

# An Assessment of the Marginal Predictive Content of Economic Uncertainty Indexes and Business Conditions Predictors \*

Yang Liu and Norman R. Swanson  
Rutgers University

this version: October 2023

## Abstract

In this paper, we evaluate the marginal predictive content of a variety of new business conditions (BC) predictors as well as nine economic uncertainty indexes (EUIs) constructed using these predictors. Our predictors are defined as observable variables and latent factors extracted from a high dimensional macroeconomic dataset and our EUIs are functions of predictive errors from models that incorporate these predictors. Estimation of the predictors is based on a number of extant and novel machine learning methods that combine dimension reduction, variable selection, and shrinkage. When predicting 14 monthly U.S. economic series selected from 8 different groups of economic variables, our new indexes and predictors are shown to result in significant improvements in forecast accuracy, relative to predictions made using benchmark models. Moreover, while inclusion of either BC predictors or EUIs often yields forecast accuracy improvements, greater predictive gains accrue when using BC predictors with real economic activity type variables.

*Keywords:* Latent factor, business conditions predictors, macroeconomic uncertainty measure, principal components analysis, least absolute shrinkage operator, high dimensional data, big data.

---

\*Yang Liu, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, yl1241@economics.rutgers.edu. Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@economics.rutgers.edu. The authors would like to thank the editor, Esther Ruiz, as well as an associate editor and two anonymous referees for useful suggestions on earlier drafts of this paper. We are also grateful to Mingmian Cheng, Valentina Corradi, Frank Diebold, Hyun Hak Kim, John Landon-Lane, Yuan Liao, Weijia Peng, and Chun Yao for useful comments and suggestions on the topics explored in this paper.

# 1 Introduction

The creation of accurate forecasts of economic variables using econometric models has been central to the economics profession for almost 100 years. In the early days, econometric models were used to examine economic concepts and model business cycles (see e.g. [Tinbergen \(1939\)](#)). In modern times, econometric forecasting models are also used to examine the impact of policy setting behavior at the government level and to aid in business decisions at the firm level. A recent iteration of the vast related literature, spurred in large part by the availability of high frequency and high dimensional data (i.e., big data) has focused on the development and assessment of various (latent) economic predictor variables as well as business conditions and economic uncertainty indexes.

One key recent business conditions index is that proposed by [Aruoba, Diebold, and Scotti \(2009\)](#), called the ADS index in the sequel. These authors extract business conditions indexes from key macroeconomic variables of different observational frequencies.<sup>1</sup> In addition to such indexes, business conditions are often also predicted using “predictor” variables extracted by implementing machine learning techniques. For example, [Cepni, Guney, and Swanson \(2019\)](#) extract latent global factors from a large-scale multi-country dataset, [Bulligan, Marcellino, and Venditti \(2015\)](#) utilize predictor selection methods coupled with latent factor estimation for forecasting Italian GDP, [Schumacher \(2007\)](#) forecasts German GDP using latent factors, [Marcellino, Porqueddu, and Venditti \(2016\)](#) specifies a mixed-frequency dynamic factor model and investigate business cycles in the euro area, and [Andreou, Ghysels, and Kourtellis \(2013\)](#) utilize the Mixed Data Sampling (MIDAS) framework developed in [Ghysels, Sinko, and Valkanov \(2007\)](#) to include daily data when forecasting lower frequency macroeconomic variables.<sup>2</sup> In all of these papers, models are specified that include latent factors which can be interpreted as business conditions predictors.

Examples of economic uncertainty indexes (EUIs) are discussed in [Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry \(2018\)](#), [Baker, Bloom, and Davis \(2016\)](#), [Carriero,](#)

---

<sup>1</sup>A 6-variable variant of this index is updated regularly by the Federal Reserve Bank of Philadelphia.

<sup>2</sup>Related papers that explore the usefulness of mixed-frequency state space models include [Mariano and Murasawa \(2003\)](#) and [Frale, Marcellino, Mazzi, and Proietti \(2008\)](#).

Clark, and Marcellino (2016), Jo and Sekkel (2017). One particularly interesting EUI that we discuss in detail in this paper, and which is called JLNI in the sequel, is due to Jurado, Ludvigson, and Ng (2015). This index is constructed using prediction errors from forecasts made using machine learning methods. In a related strand of the literature, uncertainty is instead measured directly, by examining volatilities of business conditions indexes constructed using mixed frequency models (see, e.g. Peng, Swanson, Yang, and Yao (2023)), or by directly estimating volatility using high frequency financial times series variables (see, e.g. Chauvet, Senyuz, and Yoldas (2015) and Cheng, Swanson, and Yang (2021)).

In this paper, we propose a number of new business conditions predictors, as well as nine new EUIs that are constructed using these predictors. In order to measure the usefulness of our new predictors and indices, we assess their marginal predictive content when used to forecast fourteen financial and non-financial economic variables. As a benchmark, forecasts from models that contain our predictors and indices are compared against simple autoregressive models as well as forecasting models that include the ADS index, the JLN index, and the VIX.<sup>3</sup> Our predictors and indexes are extracted from a large dimensional macroeconomic dataset using machine learning methods. Our dataset includes 122 variable partitioned into 8 groups. The groups include: output and income, labor market, housing, consumption and inventories, money and credit, interest and exchange rates, prices, and the stock market.<sup>4</sup> In our analysis, we begin by constructing various business conditions predictors by jointly applying a number of least absolute shrinkage operator (lasso) and principal components analysis (PCA) methods to a high dimensional dataset, say  $\mathbf{X}$ , where  $\mathbf{X}$  is an  $N \times T$  matrix, with  $T$  denoting the number of time series observations in the sample, and  $N$  denoting the number of variables in the dataset. These predictors are subsequently used as explanatory variables in econometric forecasting models. Forecasts errors from these models are then used to construct our EUIs, as in Jurado et al. (2015).<sup>5</sup> With these new predictors and

---

<sup>3</sup>Note that Ludvigson and Ng (2009) propose multiple indices. As our analysis is real-time, we use their “real uncertainty index” discussed in the appendix to their paper. This index is truly real-time, in the sense that model parameters are estimated in real-time using data available prior to the calendar date of each index value, and is thus directly comparable with our business conditions predictors and EUIs. This real-time index is available from Sydney Ludvigson’s website.

<sup>4</sup>The dataset used in Jurado et al. (2015) is quite similar, but contains an additional 10 variables.

<sup>5</sup>See also Ludvigson and Ng (2009) and Gu, Kelly, and Xiu (2020) for a discussion of related research

EUIs in hand, we undertake a comprehensive set of forecasting experiments to assess their marginal predictive content. These experiments involve predicting 14 variables chosen from 8 different groups of macroeconomic variables, covering areas ranging from banking and finance to employment.<sup>6</sup>

Before continuing, it is worth stressing that the predictors and EUIs examined in this paper are designed to contain useful information for forecasting economic conditions. In particular, it is posited that the predictive accuracy of standard autoregressive forecasting models may be improved by including our new variables in these models. This is important, given that governments, firms, and individuals all rely on accurate forecasts when making decisions. For example, insurance firms and banks utilize forecasts of government and corporate bond yields in order to aid them in asset allocation. When forecasts are more precise, investment decisions are improved and regulatory capital requirements can be more precisely met, leading to increased profitability. Additionally, governments utilize predictions of a whole host of economic variables when forecasting the impact of possible interest rate changes on variables like inflation and GDP growth. In our analysis, we show that a variety of macroeconomic variables can be more precisely predicted when business conditions predictors or economic uncertainty indexes are added to standard autoregressive forecasting models of the variety widely utilized in government and industry.

In summary, our paper builds on various key papers on diffusion index modeling by Jushan Bai, Serena Ng, James Stock and Mark Watson (see, e.g. [Bai and Ng \(2002\)](#), [Stock and Watson \(2002a\)](#), [Stock and Watson \(2002b\)](#), [Bai and Ng \(2008\)](#), and [Bai and Ng \(2009\)](#), and the references cited therein), in which business conditions predictors are constructed by

---

that uses machine learning methods to extract asset risk premia. Various other papers also utilize approaches that are related to ours for measuring uncertainty. For instance, [Bloom \(2009\)](#) and [Basu and Bundick \(2017\)](#) analyze the impact of uncertainty using the VIX and VXO, which are two well-known investor fear gauges measuring the stock market’s expectation of volatility based on S&P500 index options. [Gilchrist, Sim, and Zakrajšek \(2014\)](#) construct realized volatility measures using a micro-level firm-specific asset returns dataset. Additionally, [Bontempi, Michele, Golinelli, and Squadrani \(2021\)](#) propose a web search-based uncertainty index. Finally, [Carriero et al. \(2016\)](#) develop indexes using a VAR model with stochastic volatility driven by common factors, [Jo and Sekkel \(2017\)](#) extract common factors using a stochastic volatility model, and [Carriero, Clark, and Marcellino \(2015\)](#) build a mixed frequency stochastic volatility model in order to estimate latent uncertainty indexes.

<sup>6</sup>One feature of our analysis is that we assess the usefulness of pre-classifying our large dataset into sub-groups, prior to treating said data using big data methods.

extracting latent factors from economic variables. Our paper is also closely related to [Jurado et al. \(2015\)](#), whose general methodology for constructing EUIs we utilize, although our analysis differs from theirs in a number of key ways. For instance, we explore the importance of grouping variables in our analysis, which allows us to construct “group” indexes rather than constructing an index for each variable in  $\mathbf{X}$ . Additionally, we explore the usefulness of our indexes for forecasting and we explore the usefulness of using windowing techniques (e.g. rolling estimation windows), in order to construct real-time forecasts. Finally, utilize multiple different machine learning methods when constructing our EUIs, thereby building on the growing body of research that employs machine learning methods for macroeconomic forecasting (see, e.g. [Kim and Swanson \(2014\)](#), [Kim, Kim, and Swanson \(2023\)](#), and the papers cited therein).

Our main empirical findings include the following. First, the majority of our models specified to include either EUIs or business conditions (BC) predictors have marginal predictive content for the macroeconomic variables that we examine, when compared with our benchmark econometric models, at forecast horizons ranging from 1-month to 1-year ahead. This is particularly true when forecasting “real” economic activity variables, defined to include output, housing, consumption, manufacturing, and employment. Second, the very best forecasting models usually contain BC predictors; and adding EUIs to BC predictors does yield forecast improvement, but primarily for housing variables. Third, the ADS index is found to be particularly useful for forecasting “real” variables. Fourth, our big data methods perform the best when used to predict variables at shorter forecast horizons. At longer horizons, big-data methods appear to become less useful, except when predicting housing variables. Finally, we show that using the so-called hybrid factor lasso method (see [Hansen and Liao \(2019\)](#)) to construct EUIS and BC predictors sometimes often leads to predictions that are more precise than those obtained using an AR(SIC) benchmark model.

The rest of this paper is organized as follows. In Section 2, we summarize the dimension reduction and shrinkage methods used in the construction of the BC predictors and EUIs that are proposed in this paper. In Section 3, we outline the experimental setup used in order to

examine the predictive content of our indexes. Section 4 contains a description of the data used in our empirical analysis, and Section 5 contains the results of our forecasting experiments. Section 6 concludes. Supplemental tables and figures are contained in a supplemental appendix available at “<https://econweb.rutgers.edu/nswanson/papers.htm>”.

## 2 Business Conditions and Economic Uncertainty Indexes

All of the business conditions predictors that are proposed in this paper are based on the application of principal components analysis (PCA) and the least absolute shrinkage operator (lasso). These business conditions predictors are directly used in our prediction experiments. In addition, these predictors are used in the specification of forecasting models, which are in turn used to construct predictions and prediction errors. Finally, the prediction errors are used to construct our EUIs. In the remainder of this section, we briefly summarize PCA and lasso methodology (Section 2.1), explain how to construct our various proposed BC predictors (Section 2.2), and summarize the methodology used to construct our EUIs (Section 2.3).

### 2.1 Dimension Reduction (PCA and the Lasso)

**Principal Component Analysis:** A brief overview of PCA is provided in this sub-section. For further details, refer to [Stock and Watson \(1999, 2002a, 2005, 2012\)](#), [Bai and Ng \(2002, 2008, 2009\)](#), [Kim and Swanson \(2014, 2018\)](#), and the references cited therein.

Principal component analysis is an unsupervised dimension reduction method, in which a reduced rank latent predictor,  $\mathbf{F}$ , is extracted from predictor matrix  $\mathbf{X}$ . PCA method is unsupervised because estimation involves the decomposition of the predictor matrix  $\mathbf{X}$ , and does not specify any relationship between  $\mathbf{X}$  and a target forecast variable of interest, say  $\mathbf{Y}$ . For dataset  $\mathbf{X} = \{X_{it}\}$ ,  $i = 1, \dots, N$ , and  $t = 1, \dots, T$ , assume that

$$\mathbf{X} = \mathbf{F}\mathbf{\Lambda} + \mathbf{E}, \tag{1}$$

where  $\mathbf{X}$  is a  $T \times N$  matrix of predictors,  $\mathbf{F}$  is a  $T \times r$  matrix of latent common factors that

we wish to extract,  $\mathbf{\Lambda}$  is an  $r \times N$  matrix of factor loadings, and  $\mathbf{E}$  is a  $T \times N$  disturbance term. For a given time period,  $t$ , and the  $N \times 1$  vector,  $X_t$ , the objective is to find an  $r \times 1$  ( $r < N$ ) vector,  $F_t$ , so that we can write:

$$\begin{aligned} X_1 &= \lambda_{11}F_1 + \dots + \lambda_{1r}F_r + e_1 \\ X_2 &= \lambda_{21}F_1 + \dots + \lambda_{2r}F_r + e_2 \\ &\vdots \\ X_N &= \lambda_{N1}F_1 + \dots + \lambda_{Nr}F_r + e_N \end{aligned} \tag{2}$$

Here,  $\lambda_{ij}$  is an element of the factor loading matrix, and  $e$  is an uncorrelated zero-mean disturbance term. A standard approach to estimate  $\mathbf{F}$  involves carrying out PCA, where  $\hat{\mathbf{F}} = \mathbf{X}\bar{\mathbf{\Lambda}}/N$ . And  $\bar{\mathbf{\Lambda}}$  is constructed as  $\sqrt{N}$  times the eigenvectors corresponding to the  $k$  largest eigenvalues of the  $N \times N$  matrix  $\mathbf{X}'\mathbf{X}$ .

