

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

Yang Liu and Norman R. Swanson
Rutgers University

April 2023

Abstract

In this paper, we evaluate the marginal predictive content of a variety of new business conditions and economic uncertainty indexes. Our indexes are defined as latent factors extracted from a high dimensional macroeconomic dataset (business conditions indexes) and as functions of predictive errors from models that incorporate these latent factors (economic uncertainty indexes). Estimation of the indexes is based on a number of extant and novel machine learning methods that combine variable selection and shrinkage methods including principal components analysis and the least absolute shrinkage operator. When predicting 14 monthly series selected from 8 different groups of economic variables, our new indexes are shown to result in significant improvements in predictive accuracy, relative to predictions made using benchmark models. Moreover, while inclusion of either business conditions indexes or economic uncertainty indexes yields forecast accuracy improvements, greater predictive gains accrue when both types of indexes are included in predictive regressions. Additionally, our business conditions indexes are found to be particularly useful for short-term forecasting, while our economic uncertainty indexes exhibit their best predictive content for long-term forecasting. Finally, our mean-square forecast error “best” indexes are almost always constructed by utilizing a novel factor-lasso estimation method that accounts for the high level of collinearity across the variables in our macroeconomic dataset.

Keywords: Latent factor, business conditions index, 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. We are 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 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 business conditions and economic uncertainty indexes.

Examples of business conditions indexes include the index proposed in [Aruoba, Diebold, and Scotti \(2009a\)](#). These authors extract business conditions indexes from four key macroeconomic variables of different observational frequencies.¹ Business conditions indexes are also examined in [Cepni, Guney, and Swanson \(2019\)](#), who extract latent global factors from a large-scale multi-country dataset, [Bulligan, Marcellino, and Venditti \(2015\)](#), who utilize predictor selection methods coupled with latent factor estimation for forecasting Italian GDP, [Schumacher \(2007\)](#), who forecasts German GDP using latent factors, [Marcellino, Porqueddu, and Venditti \(2016\)](#), who specify a mixed-frequency dynamic factor model and investigate business cycles in the euro area, and by [Andreou, Ghysels, and Kourtellos \(2013\)](#), who utilize the Mixed Data Sampling (MIDAS) framework developed in [Ghysels, Sinko, and Valkanov \(2007\)](#) to include daily data when forecasting lower frequency macroeconomic variables.² In all of these papers, models are specified that include latent factors which can be interpreted as business conditions indexes.

Examples of economic uncertainty indexes are discussed in [Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry \(2018\)](#), [Baker, Bloom, and Davis \(2016\)](#), [Carriero, Clark, and Marcellino \(2016\)](#), [Jo and Sekkel \(2017\)](#), and [Jurado, Ludvigson, and Ng \(2015\)](#). In all of these papers, forecast errors are used to construct their uncertainty indexes. In a related

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

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

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 and economic uncertainty indexes, and assess their usefulness for predicting macroeconomic variables. Our indexes are extracted from a large dimensional macroeconomic dataset using machine learning methods, and our objective is to add to the literature that proposes new methods for constructing business conditions and economic uncertainty indexes, and to add to the literature that empirically assesses the usefulness of these types of indexes. From a methodological perspective, we construct various business conditions indexes (BCIs) 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 BCIs are subsequently used as explanatory variables in econometric forecasting models. Forecasts errors from these models are then used to construct economic uncertainty indexes (EUIs), as in Jurado et al. (2015).³ With these new BCIs and EUIs in hand, we undertake a comprehensive set of forecasting experiments to assess their marginal predictive content. In particular, we predict 14 variables chosen from 8 different groups of macroeconomic variables, covering areas ranging from banking and finance to employment.⁴

Before continuing, it is worth stressing that the latent factors (i.e., BCIs) and EUIs ex-

³See also Ludvigson and Ng (2009) and Gu, Kelly, and Xiu (2020) for a discussion of related research 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.

⁴One 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.

amined 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 as additional predictors. 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 *both* business conditions *and* uncertainty indexes are *jointly* added to standard autoregressive forecasting models of the variety widely utilized in government and industry.

We would be remiss if we did not stress that 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 indexes are constructed by extracting latent factors from economic variables. Our paper is also closely related to [Jurado et al. \(2015\)](#), whose general methodology for constructing uncertainty measures 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” BCIs rather than constructing a BCI 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. Additionally, we build 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 key findings can be summarized as follows. First, the majority of our business condi-

tions and economic uncertainty indexes contain marginal predictive content for the 14 macroeconomic variables that we examine, at forecast horizons ranging from 1-month to 1-year ahead. This result obtains, regardless of whether BCIs or EUIs are added to our benchmark econometric models. However, the overall “winners” in terms of mean square forecast error (MSFE) are models that contain both business conditions and economic uncertainty indexes. One reason for this finding is that our BCIs contain substantial marginal predictive content that is useful for short-term forecasting, while our EUIs primarily contain marginal predictive content that is useful for long-term forecasting. This is consistent with the idea that economic uncertainty, as measured by our EUIs, is a good indicator of long-term economic activity, while near-term business conditions, as measured by our BCIs, are good indicators of short-term economic activity.

