7	BLASSO-7	BaLASSO-2	BLASSO-7	LASSO-7
Turkey	ALL-3	ALL-3	LASSO-3	BaLASSO-4	BaLASSO-4	BaLASSO-4	BaLASSO-8	BLASSO-4	BLASSO-8	BLASSO-8	BLASSO-1

Note. See notes to Table 1. The abbreviations “-1” to “-8” denote Specifications 1–8. Entries in the table denote “globally best” forecasting models, in the sense that the listed models are MSFE-best, across all specification types and shrinkage methods. See Sections 3 and 4 for further details.

another. It also remains to asymptotically analyze extant shrinkage methods, with an eye to the development of new and improved estimation algorithms.

ACKNOWLEDGMENTS

The authors wish to thank Hyun Hak Kim, Mingmian Cheng, Eric Ghysels, Massimiliano Marcellino, Christian Schumacher, and Xiye Yang for useful discussions related to the content of this paper on modeling and forecasting using diffusion indexes. Many thanks are also owed to the Editor, Massimiliano Marcellino, and to two anonymous referees for their many useful comments and suggestions during the revision process of this paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from Bloomberg terminal. Restrictions apply to the availability of these data, which were used under license for this study. However, we shared the corresponding tickers in Tables A8–A14 of the Supporting Information Appendix. Hence anyone who has a Bloomberg account can easily download the same data.

ORCID

I. Ethem Guney  <https://orcid.org/0000-0001-6890-205X>

REFERENCES

- Abbate, A., & Marcellino, M. (2018). Point, interval and density forecasts of exchange rates with time varying parameter models. *Journal of the Royal Statistical Society, Series A*, 181(1), 155–179.
- Ahn, S. C., & Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3), 1203–1227.
- Alessi, L., Barigozzi, M., & Capasso, M. (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics and Probability Letters*, 80(23–24), 1806–1813.
- Alessi, L., Ghysels, E., Onorante, L., Peach, R., & Potter, S. (2014). Central bank macroeconomic forecasting during the global financial crisis: The European Central Bank and the Federal Reserve Bank of New York experiences. *Journal of Business and Economic Statistics*, 32(4), 483–500.
- Andreou, E., Ghysels, E., & Kourtellis, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business and Economic Statistics*, 31(2), 240–251.
- Artis, M. J., Banerjee, A., & Marcellino, M. (2005). Factor forecasts for the UK. *Journal of Forecasting*, 24(4), 279–298.
- Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70, 191–221.
- Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146, 304–317.
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4), 1593–1636.