**Lasso:** A key potential shortcoming characterizing the classical ridge estimator, which was developed as a way to solve the regression problem when  $N > T$ , is that no coefficient estimate shrinks identically to zero as  $T, N \rightarrow \infty$ . This lack of parsimony can be quite problematic when constructing forecasting models, particularly in the context of predicting macroeconomic aggregates, where it is well known that parsimonious models often yield superior predictive accuracy. The lasso, and in particular lasso regression is introduced in [Tibshirani \(1996\)](#), and is simply a form of restricted least squares, on the  $L_1$  norm (as opposed to the  $L_2$  norm as with ridge regression). The estimation problem is as follows:

$$\hat{\beta}_{lasso} = \arg \min_{\beta} \sum_{t=1}^n (y_{t+h} - \beta_0 - \sum_{j=1}^p x_{j,t}\beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j| \tag{3}$$

Here  $y_{t+h}$  is the target variable being predicted, the “ $\beta$ s” are parameters, the “ $x$ s” are predictors (i.e., elements of  $\mathbf{X}$ ), and the tuning parameter,  $\lambda$ , can be estimated using cross validation. This estimator supervised because a target variable is specified and explicitly appears in the least squares minimand.

## 2.2 Business Conditions (BC) Predictors

As discussed above, the BC predictors analyzed in this paper, and used in the construction of our EUIs, are latent factors extracted from  $\mathbf{X}$ . These predictors, denoted by  $\mathbf{F}_t$  and  $\mathbf{W}_t$ , in the sequel, are constructed using 10 different methods, as follows.

**1. Principal Components Analysis (PCA):** For each target variable  $y_j, j = 1, \dots, 14$ , we use PCA method to estimate  $\mathbf{F}_t$  from  $\mathbf{X}^-$ , where  $\mathbf{X}^-$  is defined to be all of the variables in  $\mathbf{X}$ , except for the target variable,  $y_j$ . The number of factors are determined by using the  $PC_{p2}$  criterion in Bai and Ng (2002). Additionally, the maximum number of the factors is set equal to eight, following the findings of McCracken and Ng (2016), who introduce and examine the dataset that we utilize in our experiments.

**2. Least Absolute Shrinkage Operator (LASSO):** This is the only procedure wherein we extract no latent factors, and instead use observable variables as our BC predictors. These observable predictors are selected anew for each target variable  $y_j$ , using predictor dataset  $\mathbf{X}^-$ . The tuning parameter,  $\lambda$ , in equation (3) is determined by applying  $h\nu$ -block cross validation, as discussed in Racine (2000). We set  $\lambda$  to be an element of 50 equally-spaced values between the smallest value required to retain all variables, and the largest value required to drop all but one variable.

In a number of the following methods, we combine the above PCA and lasso methods.

**3. Lasso-PCA (LPCA):** For each of our 14 target variables,  $y_j$ , we first carry out variable selection using lasso regression on the dataset  $\mathbf{X}^-$ , yielding a distinct subset of variables for each  $j = 1, \dots, 14$ . Then, PCA is applied to this subset in order to estimate  $\mathbf{F}_t$  for each  $y_j$ .

**4. All Predictors (AP):** For each target variable,  $y_j$ , we construct  $N - 1$  forecast models using each of the variables in  $\mathbf{X}^-$ . Namely, we fit regressions of the following form:

$$y_{jt+1} = \alpha + x_{it}\beta_i + \epsilon_{jt+1}, \quad (4)$$

where  $j \neq i, j = 1 \dots 14, i = 1, \dots, (N - 1)$ , and  $N = 122$ . Then, for each  $y_j$ , we choose top 20 predictors based on the examination of out-of-sample  $R^2$  statistics constructed from forecast



errors for real-time 1-step ahead forecasts made using models estimated using rolling windows of data, where the in-sample period,  $R = \frac{1}{2}T$ , and the out-of-sample-period  $P = \frac{1}{2}T$ .<sup>7</sup> PCA is then applied to the predictor set for  $y_j$ , yielding  $\mathbf{F}_t$ . This is done anew for each target variable being predicted.

**5. Group Prediction I (GP1):** Consider groups of variables, say  $X_k = x_1, \dots, x_{n^*}$ , where  $n^*$  defines the number of variables in a particular group, and  $k = 1, \dots, 8$ .<sup>8</sup> We forecast each variable in a given group, using all other variables in the same group. Namely, we fit regressions of the following form:

$$x_{it+1} = \alpha + X'_{kt}\beta_j + \epsilon_{it+1}, i = 1, \dots, n^*, \quad (5)$$

where the coefficients,  $\beta_j$ , and the stochastic disturbances,  $\epsilon_{it}$  are conformably defined. Then, in each group, we pick the 5 most precisely predicted variables based on the examination of out-of-sample  $R^2$  statistics constructed from forecast errors for real-time 1-step ahead forecasts made using models estimated using rolling windows of data, where the in-sample period,  $R = \frac{1}{2}T$ , and the out-of-sample-period  $P = \frac{1}{2}T$ . This is done for each group of variables, except for group 8, which only contains with 5 variables, of which 4 are selected, yielding a total of 39 predictor variable set across all 8 groups. Variable selection based on lasso regression is then applied using this predictor set for each  $y_j$ , yielding a unique vector of business conditions predictors,  $\mathbf{F}_t$ , for each target variable.<sup>9</sup> Finally, if the number of selected predictors for a given  $y_j$  is more than 8, we carry out PCA once again, but this time on the factors chosen in the previous steps of the procedure, yielding a final set of business conditions predictors,  $\mathbf{F}_t$ . Otherwise, we use the variables selected via lasso regression as our “final” set of business conditions predictors.

**6. Group Prediction II (GP2):** This method is a variation of GP1. For each group of variables and a given target variable,  $y_j$ , construct predictions using  $n^*$  different regressions

---

<sup>7</sup>Our choice of 20 predictors is based on our finding that lasso selects at most 20 variables in our experiments.

<sup>8</sup>The 8 groups into which we partition our predictor dataset are defined in Section 4.

<sup>9</sup>Note that if the target variable,  $y_j$ , is in the 39 variable predictor set, we remove it prior to carrying out lasso regression for each  $y_j$ .

of the form:

$$y_{jt+1} = \alpha + x_{it}\beta_i + \epsilon_{ijt+1}, i = 1, \dots, n^*, \quad (6)$$

where the  $x_{it}$  denotes an individual variable in a particular group. Using the approach outlined for GP1, pick the top 5 “forecast-best” predictors in each group, for each target variable. This yields a total of 40 predictors for each  $y_j$ .<sup>10</sup> Now, for each  $y_j$  predictor set, apply PCA, yielding a unique set of business conditions predictors,  $\mathbf{F}_t$ , for each target variable.

**7. Group Factors I (GF1):** First, carry out variable selection using lasso regression for each target forecast variable,  $y_j$ , on each of 8 groups of the variables in  $\mathbf{X}^-$ . This yields 8 new subsets of variables for each  $y_j$ . For each subset of variables, if the number of variables is larger than 8, then apply PCA in order to estimate “group latent factors”. Otherwise, just use selected variables themselves as group predictors. This yields 8 sets of group latent factors (i.e. group predictors), for each  $y_j$ . In order to achieve further parsimony, select a final set of predictors for each  $y_j$  by again carrying out variable selection using lasso regression on all latent factors across all 8 groups. If the number of selected group predictors is more than 8, carry out PCA once again, but this time on the factors chosen in the previous steps of this procedure, yielding a final set of BC predictors,  $\mathbf{F}_t$ . Otherwise, we use the factors selected via lasso regression as our “final” set of BC predictors.

**8. Group Factors II (GF2):** This is a simplified variant of GF1. First, carry variable selection using lasso regression for each target forecast variable,  $y_j$ , on each of 8 groups of variables in  $\mathbf{X}^-$ . This yields 8 new subsets of variables for each  $y_j$ . For each  $y_j$ , combine all of these variables into one group of variables, and apply PCA, yielding  $\mathbf{F}_t$ .

**9. PCA with Error Variables (EPCA):** As discussed in the introduction, this is a method where we address the possibility of inaccurate Lasso variable selection associated with highly correlated variables in  $\mathbf{X}^-$ , while at the same time allowing for both factors and disturbances in a factor model setup to contain useful information for forecasting a particular target variable

---

<sup>10</sup>An exception to this is the case of predicting S&P 500. For this variable, there are only 39 predictors, due to the fact that the group in which the S&P 500 is contained has only 5 variables in it, and S&P 500 cannot itself be used as a predictor in our setup.

of interest. This is done by adding additional variables to those estimated via application of PCA. The procedure follows the factor-lasso procedure outlined in [Hansen and Liao \(2019\)](#) and [Fan, Ke, and Wang \(2020\)](#). First, for each target variable,  $y_j$ , utilize PCA to estimate factors using the entire dataset,  $\mathbf{X}^-$ . Then, for each target variable, construct the predictor error term matrix,  $\mathbf{E} = \mathbf{X}^- - \widehat{\mathbf{X}}^-$ , where  $\widehat{\mathbf{X}}^- = \widehat{\mathbf{F}}\widehat{\mathbf{\Lambda}}$ ,  $\widehat{\mathbf{F}}$  is the estimated factor matrix, and  $\widehat{\mathbf{\Lambda}}$  is the factor loading matrix. Second, regress the target variable,  $y_j$  on  $\widehat{\mathbf{F}}$  and collect the residuals, say  $\widehat{\mathcal{E}}$ . Third, for each target variable,  $y_j$ , carry out variable selection using lasso regression, with dependent variable equal to  $\widehat{\mathcal{E}}$  and predictor variables  $\mathbf{E}$ . The selected new set of variables is called  $\mathbf{E}^*$ . Finally, form  $(\widehat{\mathbf{F}}, \mathbf{E}^*)$ , which is the set of business conditions predictors that we use when predicting  $y_j$ , and when constructing our EUIs.

**10. Jurado, Ludvigson and Ng Predictors (JLN):** When constructing the [Jurado et al. \(2015\)](#) uncertainty index there are essentially two steps involved. In the first step, factors for use in prediction equations are estimated. In the second step, uncertainty indexes are constructed using the forecast errors from predictive regressions in which these factors serve as predictors. In our analysis, we not only compare our EUIs with the EUI due to [Jurado et al. \(2015\)](#), which is called the JLNI in the sequel, but we also construct predictions of our 14 economic variables by directly using the business conditions predictor variables constructed using the same methodology as they use when constructing their EUI.<sup>11</sup>

More specifically, our method, called JLN, replicates the first step of the approach taken by [Jurado et al. \(2015\)](#) when constructing EUIs. In order to construct these JLN business conditions predictors, we first apply PCA to obtain our factors,  $\widehat{\mathbf{F}}_{\mathbf{t}}$ . Again following their approach, we then construct an additional predictor set,  $\mathbf{W}_{\mathbf{t}}$ . This predictor set contains the square of the first component of  $\widehat{\mathbf{F}}_{\mathbf{t}}$ , as well as the first factor extracted by applying PCA to  $X_{it}^2, i = 1, \dots, N$ . Finally, the consolidated dataset,  $(\widehat{\mathbf{F}}_{\mathbf{t}}, \mathbf{W}_{\mathbf{t}})$  is pruned by applying hard-thresholding using t-statistics on predictive regressions, for a given target variable,  $y_j$ . Four lags of the dependent variable are always included in the predictive regressions (as done in their paper).<sup>12</sup> Note that when applying PCA using their approach, the maximum number

---

<sup>11</sup>Note that our JLNI index is the “real economic uncertainty index” discussed on Sydney Ludvigson’s website, and our JLNI is downloaded directly from her website, as discussed above.

<sup>12</sup>See [Jurado et al. \(2015\)](#) for complete details.

of factors is 20 and the  $PC_{p2}$  of [Bai and Ng \(2002\)](#) is used to determine the number of factors.

In addition to using JLNI as a benchmark against which to assess predictions made using our EUIs and business conditions predictors, we also include another benchmark index in our analysis. Namely, we include the Aruoba-Diebold-Scotti (ADS) business conditions index, which is discussed in [Aruoba et al. \(2009\)](#).<sup>13</sup>

### 2.3 Economic Uncertainty Indexes (EUIs)

Our EUIs are constructed using forecast error variances from factor augmented autoregressions, where the factors are our BC predictors (as discussed in Section 2.2 above).<sup>14</sup> More specifically, our EUIs are aggregates of 14 individual uncertainty indexes that we construct. These 14 indexes correspond to the 14 target variables that we forecast, and our EUIs are in this sense “whole economy” uncertainty indexes. In total, 9 such uncertainty indexes are constructed. These EUIs use predictors from the first 9 different methods for specifying our BC predictors discussed above. These EUIs are named in similar fashion to our BC predictors above, except that an ‘I’ is added to the name, allowing us to easily differentiate between the cases where we add BC predictors to our forecasting models and cases where we add EUIs to our forecasting models. For the tenth method (i.e., the JLN method), we do not construct an EUI, but instead use the “real economic uncertainty index” from Sydney Ludvigson’s website, which we call JLNI is discussed above. In this sub-section, we summarize the method used to construct our nine original EUIs, noting that this method follows closely that of [Jurado et al. \(2015\)](#).

In the sequel,  $U_{jt}(h)$  denotes an uncertainty index for a given target variable to be predicted at forecast horizon,  $h$ , where  $j = 1, \dots, 14$ . As we are interested in forecasting 14 different target variables, we construct 14 different uncertainty indexes. Our “whole economy” uncertainty indexes are aggregated from these 14 measures, and are constructed as

---

<sup>13</sup>Our ADS data were downloaded directly from the Federal Reserve Bank of Philadelphia website. Also, note that ADS index is defined as a business conditions index and is not an uncertainty index, in the sense that it is not constructed using forecast errors.

<sup>14</sup>Note that in one case we do not use factor augmented autoregressions. Rather, we use an autoregression with additional predictor that are observable, as they are selected using the lasso. See discussion in the Section 2.2.

follows:

$$EUI = \sum_{j=1}^{14} w_j U_{jt}^y(h). \quad (7)$$

We assume that each weight,  $w_j$ , is equal to  $1/14$ .<sup>15,16</sup> In order to describe the construction of our uncertainty indexes in more detail, we draw in part from the discussion in [Jurado et al. \(2015\)](#), while highlighting key differences between their method and ours.