Second, many of our new business conditions and economic uncertainty indexes outperform the indexes discussed in [Jurado et al. \(2015\)](#). Three reasons for this success include our use of grouped variables, our use of new factor-lasso methods when applying shrinkage to \mathbf{X} , and our use of real-time parameter updating. Note that our use of grouped variables is predicated to some extent on the comment in [Jurado et al. \(2015\)](#), where it is noted that “Our measure of uncertainty conveniently aggregates uncertainty in the economy derived from all sources into one summary statistic. In some cases, it may be useful to construct subindices”. The 8 groups of variables that we specify are defined as follows: output and income, labor market, housing, consumption and inventories, money and credit, interest and exchange rates, prices, and stock market. Note also that [Jurado et al. \(2015\)](#) do not carry out forecasting experiments, while our main focus is on prediction. In addition, our dataset, while very similar to theirs, includes only 122 variables while their dataset includes 132 variables. Finally, as mentioned above, our estimation techniques are based on rolling windows with key parameters updated at each point in time, while their reported results are primarily based on an in-sample analysis of the data.

Third, we show that when PCA is used to extract latent factors that are the building blocks of our BCIs and EUIs, factors obtained using a standard eigenvalue-eigenvector de-

composition of the correlation matrix of \mathbf{X} can be easily improved upon, in the sense that superior predictions can easily be obtained using variants of these standard factors. One way in which this can be done is by adding additional variables, say \mathbf{E}^* , to the original set of factors extracted using PCA, say $\widehat{\mathbf{F}}$, when constructing BCIs. Here, \mathbf{E}^* are variables selected via use of lasso regression. In this lasso regression, the dependent variable is a vector of residuals from the regression of the variable that we want to predict on $\widehat{\mathbf{F}}$, and the predictor variables are \mathbf{E} , where $\mathbf{E} = \mathbf{X} - \widehat{\mathbf{X}}$, $\widehat{\mathbf{X}} = \widehat{\mathbf{F}}\widehat{\boldsymbol{\Lambda}}$, and $\widehat{\boldsymbol{\Lambda}}$ is the estimated factor loading matrix. This is a variant of the so-called factor-lasso method. This approach to estimating our BCIs and EUIs draws on important results in the high-dimensional data analysis literature discussed in [Kneip and Sarda \(2011\)](#), [Hansen and Liao \(2019\)](#), and [Fan, Ke, and Wang \(2020\)](#), and is predicated on the fact that the true unknown data generating process (DGP) describing our dataset may not be well approximated by a sparse model assumption, which holds that most of the explanatory power of a set of predictors concentrates within a small number of variables. Indeed, such methods may perform quite poorly if the true model is “dense”, such as when the true DGP follows a factor model, as in this case many variables “load” on the underlying latent factors. Moreover, screening methods tend to contain too many variables when covariates are strongly correlated. Needless to say, converse arguments may be made in favor of pre-screening using methods such as the lasso, and against the use of standard factor model setups. Building on these ideas, [Hansen and Liao \(2019\)](#) note that since there are instances where both factors and disturbances may be related to a target variable of interest, factor-lasso methods of the variety utilized in this paper may prove useful in empirical applications.

Indexes constructed using the factor-lasso method described above dominate in many of our forecasting experiments, when comparing predictions using MSFEs. We argue that the impressive performance that we observe is indeed due to the fact that there is a high degree of correlation across the variables in X . In closing, it is worth stressing that this approach is a hybrid one, in which lasso regression type variable selection (which is supervised) is combined with PCA (which is unsupervised), and most of the other successful methods for constructing

BCIs and EUIs that are examined in this paper also combine lasso variable selection with PCA.

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 BCIs 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 an online appendix available at “<https://econweb.rutgers.edu/nswanson/papers.htm>”.

2 Business Conditions and Economic Uncertainty Indexes

All of the business conditions indexes that are proposed in this paper are based on the application of principal components analysis (PCA) and the least absolute shrinkage operator (lasso). Additionally, all of the economic uncertainty indexes that we propose utilize our BCIs in a first step, in order to construct predictions, from which forecast errors are collected and used to construct our EUIs.⁵ In the remainder of this section, we briefly summarize PCA and lasso methodology, recap the EUI methodology that we use, and explain how to construct our various proposed BCIs and EUIs.

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

⁵As a benchmark against which to compare our economic uncertainty indexes, we also investigate the predictive content of the EUIs proposed in Jurado et al. (2015).

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

$$\mathbf{X} = \mathbf{F}\boldsymbol{\Lambda} + \mathbf{E}, \quad (1)$$

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, $\boldsymbol{\Lambda}$ 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:

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

Here, λ_{ij} is an element of the factor loading matrix, and e is an uncorrelated zero-mean disturbance term. A standard approach to estimating \mathbf{F} involves carrying out PCA. Namely, carry out an eigenvalue-eigenvector decomposition of the correlation matrix estimator of (\mathbf{XX}') , yielding r estimated eigenvalues, $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_r$, and corresponding estimated eigenvectors, $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r$. The first r principal components are estimated as $\hat{F}_j = \mathbf{X}\hat{\theta}'_j, j = 1, \dots, r$.