- Banbura, M., Giannone, D., Modugno, M., & Reichlin, L. (2013). Now-casting and the real-time data-flow. In G. Elliott, & A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Vol. 2. Amsterdam, Netherlands: North-Holland, pp. 195–237.
- Banbura, M., & Rünstler, G. (2011). A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting*, 27, 333–346.
- Banerjee, A., & Marcellino, M. (2006). Are there any reliable leading indicators for US inflation and GDP growth? *International Journal of Forecasting*, 22(1), 137–151.
- Bessec, M. (2013). Short-term forecasts of French GDP: A dynamic factor model with targeted predictors. *Journal of Forecasting*, 32(6), 500–511.
- Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28(2), 153–176.
- Boivin, J., & Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132, 169–194.
- Bontempi, M. E., Golinelli, R., & Squadrani, M. (2016). A new index of uncertainty based on internet searches: A friend or foe of other indicators? (*Quaderni-Working Paper DSE No. 1062*). Bologna, Italy: Department of Economics, University of Bologna.
- Bragoli, D. (2017). Now-casting the Japanese economy. *International Journal of Forecasting*, 33(2), 390–402.
- Bragoli, D., & Fosten, J. (2018). Nowcasting indian GDP. *Oxford Bulletin of Economics and Statistics*, 80(2), 259–282.
- Bragoli, D., Metelli, L., & Modugno, M. (2015). The importance of updating: Evidence from a Brazilian nowcasting model. *Journal of Business Cycle Measurement and Analysis*, 2015(1), 5–22.
- Bulligan, G., Marcellino, M., & Venditti, F. (2015). Forecasting economic activity with targeted predictors. *International Journal of Forecasting*, 31, 188–206.
- Caggiano, G., Kapetanios, G., & Labhard, V. (2011). Are more data always better for factor analysis? Results for the euro area, the six largest euro area countries and the UK. *Journal of Forecasting*, 30(8), 736–752.
- Carriere-Swallow, Y., & Cespedes, L. F. (2013). The impact of uncertainty shocks in emerging economies. *Journal of International Economics*, 90(2), 316–325.
- Carriero, A., Clark, T. E., & Marcellino, M. (2016). Common drifting volatility in large Bayesian VARs. *Journal of Business and Economic Statistics*, 34(3), 375–390.
- Caruso, A. (2015). Nowcasting Mexican GDP. *ECARES Working Paper*, 40.
- Cepni, O., Güney, I. E., & Swanson, N. R. (2019). Nowcasting and forecasting GDP in emerging markets using global financial and macroeconomic diffusion indexes. *International Journal of Forecasting*, 35(2), 555–572.
- Corradi, V., & Swanson, N. R. (2006). Predictive density evaluation. In G. Elliott, C. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Vol. 1. Amsterdam, Netherlands: North-Holland, pp. 197–284.
- Dahlhaus, T., Guénette, J. D., & Vasishtha, G. (2017). Nowcasting BRIC+ M in real time. *International Journal of Forecasting*, 33(4), 915–935.
- De Mol, C., Giannone, D., & Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, 146(2), 318–328.
- Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 20, 134–144.
- Fan, J., & Li, R. (2001). Variable selection via non-concave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456), 1348–1360.
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and Statistics*, 82, 540–554.
- Fu, W. J. (1998). Penalized regressions: The bridge versus the LASSO. *Journal of Computational and Graphical Statistics*, 7, 397–416.
- Gauvin, L., McLoughlin, C., & Reinhardt, D. (2014). Policy uncertainty spillovers to emerging markets: Evidence from capital flows. (*Working Paper No.512*). London, UK: Bank of England.
- Ghysels, E., Osborn, D. R., & Sargent, T. J. (2001). *The econometric analysis of seasonal time series*. Cambridge: Cambridge University Press.
- Giannone, D., Agrippino, S. M., & Modugno, M. (2013). Nowcasting China real GDP. New York, NY: Mimeo.
- Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4), 665–676.
- Girardi, A., Gayer, C., & Reuter, A. (2015). The role of survey data in nowcasting euro area GDP growth. *Journal of Forecasting*, 35(5), 400–418.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised learning, *The elements of statistical learning*. New York, NY: Springer, pp. 485–585.
- Kapetanios, G., Marcellino, M., & Papailias, F. (2016). Forecasting inflation and GDP growth using heuristic optimisation of information criteria and variable reduction methods. *Computational Statistics and Data Analysis*, 100, 369–382.
- Kim, H. H., & Swanson, N. R. (2014). Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. *Journal of Econometrics*, 178, 352–367.
- Kim, H. H., & Swanson, N. R. (2018a). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting*, 34, 339–354.
- Kim, H. H., & Swanson, N. R. (2018b). Methods for backcasting, nowcasting and forecasting using factor MIDAS: With an application to Korean GDP. *Journal of Forecasting*, 37(3), 281–302.
- Korobilis, D. (2013). Hierarchical shrinkage priors for dynamic regressions with many predictors. *International Journal of Forecasting*, 29(1), 43–59.
- Li, J., & Chen, W. (2014). Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting*, 30(4), 996–1015.
- Liu, P., Matheson, T., & Romeu, R. (2012). Real-time forecasts of economic activity for Latin American economies. *Economic Modelling*, 29(4), 1090–1098.
- Luciani, M., Pundit, M., Ramayandi, A., & Veronese, G. (2018). Nowcasting indonesia. *Empirical Economics*, 55(2), 597–619.
- Luciani, M., & Ricci, L. (2014). Nowcasting norway. *International Journal of Central Banking*, 10(4), 215–248.
- Mariano, R. S., & Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4), 427–443.
- Medeiros, M. C., & Mendes, E. F. (2016). L1-regularization of high-dimensional time-series models with non-Gaussian and heteroskedastic errors. *Journal of Econometrics*, 191(1), 255–271.
- Medeiros, M. C., & Vasconcelos, G. F. (2016). Forecasting macroeconomic variables in data-rich environments. *Economics Letters*, 138, 50–52.