Let  $\mathbf{X}_t = (X_{1t}, X_{2t}, \dots, X_{Nt})'$  denote the predictors and assumes that  $X_{it}$  has a conformably defined factor structure taking the following form, for  $i = 1, \dots, N$ :

$$\mathbf{X}_{it} = \Lambda_i' \mathbf{F}_t + e_{it} \quad (8)$$

where  $\mathbf{F}_t$  is an  $r_F \times 1$  vector of latent common factors,  $\Lambda_i$  is a corresponding  $r_F \times 1$  vector of factor loadings, and  $e_{it}$  is a vector of idiosyncratic errors. Now, let  $y_{jt}$  denote a target variable of interest, to be fitted using the following factor augmented forecasting model. This is the model from which our forecast errors are collected:

$$y_{jt} = \phi_j(L)y_{jt-1} + \gamma_j(L)\hat{\mathbf{F}}_{t-1} + \beta_j(L)\mathbf{W}_{t-1} + v_{jt}, \quad (9)$$

where  $\phi_j(L), \gamma_j(L), \beta_j(L)$  are finite-order polynomials in the lag operator  $L$  of order  $p_y, p_f, p_w$ , respectively,  $\hat{\mathbf{F}}_{t-1}$  are consistent estimates of  $\mathbf{F}_{t-1}$ , and  $\mathbf{W}_{t-1}$  contains additional predictors, to be discussed below. When the factors have autoregressive dynamics, we can write the above equation in more compact form. Let  $\mathbf{Z}_t = (\hat{\mathbf{F}}_t, \mathbf{W}_t)'$  be an  $r = r_F + r_W$  dimensional vector and let  $\mathbf{Z}_t = (\mathbf{Z}_t', \dots, \mathbf{Z}_{t-q+1}')'$ . Also, let  $Y_{jt} = (y_{jt}, y_{jt-1}, \dots, y_{jt-q+1})'$ . We can now

---

<sup>15</sup>Different weighting schemes, where weights are selected according to the “importance” of the different target variables associated with each  $U_{jt}(h)$  may be of interest, although such schemes are not examined here. For example, note that each of the 14 variables predicted in this paper are taken from a specific member of 8 “groups” of variables, into which our set of predictors are partitioned. The correlation between business conditions indexes constructed for each of these groups could be used to determine an interesting weighting scheme, whereby  $U_{jt}(h)$  variables associated with the most relevant business conditions indexes for a particular target variable are assigned a greater weight in a “data driven” weighting scheme.

<sup>16</sup>A key difference between our approach and that taken in [Jurado et al. \(2015\)](#) is that we construct only 14 individual uncertainty indexes, based on groups of variables, while they aggregate a large number of individual uncertainty indexes, one for each variable in the dataset analyzed in their paper. In order to compare the usefulness of the global uncertainty index examined in their paper, relative to the indexes proposed in this paper, we also estimate prediction models using the original EUI index, as discussed above.

write the forecasting equation for factors and target variables as follows:

$$\begin{pmatrix} \mathcal{Z}_t \\ Y_{jt} \end{pmatrix}_{\substack{rq \times 1 \\ q \times 1}} = \begin{pmatrix} \Phi^{\mathcal{Z}} & 0 \\ \Lambda'_j & \Phi_j^Y \end{pmatrix}_{\substack{qr \times qr & qr \times q \\ q \times qr & q \times q}} \begin{pmatrix} \mathcal{Z}_{t-1} \\ Y_{jt-1} \end{pmatrix} + \begin{pmatrix} \mathcal{V}_t^{\mathcal{Z}} \\ \mathcal{V}_{jt}^Y \end{pmatrix} \quad (10)$$

$$\mathcal{Y}_{jt} = \Phi_{jt}^{\mathcal{Y}} \mathcal{Y}_{jt-1} + \mathcal{V}_{jt}^{\mathcal{Y}},$$

where  $\Lambda'_j$  and  $\Phi_j^Y$  are functions of the coefficient in the lag polynomials in (6), and  $\Phi^{\mathcal{Z}}$  stacks the autoregressive coefficients of the components of  $\mathcal{Z}_t$ . We denote the individual uncertainty index of each target variable as  $U_{jt}^y(h)$ . With  $1_j$  being a selection vector, so that:

$$U_{jt}^y(h) = \sqrt{1_j' \Omega_{jt}^{\mathcal{Y}}(h) 1_j}. \quad (11)$$

Note that when  $h = 1$ ,

$$\Omega_{jt}^{\mathcal{Y}}(1) = E_t(\mathcal{V}_{jt+1}^{\mathcal{Y}} \mathcal{V}_{jt+1}^{\mathcal{Y}'}) . \quad (12)$$

For  $h > 1$ , the forecast variance of  $\mathcal{Y}_{jt+h}$  is generated according to:

$$\Omega_{jt}^{\mathcal{Y}}(h) = \Phi_{jt}^{\mathcal{Y}} \Omega_{jt}^{\mathcal{Y}}(h-1) \Phi_{jt}^{\mathcal{Y}'} + E_t(\mathcal{V}_{jt+h}^{\mathcal{Y}} \mathcal{V}_{jt+h}^{\mathcal{Y}'}) \quad (13)$$

In our experiments, we update the coefficient matrix  $\Phi_{jt}^{\mathcal{Y}}$  at each point in time, using a rolling window, prior to constructing forecasts. This setup differs from the approach taken in [Jurado et al. \(2015\)](#), where a fixed in-sample coefficient matrix is estimated using their entire data sample. Note that [Jurado et al. \(2015\)](#) do not construct forecasts using their EUIs, while in this paper our primary use of EUIs indexes is in the construction of forecasts.

In our approach as well as in the [Jurado et al. \(2015\)](#) approach, one-step ahead forecast errors associated with  $y_{jt}$ ,  $F_{l,t}$ , and  $W_{g,t}$  have time varying volatility,  $\sigma_{jt}^y$ ,  $\sigma_{lt}^F$ , and  $\sigma_{gt}^W$ , respectively. Namely, for factor  $F_t$  (and analogously for  $W_t$ ), we assume that:

$$\mathbf{F}_t = \Phi^F \mathbf{F}_{t-1} + \mathbf{v}_t^F. \quad (14)$$

This allows the shock to  $\mathbf{F}$  to exhibit time-varying stochastic volatility, such that  $v_t^F = \sigma_t^F \epsilon_t^F$ , and  $\epsilon_t^F \stackrel{iid}{\sim} N(0, 1)$ , where the log volatility has the following autoregressive structure:

$$\log(\sigma_t^F)^2 = \alpha^F + \beta^F \log(\sigma_{t-1}^F)^2 + \tau^F \eta_t^F, \eta_t^F \stackrel{iid}{\sim} N(0, 1). \quad (15)$$

Using the law of the iterated logarithm, and then taking the expectation of the above equation, yields:

$$E_t(\sigma_{t+h}^F)^2 = \exp \left[ \alpha^F \sum_{s=0}^{h-1} (\beta^F)^s + \frac{(\tau^F)^2}{2} \sum_{s=0}^{h-1} (\beta^F)^{2s} + (\beta^F)^h \log(\sigma_t^F)^2 \right] \quad (16)$$

The stochastic volatility parameters,  $\alpha$ ,  $\beta$ , and  $\tau$ , are estimated from the least square residuals of the forecasting model using Markov chain Monte Carlo methods, as detailed in footnote 9 of [Jurado et al. \(2015\)](#). Since  $\epsilon_t^F \stackrel{iid}{\sim} N(0, 1)$ , we have that  $E_t(v_{t+h}^F)^2 = E_t(\sigma_{t+h}^F)^2$ . This allows us to compute the  $h > 1$  forecast error variance for  $F$  using the recursion

$$\Omega_t^F(h) = (\Phi^F) \Omega_t^F(h-1) \Phi^{F'} + E_t(v_{t+h}^F v_{t+h}^{F'}), \quad (17)$$

with  $\Omega_t^F(1) = E_t(v_{t+1}^F)^2$ . The  $h$  period ahead predictor uncertainty at time  $t$  is the square root of the  $h$  - step forecast error variance of the predictor

$$U_t^F(h) = \sqrt{1_F' \Omega_t^F(h) 1_F}, \quad (18)$$

where  $1_F$  is an appropriate selection vector.

In most of our experiments, we use a simplified framework where only business conditions predictors constructed using PCA and the lasso are utilized. Throughout much of this paper, these BC predictors are defined to be the latent factors extracted from  $\mathbf{X}$ , and are called  $\mathbf{F}_t$ .<sup>17</sup> In this case, we have that:

$$y_{jt+1} = \phi_j^y y_{jt} + \gamma_j^F \hat{\mathbf{F}}_t + v_{jt+1}^y. \quad (19)$$

---

<sup>17</sup>The  $\mathbf{W}_t$  are also latent factors, but are constructed using the methods of [Jurado et al. \(2015\)](#), as discussed in the next section.

Here, the shock to  $y_j$  still has time-varying stochastic volatility, with  $v_{jt+1}^y = \sigma_{jt+1}^y \epsilon_{jt+1}^y$  and  $\epsilon_{jt+1}^y \stackrel{iid}{\sim} N(0, 1)$ . Moreover, log volatility follows the following process:

$$\log(\sigma_{jt+1}^y)^2 = \alpha_j^y + \beta_j^y \log(\sigma_{jt}^y)^2 + \tau_j^y \eta_{jt+1}, \eta_{jt+1} \stackrel{iid}{\sim} N(0, 1). \quad (20)$$

Again, the stochastic volatility parameters,  $\alpha_j, \beta_j$ , and  $\tau_j$ , are estimated from the least squares residuals of the forecasting model using Markov chain Monte Carlo methods.

The following decomposition illustrates how uncertainty in the predictors affects uncertainty in target variable  $y_j$ . For  $h = 1$ ,  $V_{jt+1}^y$  is only related to  $v_{jt+1}^y$ . When  $h = 2$ ,

$$V_{jt+2}^y = v_{jt+2}^y + \phi_j^y V_{jt+1}^y + \gamma_j^F V_{t+1}^F, \quad (21)$$

where  $V_{t+1}^y$  and  $V_{t+1}^F$  are uncorrelated. When  $h = 3$ , the forecast error is:

$$V_{jt+3}^y = v_{jt+3}^y + \phi_j^y V_{jt+2}^y + \gamma_j^F V_{t+2}^F \quad (22)$$

where  $V_{t+2}^y$  and  $V_{t+2}^F$  are correlated now because they both depend on  $V_{t+1}^F$ .

Recalling the general case where the predictors are  $\mathbf{Z}_t = (\mathbf{F}_t', \mathbf{W}_t')'$  and its lags, the  $h$ -step ahead forecast error variance for  $Y_{jt+h}$  is given by:

$$\Omega_{jt}^Y(h) = \Phi_j^Y \Omega_{jt}^Y(h-1) \Phi_j^{Y'} + \Omega_{jt}^Z(h-1) + E_t(\mathcal{V}_{jt+h}^Y \mathcal{V}_{jt+h}^{Y'}) + 2\Phi_j^Y \Omega_{jt}^{YZ}(h-1),$$

where  $\Omega_{jt}^{YZ}(h) = \text{cov}_t(\mathcal{V}_{jt+h}^Y, \mathcal{V}_{jt+h}^Z)$ . Here, the terms in  $E_t(\mathcal{V}_{jt+h}^Y \mathcal{V}_{jt+h}^{Y'})$  are computed using the fact that  $E_t(v_{jt+h}^y)^2 = E_t(\sigma_{jt+h}^y)^2$ ,  $E_t(v_{t+h}^F)^2 = E_t(\sigma_{t+h}^F)^2$ , and  $E_t(v_{t+h}^W)^2 = E_t(\sigma_{t+h}^W)^2$ . Thus, this equation is equivalent to equation (10), for the subvector  $\mathbf{Y}_t$ .

Notice that in the above discussion, our EUIs depend both on  $\mathbf{F}_t$  and  $\mathbf{W}_t$ . These variables are latent factors extracted from our dataset  $\mathbf{X}$ . Moreover, these latent factors are interpreted in this paper as business conditions predictors, as discussed above. Thus, by specifying the various procedures that we utilize to construct our BC predictors, we are also providing the



various inputs for the construction of our EUIs.<sup>18</sup>

In the next section, we outline the experiments that we carry out in order to assess the marginal predictive content of our BC predictors and EUIs.

### 3 Forecasting Models and Predictive Accuracy Testing

In order to examine the marginal predictive content of our BC predictors and EUIs, we carry out real-time  $h$ -month ahead predictions using an updated 122 variables of the monthly dataset examined in Jurado et al. (2015), as discussed in Section 4, and using a number of forecasting models, for  $h = 1, 3$ , and 12. Results for  $h = 6$  yield the same set of qualitative results, and so are omitted for the sake of brevity, but are available upon request from the authors. Overall, predictions are made for 14 different monthly U.S. variables that are divided into 8 groups (see Section 4 for details). In our experiments, we construct predictions using the following models:

**Autoregressive Model (AR(SIC)):** The AR model is our benchmark and is specified as follows:

$$y_{t+h} = \alpha + \beta'_h(L)y_t + \epsilon_{t+h}, \quad (23)$$

where  $y_t$  is the scalar target forecast variable,  $h$  denotes the forecast horizon,  $\hat{\beta}_h(L)$  is a finite order lag polynomial,  $\epsilon_{t+h}$  is a stochastic disturbance term. In our experiments, the lag order is selected using SIC. In all models, estimation is carried out using least squares.

**AR model with business conditions predictors (AR(SIC)+BC):** The AR model augmented with BC predictors is:

$$y_{t+h} = \alpha + \beta'_h(L)y_t + \gamma'_h(L)\widehat{BC}_t + \epsilon_{t+h}, \quad (24)$$

where  $\beta'_h(L)$  and  $\gamma'_h(L)$  are finite order lag polynomials,  $\widehat{BC}_t$  is a set of latent factors, observable variables, or latent factors and observable variables, constructed as discussed above.

**AR model with economic uncertainty indexes (AR(SIC)+EUI):** The AR model

---

<sup>18</sup>As mentioned above, in all but one of the procedures outlined above, we specify and estimate only  $\mathbf{F}_t$ .

augmented with EUIs is:

$$y_{t+h} = \alpha + \beta'_h(L)y_t + \gamma'_h(L)\widehat{\mathbf{EUI}}_t + \epsilon_{t+h}, \quad (25)$$

where  $\beta'_h(L)$  and  $\gamma'_h(L)$  are finite order lag polynomials,  $\widehat{\mathbf{EUI}}_t$  is a set of uncertainty indexes, constructed as discussed above.

**AR model with business conditions predictors and economic uncertainty indexes**

**(AR(SIC)+BC+EUI):** The AR model augmented with BC predictors and EUIs is:

$$y_{t+h} = \alpha + \beta'_h(L)y_t + \gamma'_{1h}(L)\widehat{\mathbf{BC}}_t + \gamma'_{2h}(L)\widehat{\mathbf{EUI}}_t + \epsilon_{t+h}, \quad (26)$$

where  $\beta'_h(L)$ ,  $\gamma'_{1h}(L)$ , and  $\gamma'_{2h}(L)$  are finite order lag polynomials, and  $(\widehat{\mathbf{BC}}_t, \widehat{\mathbf{EUI}}_t)$  and all factors contained in  $\widehat{\mathbf{BC}}_t$  and used in the estimation of  $\widehat{\mathbf{EUI}}_t$  are consistently estimated and, as discussed above.