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:

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

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

In the sequel, procedures utilizing PCA and the lasso are used to construct our business conditions indexes. Prior to discussing these procedures, however, we first provide an overview of the methodology utilized to construct economic uncertainty indexes.

2.2 Economic Uncertainty Indexes

Roughly speaking, our uncertainty indexes are constructed using forecast error variances from factor augmented autoregressions, where the factors are our BCIs (as discussed in Section 2.3). More specifically, our uncertainty indexes are aggregates of 14 individual uncertainty indexes. These 14 indexes correspond to the 14 target variables that we forecast, and are in this sense “whole economy” uncertainty indexes. In total, 10 such uncertainty indexes are constructed. This is because we examine 10 different types of business condition indexes (as discussed in Section 2.3), and each of these varieties of indexes are in turn used in the first stage prediction regressions from which the forecast error variances mentioned above are extracted.

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 follows:

$$MI = \sum_{j=1}^{14} w_j U_{jt}^y(h). \quad (4)$$

We assume that each weight, w_j , is equal to $1/14$.^{6,7} 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 $i = 1, \dots, N$:

$$\mathbf{X}_{it} = \Lambda'_i \mathbf{F}_t + e_{it} \quad (5)$$

where \mathbf{F}_t is an $r_F \times 1$ vector of latent common factors, Λ_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:

$$y_{jt} = \phi_j(L)y_{jt-1} + \gamma_j(L)\hat{\mathbf{F}}_{t-1} + \beta_j(L)\mathbf{W}_{t-1} + v_{jt}, \quad (6)$$

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 $\mathcal{Z}_t = (\mathbf{Z}'_t, \dots, \mathbf{Z}'_{t-q+1})$. Also, let $Y_{jt} = (y_{jt}, y_{jt-1}, \dots, y_{jt-q+1})'$. We can now write the forecasting equation for factors and target variables as follows:

⁶Different 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.

⁷A 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 construct a global uncertainty index with individual uncertainty indexes for all 122 variables in our dataset (see the subsequent section for further discussion).

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

$$\mathcal{Y}_{jt} = \Phi_j^Y \mathcal{Y}_{jt-1} + \mathcal{V}_{jt}^Y,$$
(7)

where Λ'_j and Φ_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:

$$U_{jt}^y(h) = \sqrt{1_j' \Omega_{jt}^y(h) 1_j}. \quad (8)$$

Note that when $h = 1$,

$$\Omega_{jt}^y(1) = E_t(\mathcal{V}_{jt+1}^y \mathcal{V}_{jt+1}^{y'}). \quad (9)$$

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

$$\Omega_{jt}^y(h) = \Phi_{jt}^y \Omega_{jt}^y(h-1) \Phi_{jt}^{y'} + E_t(\mathcal{V}_{jt+h}^y \mathcal{V}_{jt+h}^{y'}). \quad (10)$$

In our experiments, we update the coefficient matrix Φ_{jt}^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, σ_{jt}^y , σ_{lt}^F , and σ_{gt}^W , respectively. Namely, for factor F_t (and analogously for W_t), we assume that:

$$\mathbf{F}_t = \Phi^F \mathbf{F}_{t-1} + \mathbf{v}_t^F. \quad (11)$$

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:

$$\log(\sigma_t^F)^2 = \alpha^F + \beta^F \log(\sigma_{t-1}^F)^2 + \tau^F \eta_t^F, \eta_t^F \stackrel{iid}{\sim} N(0, 1). \quad (12)$$

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

$$E_t(\sigma_{t+h}^F)^2 = \exp \left[\alpha^F \sum_{s=0}^{h-1} (\beta^F)^s + \frac{(\tau^F)^2}{2} \sum_{s=0}^{h-1} (\beta^F)^{2s} + (\beta^F)^h \log(\sigma_t^F)^2 \right] \quad (13)$$

The stochastic volatility parameters, α , β , and τ , 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

$$\Omega_t^F(h) = (\Phi^F) \Omega_t^F(h-1) \Phi^{F'} + E_t(v_{t+h}^F v_{t+h}^{F'}), \quad (14)$$

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

$$U_t^F(h) = \sqrt{\mathbf{1}_F' \Omega_t^F(h) \mathbf{1}_F}, \quad (15)$$

where $\mathbf{1}_F$ is an appropriate selection vector.

In most of our experiments, we use a simplified framework where only business conditions indexes constructed using PCA and the lasso are utilized in this framework. Throughout this paper, these business conditions indexes have been defined to be the latent factors extracted from \mathbf{X} , and are called \mathbf{F}_t .⁸ In this case, we have that:

$$y_{jt+1} = \phi_j^y y_{jt} + \gamma_j^F \hat{\mathbf{F}}_t + v_{jt+1}^y. \quad (16)$$

⁸The \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:

$$\log(\sigma_{jt+1}^y)^2 = \alpha_j^y + \beta_j^y \log(\sigma_{jt}^y)^2 + \tau_j^y \eta_{jt+1}, \eta_{jt+1} \stackrel{iid}{\sim} N(0, 1). \quad (17)$$

Again, the stochastic volatility parameters, α_j , β_j , and τ_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 $h = 2$,

$$V_{jt+2}^y = v_{jt+2}^y + \phi_j^y V_{jt+1}^y + \gamma_j^F V_{t+1}^F, \quad (18)$$