- Modugno, M., Soybilgen, B., & Yazgan, E. (2016). Nowcasting Turkish GDP and news decomposition. *International Journal of Forecasting*, 32, 1369–1384.
- Onatski, A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica*, 77(5), 1447–1479.
- Park, T., & Casella, G. (2008). The bayesian LASSO. *Journal of the American Statistical Association*, 103(482), 681–686.
- Rossi, B. (2014). Density forecasts in economics, forecasting and policymaking. (*Working Paper No.37*). Barcelona, Spain: Centre de Recerca en Economia Internacional (CREI).
- Rossi, B., & Sekhposyan, T. (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *American Economic Review*, 105(5), 650–655.
- Schumacher, C. (2007). Forecasting German GDP using alternative factor models based on large datasets. *Journal of Forecasting*, 26(4), 271–302.
- Schumacher, C. (2010). Factor forecasting using international targeted predictors: The case of German GDP. *Economics Letters*, 107, 95–98.
- Scotti, C. (2016). Surprise and uncertainty indexes: Real-time aggregation of real-activity macro-surprises. *Journal of Monetary Economics*, 82, 1–19.
- Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, 20, 147–162.
- Swanson, N. R. (2016). Comment on: In sample inference and forecasting in misspecified factor models. *Journal of Business and Economic Statistics*, 34, 348–353.
- Swanson, N. R., & Urbach, R. (2015). Prediction and simulation using simple models characterized by nonstationarity and seasonality. *International Review of Economics and Finance*, 40, 312–323.
- Thorsrud, L. A. (2018). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business and Economic Statistics*. Advance online publication <https://doi.org/10.1080/07350015.2018.1506344>
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society, Series B*, 73(1), 273–282.
- Uhlig, H. (2009). Comment on how has the Euro changed the monetary transmission. In D. Acemoglu, K. Rogoff, & M. Woodford (Eds.), *NBER Macroeconomics Annual* (pp. 141–152). Chicago: University of Chicago Press.
- Zou, H. (2006). The adaptive LASSO and its oracle properties. *Journal of the American Statistical Association*, 101(476), 1418–1429.

AUTHOR BIOGRAPHIES

Oguzhan Cepni is a Deputy Executive Director at the Central Bank of the Republic of Turkey. His primary research interests include applied econometrics, advanced analytics, machine learning and factor models. On these topics he has published widely in *International Journal of Forecasting*, *Emerging Markets Review*, *International Review of Finance* and *Financial Research Letters*.

Ibrahim Ethem Guney has a Ph.D. in Banking & Finance from University of Zurich and Swiss Finance Institute. He is currently Executive Director of Strategy and Corporate Governance Department at the Central Bank of the Republic of Turkey. His primary research interests include monetary policy, asset pricing, applied financial economics, mathematical finance and big data applications. Dr. Guney has publications *International Journal of Forecasting*, *Emerging Markets Review*, *Journal of Mathematical Economics* and *Applied Economics Letters*.

Norman R. Swanson was educated at the University of Waterloo and the University of California, San Diego. He is currently Professor and Director of Graduate Studies in the Economics Department at Rutgers University. He has held previous positions at Pennsylvania State University (1994–1999), Texas A&M University (1999–2001), and Purdue University (2001–2002). His primary research interests include econometric theory, financial and macro econometrics, time series analysis, and forecasting. He is a fellow of the *Journal of Econometrics*, he is currently an associate editor of various econometrics journals, and he has served as guest editor for journals ranging from *Journal of Econometrics* to *Journal of Business and Economic Statistics*. Swanson has recent scholarly publications in leading economics and statistics journals including *Econometrica*, *Journal of Econometrics*, *Review of Economics and Statistics*, *Journal of Business and Economic Statistics*, and the *Journal of the American Statistical Association*, among others; and is or has acted as ongoing visiting scholar to various central banks including The Federal Reserve Bank of Philadelphia and the Bank of Canada.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Cepni O, Guney IE, Swanson NR. Forecasting and nowcasting emerging market GDP growth rates: The role of latent global economic policy uncertainty and macroeconomic data surprise factors. *Journal of Forecasting*. 2020;39:18–36. <https://doi.org/10.1002/for.2602>