As mentioned above, our dataset consists of 122 monthly variables for the period 1963:3 to 2021:6, so that  $N = 122$  and  $T = 700$ , using our above notation. Our training sample experiments, used when calibrating the parameters in the methods discussed in Sections 2.2 and 2.3 utilize data prior to 1992:6. This leaves the sample period 1992:6-2021:6 (349 observations) for our reported prediction experiments. Note that in the results reported in this paper, our predictions are carried out for a smaller period, ending in 2019:12. For an analogous set of results for the longer period ending in 2021:6 (which includes effects from Covid-19), refer to the supplemental appendix. As evidenced by comparison of these sets of results, predictions from all of our methods are affected by including Covid-19 data, as are predictions based on the use of our benchmark indices (i.e., ADS, JLNI, and the VIX). For this reason, we report results only for the shorter out-of-sample period, and we leave further examination of the importance of Covid-19 data to further research. We set the initial in-sample estimation period to be  $R = 144$  observations (i.e., for  $h = 1$  this implies that the in-sample estimation period is 1992:6 - 2004:5). Thus, our longest prediction period for  $h = 1$  is 2004:6-2021:6 (i.e.,  $P = 349 - 144 = 205$  observations). All BC predictors and EUIs are re-

estimated in real-time, using a rolling window estimation scheme, as are the above prediction models, prior to the construction of each calendar dated prediction. Moreover, the in-sample estimation periods used when constructing our  $h = 3$  and 12-step ahead forecasts are adjusted so that our forecast period remains 2004:6-2021:6, regardless of forecast horizon.

Forecasting performance is evaluated using point mean squared forecast errors (MSFEs), where  $\text{MSFE} = \frac{1}{P} \sum_{t=1}^T (y_{j,t} - \hat{y}_{j,t})^2$ , and  $\hat{y}_{j,t}$  denotes a prediction for target variable  $y_j$ , as discussed above. In our tabulated results, MSFEs, relative to that of our benchmark AR model are reported. Additionally, we report the results of Giacomini and White (GW) tests (see [Giacomini and White \(2006\)](#)), which can be viewed as conditional Diebold-Mariano (DM) predictive accuracy tests (see [Diebold and Mariano \(2002\)](#)). Recall that the null hypothesis of the DM test when formulated using the conditioning approach of Giacomini and White is:  $H_0 : E[L(\hat{\epsilon}_{t+h}^{(1)})|G_t] - E[L(\hat{\epsilon}_{t+h}^{(2)})|G_t] = 0$ , where the  $\hat{\epsilon}_{t+h}^{(i)}$  are prediction errors associated with model  $i$ , for  $i = 1, 2$ , and  $G_t$  denotes the conditioning set, which includes the model and estimated parameters. Here,  $L(\cdot)$  is a quadratic loss function, and the test statistic is:

$$\text{DM}_P = P^{-1} \sum_{t=1}^P \frac{d_{t+h}}{\hat{\sigma}_{\bar{d}}}, \quad (27)$$

where  $d_{t+h} = [\hat{\epsilon}_{t+h}^{(1)}]^2 - [\hat{\epsilon}_{t+h}^{(2)}]^2$ ,  $\bar{d}$  denotes the mean of  $d_{t+h}$ ,  $\hat{\sigma}_{\bar{d}}$  is a heteroskedasticity and autocorrelation consistent estimate of the standard deviation of  $\bar{d}$ , and  $P$  denotes the number of ex-ante predictions used to construct the test statistic.<sup>19</sup> If the statistic is significantly negative, then Model 1 is preferred to Model 2. For this test, we assume that the test statistic is asymptotically normally distributed, following [Giacomini and White \(2006\)](#). For a discussion of alternative approaches to assessing forecasting performance, see [Corradi and Swanson \(2006\)](#) and [Rossi and Sekhposyan \(2011\)](#).

---

<sup>19</sup>In this paper, we report test results for the Wald version of this test statistic (see [Giacomini and White \(2006\)](#) for further details).

## 4 Data

Our dataset consists of monthly macroeconomic variables obtained from the FRED-MD database at the St. Louis Federal Reserve Bank. In our forecasting experiments, we follow the data cleaning methods outlined in [McCracken and Ng \(2016\)](#) and on the FRED-MD data website. This yields a dataset consisting of 122 variables for the period 1963:3 to 2021:6. The full list of all macroeconomic variables and their transformations is available upon request from the authors.<sup>20</sup> We also collected VIX data from the FRED database, the Aruoba, Diebold, and Scotti (ADS) business conditions index from the website of Federal Reserve Bank of Philadelphia, and Jurado, Ludvigson and Ng economic uncertainty index data from the website of Sydney Ludvigson (see above discussion).

The 8 groups of variables into which our 122 variable dataset (called  $\mathbf{X}$  above) is subdivided include: (i) output and income, (ii) labor market, (iii) housing, (iv) consumption and inventories, (v) money and credit, (vi) interest and exchange rates, (vii) prices, and (viii) stock market. Additionally, the 14 variables for which we construct predictions are: Real Personal Income (RPI), Industrial Production (INDPRO), Civilian Unemployment Rate (UNRATE), Initial Jobless Claims (CLAIMS), Housing Starts: new, privately owned (HOUST), Housing Permits: new, privately owned (PERMIT), Real Personal Consumption Expenditures (RCON), Real Manufacturing and Trade Industries Sales (MTS), M2 Money Stock (M2), 10-Year Government Treasury Bond Rate (R10), PPI - Finished Goods (PPI), CPI - All Items (CPI), Personal Consumption Expenditures - Chain Index (PCECI), S&P Common Stock Price Index - Composite (S&P500).

---

<sup>20</sup>We use fully revised FRED data in our subsequent empirical analysis. In this sense, our dataset is pseudo real-time, and not truly real-time. Our findings, thus, are meant primarily to be a guide to understanding the potential usefulness of the various new BC predictors and EUIs that we specify and assess in our prediction experiments. For further discussion of pseudo real-time forecasting, refer to [Girardi, Golinelli, and Pappalardo \(2017\)](#).

## 5 Empirical Findings

In this section, we discuss the results of a series of forecasting experiments designed to assess the marginal predictive content of our business conditions (BC) predictors and economic uncertainty indexes (EUIs), in the context of predicting 14 target variables chosen to be representative of 8 different groups into which the FRED-MD dataset is subdivided (see the preceding section as well as Table 1 for details). In particular, and as discussed in Section 2, we first construct 10 sets of BC predictors, which are latent factors extracted using PCA, or variables selected using lasso regression, or latent factors constructed using factor-lasso methods. The 10 methods are denoted PCA, LASSO, LPCA, AP, GP1, GP2, GF1, GF2, EPCA, and JLN, as discussed in Section 2.2. BC predictors are then used in the construction of 10 global economic uncertainty indexes (EUIs). These 10 indices are denoted PCAI, LASSOI, LPCAI, API, GP1I, GP2I, GF1I, GF2I, EPCAI, and JLN, as discussed in Section 2.3. We also report results using the ADS index, which is extracted following the approach of [Aruoba et al. \(2009\)](#). Finally, we include the VIX as a benchmark EUI in our analysis.

In summary, 3 samples of  $h=1, 3$  and 12-step ahead predictions are constructed for the out-of-sample period 2004:6-2021:6, for each target variable. However, we only report results for the period 2004:6-2019:12 in the sequel. The reason for this is that including results that span the Covid-19 period markedly impact the ranking of our models, due to the fact that various methods yield a number of non-sense predictions. Findings based on the entire sample period that includes the effects of Covid-19 are included in the supplemental appendix. Predictions are made using forecasting models denoted as: AR(SIC), AR(SIC)+BC, AR(SIC)+EUI, and AR(SIC)+BC+EUI, as summarized in Table 2, and detailed in Section 3. Please refer to Section 3 for complete details of our experimental setup. The remainder of this section summarizes a number of key empirical findings, based on the results collected in Tables 1-10 and Figure 1. Additional experimental results that we discuss, but that are not reported here, are contained in the supplemental appendix.

First, we discuss results based on the MSFEs reported in Table 3. Entries in this table are relative MSFEs, when comparing forecasts based on our AR(SIC)+EUI model with those

based on our AR(SIC) model. Values less than unity indicate that the point MSFE associated with our EUI-augmented models are lower than those associated with the AR(SIC) benchmark. For each variable, 11 “methods” are used to construct our EUIs (i.e., the 10 methods discussed in Section 2.2 and the VIX), as listed in the first row of entries in the table. Bolded entries indicate our “MSFE-best” methods, for a given target variable and forecast horizon. For example, inspection of the results in this table indicate that constructing BC predictors using the Lasso yields EUIs that are included as predictors in our MSFE-best models for 9 of all variables and forecast horizon combinations. However, this finding is based solely on the examination of the point MSFEs reported in the table. As these MSFEs are relative, entries greater than unity indicate that the AR(SIC) model has a lower MSFE than the “best” of our 11 reported methods. In this sense, LASSOI never “wins”. Still, it is worth pointing out that the LASSOI method does yield an EUI that significantly improves predictions relative to the AR(SIC) benchmark in 5 cases, across all variable and forecast horizon permutations, at a 5% level of significance (based on the application of [Giacomini and White \(2006\)](#) conditional predictive ability tests). Nevertheless, other methods yield lower MSFEs in all 5 of these cases, as indicated by the fact that these 5 LASSOI entries are not bolded. Indeed, there are many cases where our EUIs do yield predictions that outperform the AR(SIC) model. EUIs constructed using the GP1I method yield MSFEs that are “MSFE-best” across all methods and beat the AR(SIC) benchmark in 6 cases. Moreover, other methods also “win” in various instances. For example, GF2I “wins” twice, JLNI “wins” 6 times, and VIX “wins” once.

Second, when comparing all of the bolded entries in Table 3, note that the AR(SIC) benchmark “wins” (since the best MSFEs are greater than unity) in 20 of 42 case, or around 1/2 of the cases, across all variables and forecast horizon combinations.<sup>21</sup> Interestingly, though, the AR(SIC) benchmark yields the lowest overall MSFE in only 3 of 24 cases, or around 1/8 of the cases, when examining results for our real economic activity variables, which include the first 8 variables in the table. For the remaining 6 variables, which are price, interest rate, and financial variables, the AR(SIC) model “wins” in 17 of 18 cases, indicating that our EUIs are

---

<sup>21</sup>Refer to Table 7 for further details, where the top 3 models are listed for each variable and forecast horizon from Table 3.

not directly useful for forecasting in these cases. Thus, we have strong evidence that EUIs are useful for predicting real economic activity type variables, and the particular EUI that is useful varies by variable and forecast horizon, although the “best” EUIs appear to be GP1I and JLNI.

Third, recall that the results in Table 3 are based on the comparison of AR(SIC)+EUI and AR(SIC) type forecasting models. In Table 4, AR(SIC)+BC and AR(SIC) forecasting models are instead compared, and in Table 5, AR(SIC)+BC+EUI and AR(SIC) forecasting models are compared. Turning first to the results collected in Table 4, note that the “MSFE-best” models with BC predictors yield lower MSFEs than the AR(SIC) benchmark in 23 of 24 cases when forecasting our real economic activity variables, and in 6 of 18 cases when forecasting our price, interest rate, and financial variables. Moreover, MSFEs associated with our best models are often significantly lower than those associated with using the AR(SIC) model. Thus, both EUIs and BC predictors appear useful for forecasting our “real” variables while only BC predictors yield useful predictive content when forecasting our “financial” variables.

Fourth, Table 8 summarizes the 3 top model/method combinations for each variable and forecast horizon reported on in Table 4. Note that ADS performs very well, winning in 11 of 24 cases for “real” variables. Other methods also do well, accounting for the other 13 “wins”, with LASSO accounting for 5 of these.

Fifth, recall that Table 5 contains results for forecasting models constructed using both EUIs and BC predictors. As previously, our big data methods yield good results when forecasting “real” variables, winning in 19 of 24 cases.<sup>22</sup> Of course, this results may be simply due to the inclusion of our BC predictors in these models. To assess whether this is the case, the overall “MSFE-best” models across all models, methods, variables, and forecast horizons are summarized in Table 6. Not surprisingly, for our real variables, BC predictor type models win in 16 of 24 cases, model incorporating EUIs win in 3 cases, and models incorporating both win in 4 cases. Finally, AR(SIC) models win in only 1 cases. When comparing results for the “financial” variables, the AR(SIC) benchmark wins in 11 of 18 cases, BC predictors

---

<sup>22</sup>For simplicity, we define “big data methods” to include the use of the ADS index, although the current variant of this index reported on the Philadelphia Federal Reserve website does not use the high dimensional dataset that is used when constructing the rest of the EUIs and BC predictor variables analyzed in this paper.

deliver 6 “wins” and EUIs deliver the remaining one.<sup>23</sup>

Sixth, it is worth noting that our two housing variables (HOUST and PERMIT) account for 3 of the 6 cases where forecasting models that include both EUIs and BC predictors. Thus, while BC predictors clearly do well in our experiments, there are select cases where using EUIs and BC predictors is useful.

Seventh, in order to assess the usefulness of model averaging and combination in our experiments, we collect results reporting the performance of six such methods in Table 10. In this table, FAV1 refers to the case where all 9 of the EUIs (excluding JLNI and VIX) discussed in Section 2.3 are averaged. FAV2 averages FAV1, JLNI, and VIX. FAV3 averages FAV1, JLNI, VIX, and ADS. Comb1 refers to the case where all 9 forecasts constructed using our 9 EUIs are averaged. Comb2 averages these 9 forecasts as well as forecasts made using the JLNI and VIX indices. Finally, Comb3 adds ADS to the set of forecasts that are combined. Results in this table are presented in similar fashion to those in Tables 3-5. However, in this table, entries are only bolded if they yield a MSFE that is lower than the best MSFE associated with our overall “winning” models reported in Table 6. Interestingly, there are only three instances in 42 cases (compare all variable and forecast horizon permutations) where three of these methods dominates our previously reported “best” model/method combinations for one target variable. In this sense, we have evidence that the degree of aggregation already inherent to our methods is sufficient for forecasting.

Finally, we graphically depict various EUIs and HOUST in Figures 1. Note that the indexes constructed for use in our 1, 3, and 12-step ahead forecasting models largely follow our housing variables, even during the 2008 financial crisis. However, note that various EUIs that performs very poorly when  $h = 12$  are not reported.

Summarizing, we find that both EUIs and BCIs are quite useful for forecasting a variety of macroeconomic variables at forecast horizons ranging from 1-month to 1-year ahead, with

---

<sup>23</sup>Additional tabulated results in which the numerator used in the construction of the entries reported in Tables 3-5 is not the AR(SIC), but is instead a competitor big-data model are contained in the supplemental appendix. For example, Table 1 in the supplemental appendix reports relative MSFEs where the numerator is the MSFE of our AR(SIC)+BC+EUI model and the denominator is the MSFE of our AR(SIC)+BC Model. As expected, results in this table indicate that MSFEs associated with our AR(SIC)+BC models are usually lower than those associated with the bigger model.



the caveat that BC predictors are the most useful, and that forecasting “real” variables using big data leads to more model “wins” than when forecasting “financial” variables. In this sense, we have new evidence of the usefulness of the types of indexes discussed in [Aruoba et al. \(2009\)](#) and [Jurado et al. \(2015\)](#), and expanded on in this paper.