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

$$V_{jt+3}^y = v_{jt+3}^y + \phi_j^y V_{jt+2}^y + \gamma_j^F V_{t+2}^F \quad (19)$$

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 uncertainty indexes 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 indexes”. Thus, by specifying the various procedures that we utilize to construct our business conditions indexes, we are also providing the inputs for the construction of our uncertainty indexes. In the next section we explain how

we specify and estimate \mathbf{F}_t and \mathbf{W}_t .⁹

2.3 Business Conditions Indexes

As discussed above, the business conditions indexes analyzed in this paper, and used in the construction of our uncertainty indexes, are latent factors extracted from \mathbf{X} . These measures, denoted by \mathbf{F}_t and \mathbf{W}_t , 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 proxies for our business conditions indexes. These observable variables are selected anew for each target variable y_j , using predictor dataset \mathbf{X}^- . The tuning parameter, λ , in equation (3) is determined by applying *hv*-block cross validation, as discussed in [Racine \(2000\)](#). We set λ to be an element of 100 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:

$$y_{jt+1} = \alpha + x_{it}\beta_i + \epsilon_{jt+1}, \quad (20)$$

⁹As mentioned above, in all but one of the procedures outlined below, we specify and estimate only \mathbf{F}_t .

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, $R = \frac{1}{2}2T$, and the out-of-sample-period $P = \frac{1}{2}T$.¹⁰ 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 $k = 1, \dots, 8$.¹¹ We forecast each variable in a given group, using all other variables in the same group. Namely, we fit regressions of the following form:

$$x_{it+1} = \alpha + X'_{kt}\beta_j + \epsilon_{it+1}, i = 1, \dots, n^*, \quad (21)$$

where the coefficients, β_j , and the stochastic disturbances, ϵ_{it} are conformably defined. Then, for each variable that we forecast in each group, we choose top 5 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 $P = \frac{1}{2}T$. In particular, we pick the 5 most precisely predicted variables, from the group X_j . 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 indexes, \mathbf{F}_t , for each target variable.¹² Finally, if the number of selected business conditions indexes for a given y_j is more than 3, 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 indexes, \mathbf{F}_t . Otherwise, we use the variables selected via lasso regression

¹⁰Our choice of 20 predictors is predicated on our finding that lasso selects at most 20 variables in our experiments.

¹¹The 8 groups into which we partition our predictor dataset are defined in Section 4.

¹²Note 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 .

as our “final” set of 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 of the form:

$$y_{jt+1} = \alpha + x_{it}\beta_i + \epsilon_{ijt+1}, i = 1, \dots, n^*, \quad (22)$$

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 .¹³ Now, for each y_j predictor set, apply PCA, yielding a unique set of business conditions indexes, \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 3, 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. business conditions indexes), \mathbf{F}_t , for each y_j . In order to achieve further parsimony, select a final set of factors (i.e., business conditions indexes) for each y_j by again carrying our variable selection using lasso regression on all latent factors across all 8 groups. If the number of selected business conditions indexes is more than 3, carry out PCA once again, but this time on the factors chosen in the previous steps of this procedure, yielding a final set of business conditions indexes, \mathbf{F}_t .¹⁴ Otherwise, we use the factors selected via lasso regression as our “final” set of BCIs.

8. Group Factors II (GF2): This is a simplified variant of GF1. First, carry variable

¹³An exception to this is the case of predicting S&P500. For this variable, there are only 39 predictors, due to the fact that the group in which the S&P500 is contained has only 5 variables in it, and S&P500 cannot itself be used as a predictor in our setup.

¹⁴Interestingly, it was found, via experiments using the first half of the data as a “training dataset”, that this sort of “double” lasso-PCA shrinkAGE yields superior predictions, when a “fixed” rule for selecting the number of factors and for selecting the tuning parameter in our lasso regression is implemented, as is done in our analysis. Needless to say, the same sort of result might be more efficiently achieved by adjusting increasing the shrinkage implied by our choice of tuning parameter, for example, on a case by case and variable by variable basis.

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 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 et al. \(2020\)](#). 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 $\widehat{\mathbf{X}}^- = \widehat{\mathbf{F}}\widehat{\Lambda}$, $\widehat{\mathbf{F}}$ is the estimated factor matrix, and $\widehat{\Lambda}$ 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 indexes that we use when predicting y_j , and when constructing EUIs.

10. Jurado, Ludvigson and Ng Indexes (JLN): This method replicates the approach taken used by [Jurado et al. \(2015\)](#) for index construction. Of note is that there are essentially two steps involved. In a first step, factors for use in prediction equations are estimated. In a second step, as outlined in the previous sub-section, uncertainty indexes are constructed using the forecast errors from predictive regressions in which these factors serve as predictors. In our analysis, we examine both the factors used in their predictive regressions (we call these business conditions indexes), and their uncertainty indexes based on predictive regressions of all 122 variables in our dataset on the factors. Indeed, we run three types of experiments using the JLN method. In a first experiment, we construct predictions using only their business conditions indexes as predictors. In a second set of experiments we construct predictions using only their uncertainty indexes as predictors. Finally, in a third set of experiments, we

construct predictions using both types of indexes as predictors.