## 6 Concluding Remarks

In this paper, we propose a number of new economic uncertainty indexes (EUIs) and business conditions (BC) predictors. Our BC predictors are constructed using variations on extant lasso regression and principal components analysis methods designed for variable selection and latent factor estimation, as well as novel factor-lasso methods that account for covariation across predictor variables in high dimensional datasets. Our EUIs are constructed using BC predictors as inputs and by employing methodology related to that discussed in [Jurado et al. \(2015\)](#). In a series of forecasting experiments in which we predict 14 macroeconomic variables, we show that our EUIs and BC predictors contain marginal predictive content. A number of outstanding issues are left to future research. It remains, for example, to carry out further empirical analysis that builds on our experiments by analyzing real-time datasets. It also remains to examine various variants of the variable “grouping” methods discussed in this paper. It also remains to carry our empirical and theoretical analysis of high-dimensional and mixed frequency models designed to estimate the types of EUIs and BC predictors discussed in this paper.

## References

- Andreou, E., Ghysels, E., & Kourtellis, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, 31(2), 240–251.
- Aruoba, S. B., Diebold, F. X., & Scotti, C. (2009). Real-time measurement of business conditions. *Journal of Business & Economic Statistics*, 27(4), 417–427.
- Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1), 191–221.
- Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, 146(2), 304–317.
- Bai, J., & Ng, S. (2009). Boosting diffusion indices. *Journal of Applied Econometrics*, 24(4), 607–629.
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593–1636.
- Basu, S., & Bundick, B. (2017). Uncertainty shocks in a model of effective demand. *Econometrica*, 85(3), 937–958.
- Bloom, N. (2009). The impact of uncertainty shocks. *econometrica*, 77(3), 623–685.
- Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., & Terry, S. J. (2018). Really uncertain business cycles. *Econometrica*, 86(3), 1031–1065.
- Bontempi, M. E., Michele, F., Golinelli, R., & Squadrani, M. (2021). Eurq: A new web search-based uncertainty index. *Economica*, 88(1), 969–1015.
- Bulligan, G., Marcellino, M., & Venditti, F. (2015). Forecasting economic activity with targeted predictors. *International Journal of Forecasting*, 31(1), 188–206.
- Carriero, A., Clark, T. E., & Marcellino, M. (2015). Realtime nowcasting with a bayesian mixed frequency model with stochastic volatility. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(4), 837–862.
- Carriero, A., Clark, T. E., & Marcellino, M. (2016). Measuring uncertainty and its impact on the economy. *Review of Economics and Statistics*, 100(5), 799–815.

- Cepni, O., Guney, E. I., & 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.
- Chauvet, M., Senyuz, Z., & Yoldas, E. (2015). What does financial volatility tell us about macroeconomic fluctuations? *Journal of Economic Dynamics and Control*, 52, 340–360.
- Cheng, M., Swanson, N. R., & Yang, X. (2021). Forecasting volatility using double shrinkage methods. *Journal of Empirical Finance*, 62(1), 46–61.
- Corradi, V., & Swanson, N. R. (2006). Predictive density evaluation. In G. Elliott & A. Timmermann (Eds.), *Handbook of forecasting* (pp. 197–284). Elsevier.
- Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1), 134–144.
- Fan, J., Ke, Y., & Wang, K. (2020). Factor-adjusted regularized model selection. *Journal of Econometrics*, 216(1), 71–85.
- Franses, C., Marcellino, M., Mazzi, G. L., & Proietti, T. (2008). A monthly indicator of the euro area gdp. *European University Institute; Series/Number: EUI ECO; 2008/32*.
- Ghysels, E., Sinko, A., & Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric Reviews*, 26(1), 53–90.
- Giacomini, R., & White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74(6), 1545–1578.
- Gilchrist, S., Sim, J. W., & Zakrajšek, E. (2014). Uncertainty, financial frictions, and investment dynamics. *NBER Working Papers 20038*.
- Girardi, A., Golinelli, R., & Pappalardo, C. (2017). The role of indicator selection in nowcasting euro-area gdp in pseudo real time. *Empirical Economics*, 53(1), 79–99.
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223–2273.
- Hansen, C., & Liao, Y. (2019). The factor-lasso and k-step bootstrap approach for inference in high-dimensional economic applications. *Econometric Theory*, 35(3), 465–509.

- Jo, S., & Sekkel, R. (2017). Macroeconomic uncertainty through the lens of professional forecasters. *Journal of Business & Economic Statistics*, 1–11.
- Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3), 1177–1216.
- Kim, H.-H., Kim, K., & Swanson, N. R. (2023). Mixing mixed frequency and diffusion indices in good times and in bad: an assessment based on historical data around the great recession of 2008. *Empirical Economics*, *forthcoming*.
- 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. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting*, 34(2), 339–354.
- Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *The Review of Financial Studies*, 22(12), 5027–5067.
- Marcellino, M., Porqueddu, M., & Venditti, F. (2016). Short-term gdp forecasting with a mixed-frequency dynamic factor model with stochastic volatility. *Journal of Business & Economic Statistics*, 34(1), 118–127.
- 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.
- McCracken, M. W., & Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4), 574–589.
- Peng, W., Swanson, N. R., Yang, X., & Yao, C. (2023). Macroeconomic and financial uncertainty measures in a big data environment. *Working Paper*.
- Racine, J. (2000). Consistent cross-validatory model-selection for dependent data: hv-block cross-validation. *Journal of econometrics*, 99(1), 39–61.
- Rossi, B., & Sekhposyan, T. (2011). Understanding models’ forecasting performance. *Journal of Econometrics*, 164(1), 158–172.

- Schumacher, C. (2007). Forecasting german gdp using alternative factor models based on large datasets. *Journal of Forecasting*, 26(4), 271–302.
- Stock, J. H., & Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44(2), 293–335.
- Stock, J. H., & Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460), 1167–1179.
- Stock, J. H., & Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2), 147–162.
- Stock, J. H., & Watson, M. W. (2005). *Implications of dynamic factor models for var analysis*. National Bureau of Economic Research Cambridge, Mass., USA.
- Stock, J. H., & Watson, M. W. (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business & Economic Statistics*, 30(4), 481–493.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 58(1), 267–288.
- Tinbergen, J. (1939). Business cycles in the usa, 1919-1932. *Statistical Testing of Business Cycle Theory, Volume II*.

Table 1: Target Forecast Variables\*

Group	Target Variable	Abbreviation	Data Transformation
Group1: Output and Income	Real Personal Income	RPI	$\Delta \log(y_t)$
Group1: Output and Income	Industrial Production	INDPRO	$\Delta \log(y_t)$
Group2: Labor Market	Civilian Unemployment Rate	UNRATE	$\Delta y_t$
Group2: Labor Market	Initial Jobless Claims	CLAIMS	$\Delta \log(y_t)$
Group3: Housing	Housing Starts (new, privately owned)	HOUST	$\log(y_t)$
Group3: Housing	Housing Permits (new, privately owned)	PERMIT	$\log(y_t)$
Group4: Consumption and Inventories	Real Personal Consumption Expenditures	RCON	$\Delta \log(y_t)$
Group4: Consumption and Inventories	Real Manufacturing and Trade Industries Sales	MTS	$\Delta \log(y_t)$
Group5: Money and Credit	M2 Money Stock	M2	$\Delta^2 \log(y_t)$
Group6: Interest and Exchange Rates	10-Year Government Treasury Bond Rate	R10	$\Delta y_t$
Group7: Prices	PPI (finished goods)	PPI	$\Delta^2 \log(y_t)$
Group7: Prices	CPI (all items)	CPI	$\Delta^2 \log(y_t)$
Group7: Prices	Personal Consumption Expenditures (chain index)	PCECI	$\Delta^2 \log(y_t)$
Group8: Stock Market	S&P Common Stock Price Index (composite)	S&P 500	$\Delta \log(y_t)$

\* Notes: This table lists the target forecast variables analyzed in our prediction experiments, along with the data transformations used for each variable. For complete details, see Section 4.

Table 2: Forecasting Models Used in Prediction Experiments\*

Forecasting Model	Description of Forecasting Model
AR(SIC)	Autoregression with lag order selected using the SIC. $y_{t+h} = \alpha + \beta_h(L)y_t + \epsilon_{t+h}$
AR(SIC) + BC	AR(SIC) model augmented with business conditions predictors extracted from one of the 10 methods outlined in Section 2.2. $y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_h(L)\widehat{BC}_t + \epsilon_{t+h}$
AR(SIC) + EUI	AR(SIC) model augmented with an economic uncertainty index constructed by applying the methods discussed in Section 2.3. $y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_h(L)\widehat{EUI}_t + \epsilon_{t+h}$
AR(SIC) + BC + EUI	AR(SIC) model augmented with both BC and an EUI, which using the methods outlined in Sections 2.2 and 2.3. $y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_{1h}(L)\widehat{BC}_t + \gamma_{2h}(L)\widehat{EUI}_t + \epsilon_{t+h}$

\* Notes: This table summarizes the forecasting models used in our prediction experiments. The notation used in the forecasting equations is defined in Section 3.

Table 3: Predictive Accuracy of Models that Include Economic Uncertainty Indexes\*

Methods Used to Construct Economic Uncertainty Indexes												
	Target Variable	PCAI	LASSOI	LPCAI	API	GP1I	GP2I	GF1I	GF2I	EPCAI	JLNI	VIX
h=1	RPI	0.986	0.972	0.973	0.981	0.981	0.979	<b>0.952</b>	1	0.989	0.977	0.979
	INDPRO	0.952 *	0.975 **	0.976 *	0.973 *	<b>0.889</b>	0.966 **	0.968 *	0.982 **	0.959	0.978	0.931
	UNRATE	0.942	0.997	0.994	0.995	0.95	0.971	0.949	1.027	0.932	<b>0.854 **</b>	0.887 *
	CLAIMS	1.004 *	1.01 **	1.014 *	1.009 *	0.991	1.006 *	1.008	1.016 **	1.008 *	1.004	<b>0.981</b>
	HOUST	0.967 *	0.973 **	0.973 *	0.964	<b>0.941</b>	0.978	0.957 *	0.969 *	0.968 *	0.952	0.953 **
	PERMIT	0.974 **	0.973 **	0.98 *	0.965 *	<b>0.941 *</b>	0.981 *	0.961 *	0.956 *	0.972 **	0.966	0.974 *
	RCON	0.87	0.887	0.906	0.878	<b>0.831</b>	0.906	0.832	0.879	0.874	0.867	0.955
	MTS	0.908	0.927 **	0.943 *	0.938	<b>0.834</b>	0.934 *	0.906	0.96 **	0.903 *	0.872	0.931
	M2	1.018	1.02	1.02	1.02	1.029	<b>1.016</b>	1.027	1.018	1.018	1.044	1.031
	R10	1.026	1.023	1.032	1.035	1.035	1.022	1.035	1.03	1.026	1.06	<b>1.015</b>
	PPI	1.027	1.025	1.033	1.034	1.04	<b>1.022</b>	1.04	1.027	1.029	1.061	1.028
	CPI	1.032	<b>1.027</b>	1.039	1.037	1.048	<b>1.027</b>	1.05	1.03	1.034	1.086	1.047
	PCECI	1.03	1.027	1.036	1.037	1.048	<b>1.026</b>	1.047	1.029	1.032	1.074	1.038
	S&P 500	1.027	1.019	1.035	1.032	<b>1.014</b>	1.024	1.025	1.042	1.029	1.043	1.027
h=3	RPI	0.979	0.985	0.955	0.995	1.011 *	0.981	0.945	1.006	0.975	<b>0.937</b>	0.958
	INDPRO	1.006	1.014	1.012	1.045	1.012	1.006	1.003	1.051	1.005	1.005	<b>1.002</b>
	UNRATE	0.941	1.002	0.922	0.995	1.018	0.964	0.942	0.999	0.931	<b>0.888</b>	0.938
	CLAIMS	1.013	1.017	1.021	1.029	1.014	1.01	1.015	1.027	1.017	1.013	<b>1.007</b>
	HOUST	0.931 **	1.002	0.87 ***	0.922 *	1.003	0.937 *	0.868 **	<b>0.849 ***</b>	0.915 *	0.872 ***	0.977
	PERMIT	0.932 ***	0.973 *	0.903 **	0.938	1.004	0.922 **	0.897 **	<b>0.86 ***</b>	0.933 ***	0.894 **	0.975
	RCON	0.995	1.022 *	0.981	0.985	<b>0.946</b>	1	0.971	0.977	<b>0.984</b>	0.969	1
	MTS	0.991	1.012	1.017	1.043	1.007	0.995	1.021	1.052	<b>0.984</b>	1.007	1.006
	M2	1.018	<b>1.003 **</b>	1.06	1.036	1.018 ***	1.016 ***	1.052	1.052	1.026	1.046	1.017
	R10	1.02	<b>1.003</b>	1.076	1.05	1.048	1.022 *	1.067	1.076	1.034	1.062	1.013
	PPI	1.019 **	<b>1.002</b>	1.049	1.038	1.024 ***	1.015 *	1.043	1.052	1.031 *	1.043	1.019
	CPI	1.022	<b>1.002</b>	1.079	1.047	1.023 **	1.021 *	1.062	1.073	1.04	1.065	1.013
	PCECI	1.02	<b>1.002</b>	1.074	1.044	1.025 **	1.019	1.061	1.069	1.037	1.06	1.01
	S&P 500	1.05	1.023	1.081	1.108	1.033	1.035	1.118	1.11	1.073	1.087	<b>1.022</b>
h=12	RPI	<b>0.985</b>	1	1.092	0.998	1.135	0.99 **	1.051 *	1.012	1.098	1.013	1.007
	INDPRO	1.011	1.017	1.086	1.021	1.048	<b>1.008</b>	1.062	1.045	1.104	<b>1.008</b>	1.01 *
	UNRATE	0.996	1.019	1.084	1.036 *	1.128	<b>0.989</b>	1.121	1.081	1.087	1.019	1.009
	CLAIMS	1.014	1.003	1.092	1.011	1.029	1.007	1.029	1.033	1.093	<b>0.999</b>	1.003
	HOUST	1.066	1.005	1.097	0.959	1.183	1.069	1.102	1.054	1.069	<b>0.871</b>	0.97
	PERMIT	1.061	1.008	1.123	0.926	1.231	1.065	1.099	1.04	1.084	<b>0.894</b>	0.967
	RCON	<b>0.984</b>	1.026 *	1.081 ***	1.002	1.147 *	0.988	1.105 **	1.053 *	1.057	1.021	0.997
	MTS	<b>0.999</b>	1.015	1.087	1.001	1.077	1.001	1.081 **	1.063 *	1.097	1.014	1.012 **
	M2	1.01	<b>1</b>	1.062 *	1.022	1.063 **	1.004 ***	1.058	1.055	1.053 ***	1.015	1.021
	R10	<b>1.004</b>	1.005	1.056	1.013	1.098	1.008 ***	1.076 *	1.045	1.031	1.019	1.006
	PPI	1.012	<b>1.002</b>	1.067 **	1.023	1.046 ***	1.004 ***	1.079 *	1.038 *	1.074 **	1.044	1.012 *
	CPI	1.011	<b>1.002</b>	1.031 **	1.02 *	1.049	1.003 *	1.034 **	1.022 **	1.046 **	1.026	1.008
	PCECI	1.011	<b>1.002</b>	1.036 ***	1.016 *	1.04	1.004	1.045 ***	1.029 **	1.041 **	1.022	1.009 ***
	S&P 500	1.006	1.009	1.059 ***	<b>1</b>	1.059	1.01	1.093 *	1.065 *	1.076 ***	1.04	1.001

\* Notes: See notes to Table 2. The column headers define the big-data method used to construct our EUIs. For example, PCAI denotes EUIs constructed by utilizing business conditions predictors extracted using the PCA method. Likewise, LASSOI denotes EUIs constructed by utilizing business conditions predictors extracted using the LASSO method, etc. All methods are implemented in this way, except JLNI and VIX, for which indexes are collected from the internet, as discussed in Section 2.2. Tabulated entries are relative mean squared forecast error (MSFEs) for our 14 target variables, and for forecast horizons of h=1,3 and 12 months ahead. The AR(SIC) benchmark model is in the denominator of the reported statistics, so that entries less than unity indicate that our more complex models which include economic uncertainty indexes (EUIs) have lower MSFEs. The forecast period is 2004:6-2019:12, and all models are estimated anew prior to the construction of each forecast. Entries in bold denote method with lowest relative MSFE for a given target variable and forecast horizon. Starred entries indicate rejection of the null hypothesis of equal conditional predictive ability using the [Giacomini and White \(2006\)](#) conditional predictive accuracy test. Significance levels for the test include “\*\*\*” for  $p < 0.01$ , “\*\*” for  $p < 0.05$ , “\*” for  $p < 0.1$ . See Sections 2 and 3 for complete details.

Table 4: Predictive Accuracy of Models That Include Business Conditions Predictors\*

Methods used to Construct Business Conditions Predictors												
	Target Variable	PCA	LASSO	LPCA	AP	GP1	GP2	GF1	GF2	EPCA	JLN	ADS
h=1	RPI	0.994	1.043	1.032	1.032	1.002	1.011	1.03	1.048	0.994	1.031	<b>0.922</b>
	INDPRO	1.012	1.126	0.976	1.038	0.922	1.082	0.914	0.938	1.021	1.614	<b>0.772</b>
	UNRATE	0.822 *	0.844	<b>0.8 **</b>	0.817 *	0.843 **	0.825 *	0.837 **	0.882	0.834	0.835	0.803 **
	CLAIMS	0.957	0.995	1.036	0.987	0.979	0.99	0.975	0.962	1.106	<b>0.915</b>	
	HOUST	0.761 **	0.727 **	0.882 *	0.798 *	<b>0.66 ***</b>	0.838	0.718 ***	0.884	0.761 **	0.858 *	0.964
	PERMIT	<b>0.904</b>	1.203	1.007	0.971	1.085	0.972	1.052	0.998	1.023	0.977	0.98
	RCON	1.007	1.026 *	1.026 *	0.978	1.027	0.918	1.083	1.072	1	0.915	<b>0.839 *</b>
	MTS	0.908	1.012	1.012	1.025	0.885 **	0.91	0.955	0.941	0.931	0.896	<b>0.687 **</b>
	M2	1.125	1.042	1.042	1.056	<b>1.002</b>	1.162	1.036	1.06	1.249 **	1.035 *	1.039
	R10	1.096	1.4 ***	<b>0.906</b>	1.04	0.999 *	1.036	1.198	0.936	1.11	1.215 *	1.07
	PPI	1.061	1.009	1.009	<b>0.921 **</b>	0.967	0.955 *	0.979	0.979	1.076	1.168	1.051
	CPI	1.168	0.94	0.94	1.031	1.014	1.059	0.954	<b>0.912</b>	1.073	1.079	1.067
	PCECI	1.132 *	<b>0.896</b>	<b>0.896</b>	0.973	1.408	1.064	0.924	0.924	1.132 *	1.024	1.055
	S&P 500	1.194	1.075	1.075	1.139 *	1.026	1.098	1.219 *	1.163	1.194	1.058	<b>0.966</b>
h=3	RPI	0.967	1.046	0.997	1.009	0.996	0.955	0.992	0.962	0.967	0.991	<b>0.93</b>
	INDPRO	0.96	1.191	1.099	1.117	1.017	1.002	0.996	1.045	0.977	0.999	<b>0.943</b>
	UNRATE	0.884	0.968	0.903	0.934	0.837	0.915	0.934	0.949	0.905	0.976	<b>0.822</b>
	CLAIMS	1.091 *	1.058	1.058	1.095 *	1.002	1.075 *	1.031	<b>0.985</b>	1.099	1.118	0.996
	HOUST	0.796	<b>0.745 **</b>	0.822 **	0.864	0.786 **	0.921	0.828	0.88	0.784 **	0.763 *	0.944 *
	PERMIT	0.895	<b>0.812 *</b>	0.846 ***	0.894	0.984	0.896	1.072	0.976	0.883	1.011	0.945 **
	RCON	1.14 *	1.007	1.007	1.058	1.12 **	1.085	1.079	1.085	1.162 *	1.054	<b>0.964</b>
	MTS	1.148	1.005	1.005	1.113	1.02	1.122	1.112	1.147	1.173	1.07	<b>0.986</b>
	M2	1.14 **	1.016	1.016	1.035	<b>1.01</b>	1.108	1.036	1.06	1.177 **	1.09	1.029
	R10	1.363 *	1.58 ***	1.015	1.041	1.031 *	1.181 **	1.094	<b>1.009</b>	1.399 **	1.613	1.056
	PPI	1.165 *	1.026 *	1.026 *	1.033	<b>1.011</b>	1.106 *	1.059 ***	1.059 ***	1.192 **	1.218 **	1.022
	CPI	1.223	1.074	1.074	1.195 *	<b>1.034</b>	1.129 *	1.067	1.063 *	1.26 *	1.129	1.039
	PCECI	1.215 **	1.043	1.043	1.125	1.236	1.128	1.066	1.066	1.215 **	1.069	<b>1.037</b>
	S&P 500	1.19	1.069	1.069	1.083	<b>1.007</b>	1.153	1.105	1.151	1.19	1.092	1.03
h=12	RPI	1.018	1.021	1.025	1.114	<b>0.955</b>	1.076	1.022	1.073	1.018	1.079	0.983
	INDPRO	1.17	1.047	1.094	1.094	1.022 ***	1.098	1.008	1.028	1.182	1.157	<b>1.005</b>
	UNRATE	1.007	<b>0.915</b>	0.942	1.231	1.025	1.004	1.242	1.124	1.021	1.17	0.994
	CLAIMS	1.073	1.006	1.006	1.01	0.994	1.039	<b>0.992</b>	1.003	1.074	1.123	0.995
	HOUST	0.712 *	<b>0.388 ***</b>	0.778	0.731	0.728 *	0.816	0.821 *	1.022	0.678 *	0.696 **	1.014
	PERMIT	0.719 *	0.568 **	0.637 *	0.74	0.707	1.033	1.027	1.068	<b>0.545 **</b>	0.969	1.019
	RCON	1.159	<b>0.964</b>	<b>0.964</b>	1.017	1.104	1.077	1.026	1.041	1.174	1.031	1.016
	MTS	1.063	1.002	1.002	0.993	1.001	1.037	1.083	1.106	1.046	1.044	<b>0.992</b>
	M2	1.336 *	1.144 *	1.144 *	1.258 *	1.02	1.407 *	1.031	1.244 *	1.434	1.077 ***	1.012
	R10	1.18 **	1.617 ***	<b>1.012</b>	1.014	1.024	1.086	1.154 ***	1.02 *	1.184 ***	1.221 *	1.016
	PPI	1.12 ***	1.012	1.012	1.034	<b>1.004</b>	1.114 *	1.037	1.037	1.129 ***	1.215	1.034
	CPI	1.119 **	1.056 *	1.056 *	1.11	<b>1.014</b>	1.082 *	1.056	1.063	1.145 ***	1.085 *	1.015
	PCECI	1.109	1.07	1.07	1.063	1.211	1.072	1.151	1.151	1.109	1.067	<b>1.011</b>
	S&P 500	1.081	0.988	0.988	1.086 *	<b>0.937</b>	1.137	0.997	1.02	1.081	1.028	1.014

\* Notes: See notes to Table 3. In this table, the benchmark model is again the AR(SIC) model. However, the alternative is our forecasting model that includes business conditions predictors (BCs), constructed using the big-data methods outlined in Section 2.2, including PCA, LASSO, etc. This nomenclature is used for all predictor extraction methods, including our benchmark ADS method, which is discussed in Section 2.2. Note that although the ADS is an index, results based on its use are included in this table. This is because it is not an uncertainty index, in the sense that it is not constructed using forecast errors. Also, the ADS is defined as a business conditions index on the website of Federal Reserve Bank of Philadelphia. See Section 4 for further details.



Table 5: Predictive Accuracy of Models That Include Both Economic Uncertainty Indexes and Business Conditions Predictors\*

Methods used to Construct Economic Uncertainty Indexes and Business Conditions Predictors											
	Target Variable	PCA(I)	LASSO(I)	LPCA(I)	AP(I)	GP1(I)	GP2(I)	GF1(I)	GF2(I)	EPCA(I)	JLN(I)
h=1	RPI	1.034	1.048	1.031	1.048	<b>0.992</b>	1.05	1.023	1.064	1.037	1.077
	INDPRO	1.024	1.163	0.994	1.033	<b>0.894</b>	1.092	0.929	0.957 **	1.037	1.402
	UNRATE	0.838	0.891	<b>0.823 *</b>	0.848	0.853 *	0.843	0.861	0.918	0.845	0.832
	CLAIMS	<b>0.983</b>	1.024	1.021	1.057	1.002	0.991	1.024	0.998	0.992	1.06
	HOUST	0.781 **	0.737 **	0.891 *	0.803 *	<b>0.658 ***</b>	0.833 *	0.731 ***	0.869	0.785 *	0.863
	PERMIT	<b>0.869 *</b>	1.187	1.011	0.952	1.004	0.972	1.004	0.933	0.984	0.915
	RCON	1.014	0.906	0.922	0.954	<b>0.905</b>	0.952	0.927	0.977	1.007	0.915
	MTS	0.9	0.944 **	0.96 *	0.996	<b>0.806</b>	0.938	0.929	0.942	0.922	0.888
	M2	1.153	1.058	1.055	1.073	<b>1.022</b>	1.199	1.064	1.077	1.287 **	1.09
	R10	1.196	1.411 ***	<b>0.928</b>	1.053	1.034	1.099	1.221 **	0.963	1.203	1.364 *
	PPI	1.086	1.034	1.042	<b>0.939</b>	1.005	0.968 *	1.023	1.007	1.103	1.184
	CPI	1.194	0.96	0.971	1.037	1.057	1.072	0.997	<b>0.934</b>	1.107	1.098
h=3	PCECI	1.152 **	<b>0.928</b>	0.937	0.996	1.442	1.074	0.934	0.929	1.153 **	1.036
	S&P 500	1.194 *	1.079	1.082	1.173 *	<b>1.032</b>	1.095	1.226 *	1.158	1.194 *	1.083
	RPI	0.985	1.052	0.981	1.029	1.011	<b>0.971</b>	0.974	0.974	0.985	1
	INDPRO	0.982	1.223	1.074	1.075	1.032	1.034	1.018	1.077	<b>0.974</b>	1.104
	UNRATE	0.881	0.964	0.876	0.927	<b>0.862</b>	0.893	0.927	0.934	0.897	0.943
	CLAIMS	1.088 *	1.065	1.103	1.102 *	<b>1.01</b>	1.058	1.074	1.019	1.092	1.117 *
	HOUST	0.77 *	0.78 **	0.786 ***	0.774 **	0.772 **	0.864	0.797 *	0.78	<b>0.762 *</b>	0.774 *
	PERMIT	0.855	0.814 *	<b>0.758 **</b>	0.826	0.995	0.88	0.929	0.842	0.806	0.867
	RCON	1.16	1.034	<b>0.985</b>	1.079	1.05	1.1	1.037	1.032	1.192 *	1.112
	MTS	1.177	<b>1.016</b>	1.024	1.106	1.03	1.143	1.126	1.163	1.176	1.115
	M2	1.174 **	<b>1.02</b>	1.057	1.058	1.027	1.115 *	1.09	1.101	1.194 **	1.146
	R10	1.33	1.545 ***	1.094	<b>1.069</b>	1.078 **	1.203 **	1.14	1.092	1.346 *	1.895
h=12	PPI	1.18 **	<b>1.028 *</b>	1.075 **	1.065 ***	1.036	1.117 *	1.098 ***	1.111 ***	1.213 **	1.188 *
	CPI	1.241	1.078 *	1.143	1.218 *	<b>1.053</b>	1.142 **	1.128	1.129	1.284 *	1.187
	PCECI	1.232 **	1.046	1.113	1.175	1.266	1.142	1.109	1.126	1.235 **	<b>1.025</b>
	S&P 500	1.201	1.098	1.125	1.201	<b>1.024</b>	1.139	1.195	1.242	1.2	1.177
	RPI	<b>1.024</b>	1.032	1.136	1.067	1.135	1.082	1.077	1.078	1.178	1.109
	INDPRO	1.17	1.074	1.122	<b>1.03</b>	1.071 **	1.103	1.072	1.086 ***	1.241	1.108
	UNRATE	0.987	<b>0.924</b>	1.045 *	1.125	1.142	1.013	1.389 **	1.174	1.116	1.224
	CLAIMS	1.072	<b>1.01</b>	1.113	1.018	1.028	1.056	1.025	1.049	1.21 ***	1.102
	HOUST	0.668 **	<b>0.389 ***</b>	0.831	0.643	0.878	0.811	0.863	1.006	0.741	0.697 **
	PERMIT	0.666 **	0.556 **	0.766 *	0.686	0.869	0.897	1.119	1.034	<b>0.55 **</b>	0.933
	RCON	1.14	<b>0.998</b>	1.038 **	<b>0.998</b>	1.198	1.073	1.099 **	1.079	1.217	1.073
	MTS	1.065	1.018	1.091	<b>0.979</b>	1.072	1.056	1.141	1.154	1.119	1.064
h=12	M2	1.349 *	1.145 *	1.209 ***	1.301 **	<b>1.075</b>	1.413 *	1.09	1.315 **	1.512 *	1.097 ***
	R10	1.17 **	1.63 ***	1.069 *	<b>1.034 ***</b>	1.12	1.089	1.249 ***	1.066 **	1.243 ***	1.226 **
	PPI	1.129 ***	<b>1.014</b>	1.078 ***	1.054 **	1.053 **	1.122 **	1.115 **	1.075 ***	1.204 ***	1.262
	CPI	1.13 **	<b>1.057 *</b>	1.089 ***	1.13 *	1.066 **	1.099 **	1.097 **	1.086 *	1.197 ***	1.112 *
	PCECI	1.121	<b>1.071</b>	1.108 *	1.08	1.262	1.083	1.196 *	1.188 *	1.157 **	1.086
	S&P 500	1.104	<b>1.004</b>	1.045	1.117 ***	1.035	1.139	1.094 *	1.096	1.184 *	1.084

\* Notes: See notes to Tables 3 and 4. In this table, the benchmark model is again the AR(SIC) model. However, the alternative is our forecasting model that includes both business conditions predictors (BCs) and economic uncertainty indexes (EUIs), which are constructed using the methods outlined in Section 2.2 and Section 2.3. Note that the case where both BC predictors and EUIs are constructed using PCA is denoted as PCA(I). This notation allows us to differentiate between the method called “PCA”, where only BC predictors are used in our predictive regressions, and said predictors are extracted using PCA, and the method called “PCAI”, where EUIs are used in our predictive regressions, and said EUIs are constructed using BC predictors extracted using PCA.