In order to construct what we term the JLN business conditions indexes, we forecast all 122 variables in our dataset, $X_{it}, i = 1, \dots, N$ using a AR model with 4 lags (as done in their paper). We then apply PCA in order to obtain our factors, $\widehat{\mathbf{F}_t}$. Again following their approach, we then construct an additional predictor set, \mathbf{W}_t . This predictor set contains the square of the first component of $\widehat{\mathbf{F}_t}$, as well as the first factor extracted by applying PCA to $X_{it}^2, i = 1, \dots, N$. Finally, the consolidated dataset, $(\widehat{\mathbf{F}_t}, \mathbf{W}_t)$ is pruned by applying hard-thresholding using t-statistics on predictive regressions, for a given target variable, y_j .¹⁵ Note that when applying PCA using their approach, the maximum number of factors is 20 and the PC_{p2} of [Bai and Ng \(2002\)](#) is used to determine the number of factors.

Summarizing the discussion in this section, note that we outline methods for constructing 10 different varieties of business conditions indexes (BCIs). These are defined to be the factors extracted using the various methods ranging from PCA and lasso to GP1 and JLN. Additionally, all of these methods form a first step in the construction of the economic uncertainty indexes (EUIs) discussed in the previous section. In this sense we have also specified 10 different varieties of economic uncertainty indexes. In the next section, we outline the experiments that we carry out in order to assess the marginal predictive content of these indexes.

3 Forecasting Models and Predictive Accuracy Testing

In order to examine the marginal predictive content of our BCIs and EUIs, we carry out real-time h -month ahead predictions using an updated 122 variable variant of the monthly dataset examined in [Jurado et al. \(2015\)](#), as discussed in Section 4, and using a number of forecasting models, for $h = 1, 3, 6$, and 12. Predictions are made for 14 different monthly U.S. variables, including: 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),

¹⁵See [Jurado et al. \(2015\)](#) for complete details.

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).¹⁶ In addition, the 8 groups of variables into which our 122 variable dataset (called \mathbf{X} above) are sub-divided 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. From each group, we choose 1 to 3 variables, yielding our 14 target variables that we forecast (see Section 4 for further details).

In all of our experiments, we construct predictions using the following models:

Autoregressive Model (AR): 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 target forecast variable, h denotes the forecast horizon, $\hat{\beta}_h(L)$ is a finite order lag polynomial, ϵ_{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 indexes (AR+BCI): The AR model augmented with BCIs is:

$$y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_h(L)\widehat{BCI}_t + \epsilon_{t+h}, \quad (24)$$

where $\beta_h(L)$ and $\gamma_h(L)$ are finite order lag polynomials, \widehat{BCI}_t is a set of latent factors, constructed as discussed above.

AR model with economic uncertainty indexes (AR+EUI): The AR model augmented with EUIs is:

$$y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_h(L)\widehat{EUI}_t + \epsilon_{t+h}, \quad (25)$$

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

AR model with business conditions and economic uncertainty indexes (AR+BCI+EUI):

¹⁶See Table 1 for details on the data transformations used.

The AR model augmented with BCIs and EUIs is:

$$y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_{1h}(L)\widehat{\mathbf{BCI}_t} + \gamma_{2h}(L)\widehat{\mathbf{EUI}_t} + \epsilon_{t+h}, \quad (26)$$

where $\beta_h(L)$, $\gamma_{1h}(L)$, and $\gamma_{2h}(L)$ are finite order lag polynomials, and $(\widehat{\mathbf{BCI}_t}, \widehat{\mathbf{EUI}_t})$ is a set of consistently estimated factors and uncertainty indexes, constructed 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. 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 prediction period for $h = 1$ is 2004:6-2021:6 (i.e., $P = 349 - 144 = 205$ observations). All BCIs 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 prediction. The in-sample estimation periods used when constructing our $h = 3$, 6, 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.¹⁷ 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 \(2011\)](#).

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.¹⁸

As mentioned above, the 8 groups of variables into which our 122 variable dataset (called \mathbf{X} above) is sub-divided 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

¹⁷In this paper, we report test results for the Wald version of this test statistic (see [Giacomini and White \(2006\)](#) for further details).

¹⁸We use fully revised FRED data in our subsequent empirical analysis. In this sense, our results should be viewed as 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 BCIs 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 \(2017\)](#).

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).

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 indexes (BCIs) 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 BCIs, 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 as PCA, Lasso, LPCA, AP, GP1, GP2, GF1, GF2, EPCA, and JLN, as discussed in Section 2.3. Our BCIs are then used to construct 10 global economic uncertainty indexes (EUIs).¹⁹ In summary, 4 samples of $h=1, 3, 6$, and 12-step ahead predictions are constructed for the sample period 2004:6-2021:6, for each target variable. The 4 different samples of predictions are made using forecasting models denoted as: AR(SIC), AR(SIC)+BCI, AR(SIC)+EUI, and AR(SIC)+BCI+EUI, as summarized in Table 2, and detailed in Section 3. Please also 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 1-9 and Figures 1-3. Additional experimental results that we discuss, but that are not reported here, are contained in a supplemental appendix available at <https://econweb.rutgers.edu/nswanson/papers.htm>.