Table 6: Overall “MSFE-Best” Forecasting Models and Methods\*

Target Variable	Rank	h=1	h=3	h=12
RPI	1	AR+BC (ADS)	AR+BC (ADS)	AR+BC (GP1)
	2	AR+EUI (GF1I)	AR+EUI (JLNI)	AR+BC (ADS)
	3	AR+EUI (LASSOI)	AR+EUI (GF1I)	AR+EUI (PCAI)
INDPRO	1	AR+BC (ADS)	AR+BC (ADS)	AR
	2	AR+EUI (GP1I)	AR+BC (PCA)	AR+BC (ADS)
	3	AR+BC+EUI (GP1(I))	AR+BC+EUI (EPCA(I))	AR+BC (GF1)
UNRATE	1	AR+BC (LPCA) **	AR+BC (ADS)	AR+BC (LASSO)
	2	AR+BC (ADS) **	AR+BC (GP1)	AR+BC+EUI (LASSO(I))
	3	AR+BC (AP) *	AR+BC+EUI (GP1(I))	AR+BC (LPCA)
CLAIMS	1	AR+BC (ADS)	AR+BC (GF2)	AR+BC (GF1)
	2	AR+BC (PCA)	AR+BC (ADS)	AR+BC (GP1)
	3	AR+BC (EPCA)	AR	AR+BC (ADS)
HOUST	1	AR+BC+EUI (GP1(I)) ***	AR+BC (LASSO) **	AR+BC (LASSO) ***
	2	AR+BC (GP1) ***	AR+BC+EUI (EPCA(I)) *	AR+BC+EUI (LASSO(I)) ***
	3	AR+BC (GF1) ***	AR+BC (JLN) *	AR+BC+EUI (AP(I))
PERMIT	1	AR+BC+EUI (PCA(I)) *	AR+BC+EUI (LPCA(I)) **	AR+BC (EPCA) **
	2	AR+BC (PCA)	AR+BC+EUI (EPCA(I))	AR+BC+EUI (EPCA(I)) **
	3	AR+BC+EUI (JLN(I))	AR+BC (LASSO) *	AR+BC+EUI (LASSO(I)) **
RCON	1	AR+EUI (GP1I)	AR+EUI (GP1I)	AR+BC (LASSO)
	2	AR+EUI (GF1I)	AR+BC (ADS)	AR+BC (LPCA)
	3	AR+BC (ADS) *	AR+EUI (JLNI)	AR+EUI (PCAI)
MTS	1	AR+BC (ADS) **	AR+EUI (EPCAI)	AR+BC+EUI (AP(I))
	2	AR+BC+EUI (GP1(I))	AR+BC (ADS)	AR+BC (ADS)
	3	AR+EUI (GP1I)	AR+EUI (PCAI)	AR+BC (AP)
M2	1	AR	AR	AR+EUI (LASSOI)
	2	AR+BC (GP1)	AR+EUI (LASSOI) **	AR
	3	AR+EUI (GP2I)	AR+BC (GP1)	AR+EUI (GP2I) ***
R10	1	AR+BC (LPCA)	AR	AR
	2	AR+BC+EUI (LPCA(I))	AR+EUI (LASSOI)	AR+EUI (PCAI)
	3	AR+BC (GF2)	AR+BC (GF2)	AR+EUI (LASSOI)
PPI	1	AR+BC (AP) **	AR	AR
	2	AR+BC+EUI (AP(I))	AR+EUI (LASSOI)	AR+EUI (LASSOI)
	3	AR+BC (GP2) *	AR+BC (GP1)	AR+EUI (GP2I) ***
CPI	1	AR+BC (GF2)	AR	AR
	2	AR+BC+EUI (GF2(I))	AR+EUI (LASSOI)	AR+EUI (LASSOI)
	3	AR+BC (LASSO)	AR+EUI (VIX)	AR+EUI (GP2I) *
PCECI	1	AR+BC (LASSO)	AR	AR
	2	AR+BC (LPCA)	AR+EUI (LASSOI)	AR+EUI (LASSOI)
	3	AR+BC (GF1)	AR+EUI (VIX)	AR+EUI (GP2I)
S&P 500	1	AR+BC (ADS)	AR	AR+BC (GP1)
	2	AR	AR+BC (GP1)	AR+BC (LASSO)
	3	AR+EUI (GP1I)	AR+EUI (VIX)	AR+BC (LPCA)

\* Notes: See notes to Table 5. Entries in this table are the top 3 “MSFE-best” forecasting models, selected by comparing forecasts from our benchmark AR(SIC) model with our AR(SIC)+EUI, AR(SIC)+BC, and AR(SIC)+BC+EUI models. For each of these models, the associated predictors extraction method used when constructing the EUIs and BCs is also given, in parentheses. Starred entries indicate rejection of the null hypothesis of equal conditional predictive ability, and indicate that our alternative models that include EUIs and BCs are yield more accurate predictions than our AR benchmark model, using the [Giacomini and White \(2006\)](#) conditional predictive accuracy test. Significance levels for the test include “\*\*\*” for  $p < 0.01$ , “\*\*” for  $p < 0.05$ , “\*” for  $p < 0.1$ . See Section 3 for complete details.

Table 7: Top 3 “MSFE-Best” Methods for Models Including Economic Uncertainty Indexes\*

Target Variable	Rank	h=1	h=3	h=12
RPI	1	AR+EUI (GF1I)	AR+EUI (JLNI)	AR+EUI (PCAI)
	2	AR+EUI (LASSOI)	AR+EUI (GF1I)	AR+EUI (GP2I) **
	3	AR+EUI (LPCAI)	AR+EUI (LPCAI)	AR+EUI (API)
INDPRO	1	AR+EUI (GP1I)	AR	AR
	2	AR+EUI (VIX)	AR+EUI (VIX)	AR+EUI (JLNI)
	3	AR+EUI (PCAI) *	AR+EUI (GF1I)	AR+EUI (GP2I)
UNRATE	1	AR+EUI (JLNI) **	AR+EUI (JLNI)	AR+EUI (GP2I)
	2	AR+EUI (VIX) *	AR+EUI (LPCAI)	AR+EUI (PCAI)
	3	AR+EUI (EPCAI)	AR+EUI (EPCAI)	AR
CLAIMS	1	AR+EUI (VIX)	AR	AR+EUI (JLNI)
	2	AR+EUI (GP1I)	AR+EUI (VIX)	AR
	3	AR	AR+EUI (GP2I)	AR+EUI (VIX)
HOUST	1	AR+EUI (GP1I)	AR+EUI (GF2I) ***	AR+EUI (JLNI)
	2	AR+EUI (JLNI)	AR+EUI (GF1I) **	AR+EUI (API)
	3	AR+EUI (VIX) **	AR+EUI (LPCAI) ***	AR+EUI (VIX)
PERMIT	1	AR+EUI (GP1I) *	AR+EUI (GF2I) ***	AR+EUI (JLNI)
	2	AR+EUI (GF2I) *	AR+EUI (JLNI) **	AR+EUI (API)
	3	AR+EUI (GF1I) *	AR+EUI (GF1I) **	AR+EUI (VIX)
RCON	1	AR+EUI (GP1I)	AR+EUI (GP1I)	AR+EUI (PCAI)
	2	AR+EUI (GF1I)	AR+EUI (JLNI)	AR+EUI (GP2I)
	3	AR+EUI (JLNI)	AR+EUI (GF1I)	AR+EUI (VIX)
MTS	1	AR+EUI (GP1I)	AR+EUI (EPCAI)	AR+EUI (PCAI)
	2	AR+EUI (JLNI)	AR+EUI (PCAI)	AR
	3	AR+EUI (EPCAI) *	AR+EUI (GP2I)	AR+EUI (API)
M2	1	AR	AR	AR+EUI (LASSOI)
	2	AR+EUI (GP2I)	AR+EUI (LASSOI) **	AR
	3	AR+EUI (EPCAI)	AR+EUI (GP2I) ***	AR+EUI (GP2I) ***
R10	1	AR	AR	AR
	2	AR+EUI (VIX)	AR+EUI (LASSOI)	AR+EUI (PCAI)
	3	AR+EUI (GP2I)	AR+EUI (VIX)	AR+EUI (LASSOI)
PPI	1	AR	AR	AR
	2	AR+EUI (GP2I)	AR+EUI (LASSOI)	AR+EUI (LASSOI)
	3	AR+EUI (LASSOI)	AR+EUI (GP2I) *	AR+EUI (GP2I) ***
CPI	1	AR	AR	AR
	2	AR+EUI (LASSOI)	AR+EUI (LASSOI)	AR+EUI (LASSOI)
	3	AR+EUI (GP2I)	AR+EUI (VIX)	AR+EUI (GP2I) *
PCECI	1	AR	AR	AR
	2	AR+EUI (GP2I)	AR+EUI (LASSOI)	AR+EUI (LASSOI)
	3	AR+EUI (LASSOI)	AR+EUI (VIX)	AR+EUI (GP2I)
S&P 500	1	AR	AR	AR
	2	AR+EUI (GP1I)	AR+EUI (VIX)	AR+EUI (API)
	3	AR+EUI (LASSOI)	AR+EUI (LASSOI)	AR+EUI (VIX)

\* Notes: See notes to Table 6. This table lists the top 3 predictors extraction methods used in our “MSFE-best” forecasting models, when comparing our benchmark AR(SIC) model with our AR(SIC)+EUI model.

Table 8: Top 3 “MSFE-Best” Methods for Models Including Business Conditions Predictors\*

Target Variable	Rank	h=1	h=3	h=12
RPI	1	AR+BC (ADS)	AR+BC (ADS)	AR+BC (GP1)
	2	AR+BC (PCA)	AR+BC (GP2)	AR+BC (ADS)
	3	AR+BC (EPCA)	AR+BC (GF2)	AR
INDPRO	1	AR+BC (ADS)	AR+BC (ADS)	AR
	2	AR+BC (GF1)	AR+BC (PCA)	AR+BC (ADS)
	3	AR+BC (GP1)	AR+BC (EPCA)	AR+BC (GF1)
UNRATE	1	AR+BC (LPCA) **	AR+BC (ADS)	AR+BC (LASSO)
	2	AR+BC (ADS) **	AR+BC (GP1)	AR+BC (LPCA)
	3	AR+BC (AP) *	AR+BC (PCA)	AR+BC (ADS)
CLAIMS	1	AR+BC (ADS)	AR+BC (GF2)	AR+BC (GF1)
	2	AR+BC (PCA)	AR+BC (ADS)	AR+BC (GP1)
	3	AR+BC (EPCA)	AR	AR+BC (ADS)
HOUST	1	AR+BC (GP1) ***	AR+BC (LASSO) **	AR+BC (LASSO) ***
	2	AR+BC (GF1) ***	AR+BC (JLN) *	AR+BC (EPCA) *
	3	AR+BC (LASSO) ***	AR+BC (EPCA) **	AR+BC (JLN) **
PERMIT	1	AR+BC (PCA)	AR+BC (LASSO) *	AR+BC (EPCA) **
	2	AR+BC (AP)	AR+BC (LPCA) ***	AR+BC (LASSO) **
	3	AR+BC (GP2)	AR+BC (EPCA)	AR+BC (LPCA) *
RCON	1	AR+BC (ADS) *	AR+BC (ADS)	AR+BC (LASSO)
	2	AR+BC (JLN)	AR	AR+BC (LPCA)
	3	AR+BC (GP2)	AR+BC (LASSO)	AR
MTS	1	AR+BC (ADS) **	AR+BC (ADS)	AR+BC (ADS)
	2	AR+BC (GP1) **	AR	AR+BC (AP)
	3	AR+BC (JLN)	AR+BC (LASSO)	AR
M2	1	AR	AR	AR
	2	AR+BC (GP1)	AR+BC (GP1)	AR+BC (ADS)
	3	AR+BC (JLN) *	AR+BC (LASSO)	AR+BC (GP1)
R10	1	AR+BC (LPCA)	AR	AR
	2	AR+BC (GF2)	AR+BC (GF2)	AR+BC (LPCA)
	3	AR+BC (GP1) *	AR+BC (LPCA)	AR+BC (AP)
PPI	1	AR+BC (AP) **	AR	AR
	2	AR+BC (GP2) *	AR+BC (GP1)	AR+BC (GP1)
	3	AR+BC (GP1)	AR+BC (ADS)	AR+BC (LASSO)
CPI	1	AR+BC (GF2)	AR	AR
	2	AR+BC (LASSO)	AR+BC (GP1)	AR+BC (GP1)
	3	AR+BC (LPCA)	AR+BC (ADS)	AR+BC (ADS)
PCECI	1	AR+BC (LASSO)	AR	AR
	2	AR+BC (LPCA)	AR+BC (ADS)	AR+BC (ADS)
	3	AR+BC (GF1)	AR+BC (LASSO)	AR+BC (AP)
S&P 500	1	AR+BC (ADS)	AR	AR+BC (GP1)
	2	AR	AR+BC (GP1)	AR+BC (LASSO)
	3	AR+BC (GP1)	AR+BC (ADS)	AR+BC (LPCA)

\* Notes: See notes to Table 6. This table lists the top 3 predictors extraction methods used in our “MSFE-best” forecasting models, when comparing our benchmark AR(SIC) model with our AR(SIC)+BC model.