¹⁹Note that the way that we construct EUIs involves utilizing 1-step ahead predictions of the BCIs, using the methodology discussed in Section 2.2.

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, 10 “methods” are used to construct our EUIs, 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 BCIs using the Lasso yields EUIs that are included as predictors in our MSFE-best models for 11 of 14 variables, when $h = 1$. However, this finding is based solely on the examination of the point MSFEs reported in the table. As none of these MSFE entries are starred, there is nothing to choose between the Lasso and our benchmark AR(SIC) model, at a 5% level of significance (based on the application of [Giacomini and White \(2006\)](#) conditional predictive ability tests). Interestingly, though, the AR(SIC) model never yields the lowest MSFE in our experiments, regardless of forecast horizon. Moreover, the method yielding the MSFE-best EUI index varies when $h > 1$. For example, for $h = 12$ we see that the use of PCA, Lasso, LPCA, and GF1 all lead to superior forecasting using EUIs, with the “winner” depending on which target variable is being predicted. Moreover, various MSFE entries for $h = 3, 6$, and 12 are starred, indicating that there are various EUI-type forecasting models that are statistically superior to the AR(SIC) benchmark.

Second, 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)+BCI and AR(SIC) forecasting models are instead compared, and in Table 5, AR(SIC)+EUI+BCI and AR(SIC) forecasting models are compared. Turning first to the results collected in Table 4, note that models with BCIs yield statistically superior predictions, when compared with predictions made using our AR(SIC) benchmark, for all forecast horizons. Also, all 10 methods used to construct our BCIs result in statistically superior predictions for some of our 14 target variables. However, use of the EPCA method leads quite decisively to the highest number of “wins”. For example, for $h = 1$ and $h = 3$, use of EPCA when constructing BCIs leads to statistically superior

performance (relative to the AR(SIC)) model for 10 of 14 and 7 of 14 variables, at a 10% significance level, respectively. The second best method for constructing BCIs is the Lasso (see results in Tables 3 and 4).

In addition, it is noteworthy that the incidence of rejections of the null hypothesis of equal predictive accuracy (when comparing AR(SIC)+BCI model forecasts with AR(SIC) benchmark forecasts) increases in the forecast horizon, h (see Table 3). Conversely, in Table 4, where AR(SIC)+EUI models are compared with the AR(SIC) benchmark, the incidence of rejections decreases in the forecast horizon (see Table 4). Moreover, comparison of the results in these two tables indicates that there are many more rejections, across all forecast horizons, for AR(SIC)+BCI models than for AR(SIC)+EUI models. However, the number of test rejections for AR(SIC)+EUI models in Table 3 is greater than the number of rejections for AR(SIC)+BCI models in Table 4, for $h = 12$. Thus, if one wishes only to use a single type of index for forecasting, our BCI indexes are superior at shorter forecast horizons. However, our EUI indexes yield a greater number of MSFE “wins” at our longest forecast horizon.²⁰ This finding is consistent with the idea that economic uncertainty, as measured by our EUIs, is a good indicator of long-term economic activity, while near-term business conditions, as measured by our BCIs, are good indicators of near-term economic activity.

Third, inspection of Table 5, which contains MSFEs comparing AR(SIC)+EUI+BCI with AR(SIC) forecasting models reveals that indexes constructed using EPCA and Lasso methods are again important components of our MSFE-best forecasting models. Importantly, MSFEs in this table are lower than those reported in Tables 3 and 4, in most cases. This indicates that the use of both EUIs and BCIs in our forecasting models is preferable to the use of either EUIs or BCIs. This result is interesting, as it suggests that there is quantifiable predictive information in both “level” business conditions indexes, such as those proposed by [Aruoba et al. \(2009a\)](#) and economic uncertainty indexes, such as those proposed by [Jurado et al. \(2015\)](#).²¹ Moreover, it is of note that the JLN method reported in Tables 3-5 only yields “MSFE-best”

²⁰However as discussed below, models that include both indexes are our overall MSFE-best models, in almost all cases.

²¹Our pseudo real-time forecasting experiments build on the work of these two papers, although neither of the two papers contain results based on prediction experiments.

predictions in 1 case, across all 168 target variable, forecast horizons, and prediction model permutations.²² Thus, while the methodology proposed in Jurado et al. (2015) is clearly very useful for the construction of EUIs, our alternative BCI construction methods, coupled with our use of grouped data and parameter updating, tend to result in superior performing EUIs than those discussed in their paper. In particular, the use of our alternative BCI construction methods in addition to our approach of constructing EUIs by aggregating subindices based on grouped variables rather than aggregating across all individual variables yields EUIs with superior marginal predictive content.

Fourth, note that Table 6 lists the top 3 MSFE-best forecasting models and BCI (factor) extraction method combinations, for all target variables and forecast horizons. In this sense, the very best results from Tables 3-5 are reported in this table. Summarizing, forecasting models are listed, along with BCI extraction methods (in parentheses). Starred entries indicate models yielding statistically superior predictions, when compared with our AR(SIC) benchmark. Inspection of the models ranked as number “1” indicates that 11 of 14 variables ($h = 1$), 12 of 14 variables ($h = 3$), 12 of 14 variables ($h = 6$), and 12 of 14 variables ($h = 12$) that are MSFE-best are predicted using our AR(SIC)+EUI+BCI model. This result confirms our findings based on the examination of the results in Table 5. Namely, our best forecasting models include both EUIs and BCIs.²³

Interestingly, while using EPCA and Lasso methods for constructing BCIs and EUIs still results in many MSFE-best models, a number of other methods are also clearly useful. In particular, comparing only the models ranked number “1” indicates that for our 14 target variables, EPCA, Lasso, PCA, GF1, GF2, GP1, GP2, and AR(SIC) yield the most predictively useful BCIs and EUIs in: 8, 4, 2, 0, 0, 0, 0, and 0 cases for $h = 1$, respectively; and in 5, 2, 4, 1, 1, 0, 1, and 0 cases for $h = 3$, respectively; and in 8, 4, 1, 1, 0, 0, 0, and 0 cases for $h = 6$, respectively; and in 7, 4, 0, 1, 0, 1, 0, and 1 cases for $h = 12$, respectively. Evidently, EPCA and the Lasso win in 12, 7, 12, and 11 of 14 cases, for forecast horizons $h = 1, 3, 6$, and 12 , respectively. The rest of the wins (i.e. 2, 7, 2, and 3 of 14 cases, for forecast horizons $h = 1, 3,$

²²This result can also be seen by examining the results in Table 6.

²³Across all 168 permutations in this table, the AR(SIC) benchmark model is MSFE-best in only one case.

6, and 12, respectively) are spread across the other methods, with each of PCA, GF1, GF2, GP1, GP2 and AR(SIC) occasionally providing superior indexes.

Fifth, note that the performance of our factor extraction methods is summarized in Tables 7-9, where MSFE-best methods from comparisons of our AR(SIC)+EUI and AR(SIC) forecasting models (Table 7), our AR(SIC)+BCI and AR(SIC) forecasting models (Table 8), and our AR(SIC)+EUI+BCI and AR(SIC) forecasting models (Table 9) are compared. Interestingly, inspection of Table 7 indicates that EPCA and Lasso methods for constructing BCIs are not dominant when only economic uncertainty indexes are used in our prediction models. The dominance of these methods only arises when either BCIs or BCIs+EUIs are used for forecasting our target variables (see results in Tables 8 and 9).

Finally, we graphically depict our EUIs, two of our target housing variables (HOUST and PERMIT), and the widely analyzed VIX series in Figures 1-3. In Figure 1, we see that indexes constructed for use in our 1-step ahead forecasting models largely follow our housing variables, even during the 2008 financial crisis, and the Covid-19 period that began in 2020. The story when $h = 12$ is somewhat different, as our EUIs do not as clearly follow our housing variables (see Figure 2). However, in Figure 3 it is quite apparent that our EUIs are rather closely related to the VIX, which is promising, and to some extent explains why they are quite useful for forecasting.²⁴

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. In this sense, we have new evidence of the usefulness of the types of indexes discussed in [Aruoba et al. \(2009a\)](#) and [Jurado et al. \(2015\)](#). Interestingly, when forecasting using either EUIs or BCIs (but not both), we have evidence that BCIs are particularly useful for short-term prediction, while EUIs are primarily useful for long-term prediction. Additionally, we find that EPCA, and to a lesser extent Lasso, are particularly useful methods for selecting variables from which to construct business conditions indexes and from which to form predictions used in the construction of economic uncertainty indexes. Thus, both lasso and factor-lasso methods

²⁴[Jurado et al. \(2015\)](#) discuss reasons why EUIs such as those analyzed in our paper might be expected to exhibit somewhat different behavior than financial market uncertainty indexes such as the VXO and VIX.

dominate in our experiments.²⁵

6 Concluding Remarks

In this paper, we propose a number of new business conditions indexes (BCIs) and economic uncertainty indexes (EUIs). Our BCIs are constructed using variations on extant lasso regression and principal components analysis methods 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 BCIs 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 BCIs and EUIs contain marginal predictive content, and forecasting models that utilize both BCIs and EUIs are superior to a variety of alternatives. Additionally, we find that our very best indexes are based on lasso and factor-lasso methods, with the latter involving hybrid estimation that incorporates supervised and as well as unsupervised learning. A number of outstanding issues are left to future research. It remains, for example, to carry out further empirical analysis that builds on our pseudo real-time experiments by analyzing real-time datasets. It also remains to carry out empirical and theoretical analysis of high-dimensional and mixed frequency models designed to estimate the types of BCIs and EUIs discussed in this paper.

²⁵Lasso implements lasso regression, while our overall “winner”, EPCA, utilizes both lasso regression and PCA.