Table 9: Top 3 “MSFE-Best” Methods for Models Including Both Economic Uncertainty Indexes and Business Conditions Predictors\*

Target Variable	Rank	h=1	h=3	h=12
RPI	1	AR+BC+EUI (GP1(I))	AR+BC+EUI (GP2(I))	AR
	2	AR	AR+BC+EUI (GF1(I))	AR+BC+EUI (PCA(I))
	3	AR+BC+EUI (GF1(I))	AR+BC+EUI (GF2(I))	AR+BC+EUI (LASSO(I))
INDPRO	1	AR+BC+EUI (GP1(I))	AR+BC+EUI (EPCA(I))	AR
	2	AR+BC+EUI (GF1(I))	AR+BC+EUI (PCA(I))	AR+BC+EUI (AP(I))
	3	AR+BC+EUI (GF2(I)) **	AR	AR+BC+EUI (GP1(I)) **
UNRATE	1	AR+BC+EUI (LPCA(I)) *	AR+BC+EUI (GP1(I))	AR+BC+EUI (LASSO(I)) **
	2	AR+BC+EUI (JLN(I))	AR+BC+EUI (LPCA(I))	AR+BC+EUI (PCA(I))
	3	AR+BC+EUI (PCA(I))	AR+BC+EUI (PCA(I))	AR
CLAIMS	1	AR+BC+EUI (PCA(I))	AR	AR
	2	AR+BC+EUI (GP2(I))	AR+BC+EUI (GP1(I))	AR+BC+EUI (LASSO(I))
	3	AR+BC+EUI (EPCA(I))	AR+BC+EUI (GF2(I))	AR+BC+EUI (AP(I))
HOUST	1	AR+BC+EUI (GP1(I)) ***	AR+BC+EUI (EPCA(I)) *	AR+BC+EUI (LASSO(I)) ***
	2	AR+BC+EUI (GF1(I)) ***	AR+BC+EUI (PCA(I)) *	AR+BC+EUI (AP(I))
	3	AR+BC+EUI (LASSO(I)) **	AR+BC+EUI (GP1(I)) **	AR+BC+EUI (PCA(I)) **
PERMIT	1	AR+BC+EUI (PCA(I)) *	AR+BC+EUI (LPCA(I)) **	AR+BC+EUI (EPCA(I)) **
	2	AR+BC+EUI (JLN(I))	AR+BC+EUI (EPCA(I))	AR+BC+EUI (LASSO(I)) **
	3	AR+BC+EUI (GF2(I))	AR+BC+EUI (LASSO(I)) *	AR+BC+EUI (PCA(I)) **
RCON	1	AR+BC+EUI (GP1(I))	AR+BC+EUI (LPCA(I))	AR+BC+EUI (LASSO(I))
	2	AR+BC+EUI (LASSO(I))	AR	AR+BC+EUI (AP(I))
	3	AR+BC+EUI (JLN(I))	AR+BC+EUI (GF2(I))	AR
MTS	1	AR+BC+EUI (GP1(I))	AR	AR+BC+EUI (AP(I))
	2	AR+BC+EUI (JLN(I))	AR+BC+EUI (LASSO(I))	AR
	3	AR+BC+EUI (PCA(I))	AR+BC+EUI (LPCA(I))	AR+BC+EUI (LASSO(I))
M2	1	AR	AR	AR
	2	AR+BC+EUI (GP1(I))	AR+BC+EUI (LASSO(I))	AR+BC+EUI (GP1(I))
	3	AR+BC+EUI (LPCA(I))	AR+BC+EUI (GP1(I))	AR+BC+EUI (GF1(I))
R10	1	AR+BC+EUI (LPCA(I))	AR	AR
	2	AR+BC+EUI (GF2(I))	AR+BC+EUI (AP(I))	AR+BC+EUI (AP(I)) ***
	3	AR	AR+BC+EUI (GP1(I)) **	AR+BC+EUI (GF2(I)) **
PPI	1	AR+BC+EUI (AP(I))	AR	AR
	2	AR+BC+EUI (GP2(I)) *	AR+BC+EUI (LASSO(I)) *	AR+BC+EUI (LASSO(I))
	3	AR	AR+BC+EUI (GP1(I))	AR+BC+EUI (GP1(I)) **
CPI	1	AR+BC+EUI (GF2(I))	AR	AR
	2	AR+BC+EUI (LASSO(I))	AR+BC+EUI (GP1(I))	AR+BC+EUI (LASSO(I)) *
	3	AR+BC+EUI (LPCA(I))	AR+BC+EUI (LASSO(I)) *	AR+BC+EUI (GP1(I)) **
PCECI	1	AR+BC+EUI (LASSO(I))	AR	AR
	2	AR+BC+EUI (GF2(I))	AR+BC+EUI (JLN(I))	AR+BC+EUI (LASSO(I))
	3	AR+BC+EUI (GF1(I))	AR+BC+EUI (LASSO(I))	AR+BC+EUI (AP(I))
S&P 500	1	AR	AR	AR
	2	AR+BC+EUI (GP1(I))	AR+BC+EUI (GP1(I))	AR+BC+EUI (LASSO(I))
	3	AR+BC+EUI (LASSO(I))	AR+BC+EUI (LASSO(I))	AR+BC+EUI (GP1(I))

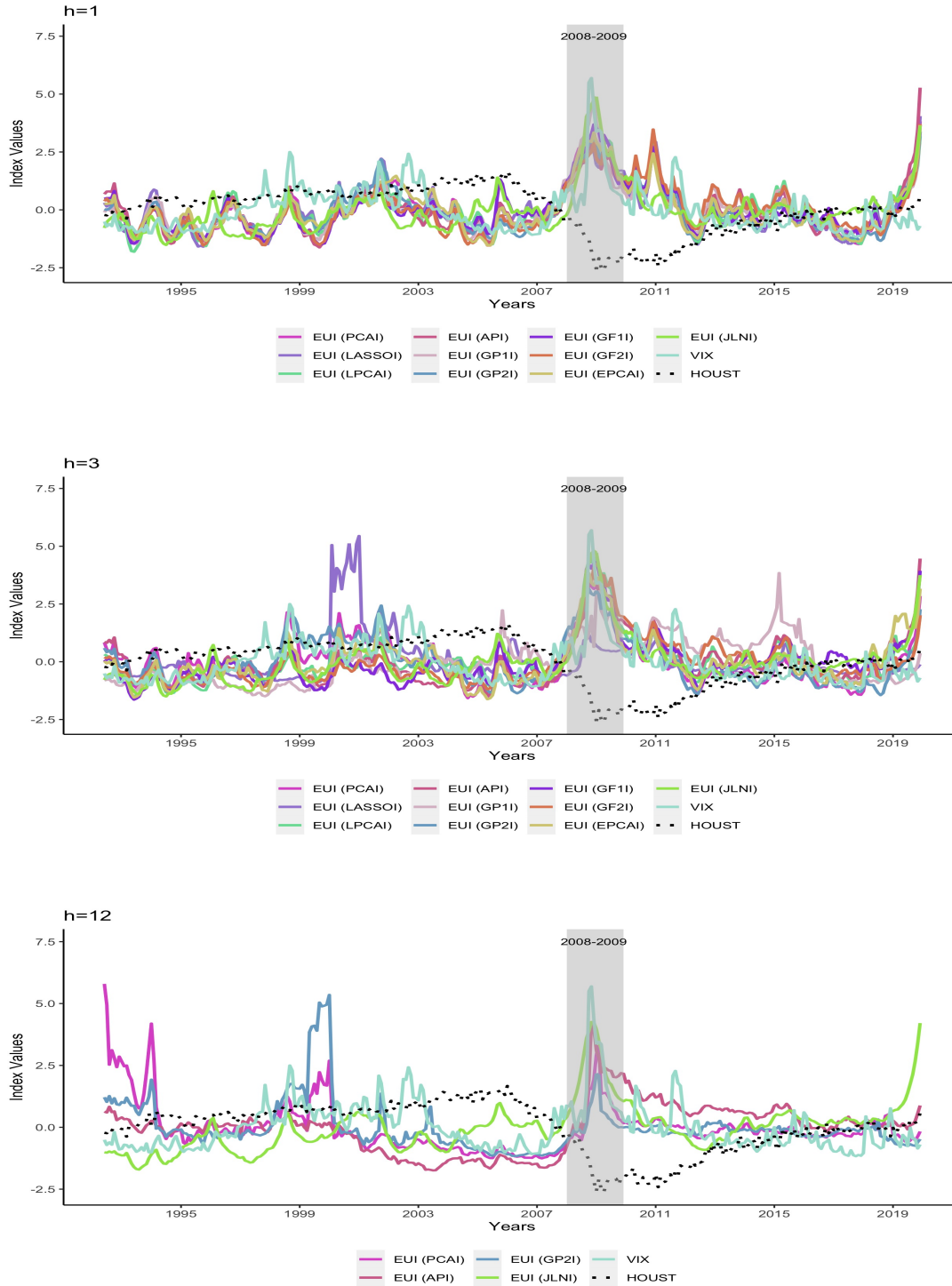
\* Notes: See notes to Table 6. This table lists the top 3 predictors extraction methods used in our “MSFE-best” forecasting models, when comparing our benchmark AR(SIC) model with our AR(SIC)+BC+EUI model.

Table 10: Predictive Accuracy of Average Economic Uncertainty Indexes Forecasts and Forecasts Combinations\*

		Methods for Indexes Average and Forecasts Combinations					
	Target Variable	FAV1	FAV2	FAV3	Comb1	Comb2	Comb3
h=1	RPI	0.977	0.979	0.985	0.974	0.968	0.959
	INDPRO	0.956 *	0.931	0.942	0.948	0.936	0.909
	UNRATE	0.97	0.886 *	0.901 *	0.957	0.926	0.902
	CLAIMS	1.006 *	0.981	0.987	1.004	0.997	0.985
	HOUST	0.96 *	0.953 **	0.956 **	0.959	0.954	0.952
	PERMIT	0.963 **	0.974 *	0.976 **	0.959	0.956	0.955
	RCON	0.863	0.954	0.966	0.859	0.851	0.842
	MTS	0.909 *	0.93	0.949	0.903	0.887	0.856
	M2	1.021	1.031	1.03	1.019	1.021	1.022
	R10	1.031	1.016	1.014	1.027	1.027	1.029
	PPI	1.032	1.029	1.026	1.030	1.031	1.032
	CPI	1.037	1.047	1.044	1.034	1.037	1.039
	PCECI	1.036	1.039	1.035	1.034	1.036	1.037
	S&P 500	1.029	1.028	1.022	1.022	1.015	1.008
h=3	RPI	0.976	0.957	0.965	0.961	0.955	0.949
	INDPRO	0.988	1.002	1.004	0.979	0.976	0.971
	UNRATE	0.969	0.938	0.95	0.922	0.912	0.892
	CLAIMS	1.017	1.007	1.008	1.008	1.006	1.005
	HOUST	0.945 *	0.976	0.977	0.892	0.890	0.892
	PERMIT	0.93	0.974	0.975	0.901	0.897	0.898
	RCON	0.987	1	1.002	0.956	0.956	0.955
	MTS	0.984	1.006	1.008	<b>0.979</b>	<b>0.979</b>	<b>0.977</b>
	M2	1.008 **	1.017	1.016	1.030	1.029	1.029
	R10	1.011	1.013	1.008	1.033	1.032	1.033
	PPI	1.008	1.019	1.019	1.028	1.028	1.027
	CPI	1.009	1.013	1.013	1.038	1.037	1.037
	PCECI	1.008	1.01	1.01	1.035	1.034	1.033
	S&P 500	1.03	1.022	1.021	1.043	1.040	1.038
h=12	RPI	1.053	1.053	1.053	0.998	0.995	0.989
	INDPRO	1.064 **	1.064 **	1.064 **	1.009	1.004	1.003
	UNRATE	1.145 **	1.145 **	1.145 **	0.998	0.991	0.988
	CLAIMS	1.012	1.012	1.012	1.015	1.009	1.007
	HOUST	1.133	1.133	1.133	0.932	0.905	0.909
	PERMIT	1.184	1.184	1.184	0.915	0.891	0.896
	RCON	1.106	1.106	1.106	0.999	0.995	0.996
	MTS	1.073 *	1.073 *	1.073 *	1.013	1.009	1.007
	M2	1.021	1.021	1.021	1.029	1.026	1.024
	R10	1.031	1.031	1.031	1.004	1.003	1.003
	PPI	1.032 *	1.032 *	1.032 *	1.029	1.028	1.028
	CPI	1.031	1.031	1.031	1.018	1.017	1.017
	PCECI	1.043	1.043	1.043	1.017	1.016	1.016
	S&P 500	1.029	1.029	1.029	1.023	1.018	1.017

\* Notes: The column headers define three average indexes forecasts and three forecasts combinations. Tabulated entries are relative mean squared forecast error (MSFEs) for our 14 target variables, and for forecast horizons of h=1,3 and 12 months ahead. The AR(SIC) benchmark model is in the denominator of the reported statistics, so that entries less than unity indicate that our more complex models which include average of economic uncertainty indexes (EUIs) or forecasts combinations have lower MSFEs. The forecast period is 2004:6-2019:12, and all models are estimated anew prior to the construction of each forecast. Entries in bold denote method can beat all other forecast models summarized in Table 6 for a given target variable and forecast horizon. Starred entries indicate rejection of the null hypothesis of equal conditional predictive ability using the [Giacomini and White \(2006\)](#) conditional predictive accuracy test. Significance levels for the test include “\*\*\*” for  $p < 0.01$ , “\*\*” for  $p < 0.05$ , “\*” for  $p < 0.1$ . See Sections 2 and 3 for complete details.

Figure 1: Plots of Housing Starts, Various Economic Uncertainty Indexes, and the VIX\*



\*Notes: See notes to Table 3. The above three figures include various plots of our economic uncertainty indexes (EUIs), HOUST, and the VIX. All EUI plots depict real-time measures constructed for the period 1992:6-2019:12, and updated prior to the construction of each new forecast reported on in Tables 3-9. The shaded regions in the plots correspond roughly to the 2008 Financial Crisis (2008:1-2009:12). And we only keep some reasonable EUIs in  $h=12$  because EUIs do not perform well in  $h=12$  forecasting experiments.