References

- Andreou, E., Ghysels, E., & Kourtellos, 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. (2009a). 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.
- Frale, 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.
- Kneip, A., & Sarda, P. (2011). Factor models and variable selection in high-dimensional regression analysis. *The Annals of Statistics*, 39(5), 2410–2447.
- 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	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	$\Delta \log(y_t)$
Group6: Interest and Exchange Rates	10-Year Government Treasury Bond Rate	R10	y_t
Group7: Prices	PPI (finished goods)	PPI	$\Delta \log(y_t)$
Group7: Prices	CPI (all items)	CPI	$\Delta \log(y_t)$
Group7: Prices	Personal Consumption Expenditures (chain index)	PCECI	$\Delta \log(y_t)$
Group8: Stock Market	S&P Common Stock Price Index (composite)	S&P500	$\Delta \log(y_t)$

* 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) + BCI	AR(SIC) model augmented with business conditions indexes extracted from one of the 10 methods outlined in Section 2.3. $y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_h(L)\widehat{BCI}_t + \epsilon_{t+h}$
AR(SIC) + EUI	AR(SIC) model augmented with a global economic uncertainty index constructed by applying the methods discussed in Section 2.2 using BCIs constructed using one of the 10 methods outlined in Section 2.3. $y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_h(L)\widehat{EUI}_t + \epsilon_{t+h}$
AR(SIC) + BCI + EUI	AR(SIC) model augmented with both BCIs and an EUI, with indexes constructed using the methods outlined in Sections 2.2 and 2.3. $y_{t+h} = \alpha + \beta_h(L)y_t + \gamma_{1h}(L)\widehat{BCI}_t + \gamma_{2h}(L)\widehat{EUI}_t + \epsilon_{t+h}$

* 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: Comparison of the Predictive Accuracy of AR(SIC) Models with AR(SIC) + EUI Models*

		Factor Extraction Method used to Construct Business Conditions Indexes									
Target Variable	PCA	Lasso	LPCA	AP	GP1	GP2	GF1	GF2	EPCA	JLN	
h=1	RPI	0.863	0.83	0.849	0.866	0.823	0.885	0.862	0.894	0.867	0.921 *
	INDPRO	0.88	0.817	0.84	0.859	0.824	0.904	0.866	0.926	0.88	0.931
	UNRATE	0.831	0.818	0.829	0.855	0.817	0.866	0.861	0.861	0.876	0.943
	CLAIMS	0.956	0.935	0.945	0.967	0.941	0.976	0.975	0.97	0.962	0.978
	HOUST	0.937	0.901	0.911	0.931	0.92	0.931	0.92	0.933	0.935	0.961
	PERMIT	0.944	0.905	0.917	0.931	0.928	0.937	0.923	0.951	0.94	0.96
	RCON	0.769	0.655	0.703	0.747	0.677	0.774	0.74	0.822	0.743	0.788
	MTS	0.892	0.785	0.829	0.873	0.788	0.914	0.866	0.954	0.865	0.923
	M2	0.887	0.816	0.851	0.889	0.82	0.911	0.891	0.936	0.873	0.915
	R10	0.985	0.942	0.947	0.973	0.954	0.976	0.962	0.962	0.985	0.975
h=3	PPI	0.961	0.939	0.936	0.956	0.943	0.957	0.951	0.949	0.968	0.961
	CPI	0.975	0.944	0.949	0.971	0.953	0.967	0.96	0.959	0.978	0.969
	PCECI	0.977	0.946	0.953	0.971	0.956	0.968	0.961	0.96	0.98	0.97
	S&P500	0.928	0.869	0.892	0.918	0.89	0.935	0.914	0.937	0.933	0.947
	RPI	0.914	0.949	0.952	0.96	0.961	0.906	0.979	0.905	0.971	0.927
	INDPRO	0.818	0.927	0.9	0.903	0.923	0.864	0.98	0.825	0.894	0.904
	UNRATE	0.723	0.896	0.857	0.792	0.866	0.791	0.987	0.744	0.823	0.883
	CLAIMS	0.749	0.885	0.882	0.755	0.888	0.779	0.98	0.735	0.877	0.889
	HOUST	0.835	0.728 ***	0.872	0.894	0.849	0.738 **	0.963	0.758 *	0.875	0.94
	PERMIT	0.868	0.726 **	0.888	0.905	0.843	0.753 ***	0.987	0.798 *	0.889	0.955
h=6	RCON	0.675	0.889	0.872	0.78	0.9	0.818	0.941	0.727	0.853	0.792
	MTS	0.774	0.951	0.901	0.9	0.944	0.857	0.987	0.801	0.896	0.89
	M2	0.873	0.946	0.939	0.91	0.941	0.914	0.96	0.868	0.94	0.931
	R10	0.977	0.951	0.956	0.977	0.943	0.957	0.939	0.937	0.982	0.985
	PPI	0.919	0.944 *	0.936 *	0.927 **	0.942	0.928	0.949	0.908	0.953	0.961 *
	CPI	0.957	0.938	0.934 *	0.956 *	0.939	0.946 *	0.922	0.924 *	0.967	0.969 **
	PCECI	0.961	0.943	0.939 *	0.957 **	0.945 *	0.95 *	0.927	0.929 *	0.972	0.969 **
	S&P500	0.969	0.925	0.961	0.996	0.963	0.95	0.93	0.957	0.988	0.955
	RPI	1.005	0.979	0.971	1.002	0.987	0.963	0.976	0.984	0.965	0.958
	INDPRO	0.994	0.947	0.949	0.997	0.953	0.994	0.951	0.976	0.947	0.921
h=12	UNRATE	0.915	0.985	0.979	0.962	0.981	0.934	0.979	0.974	0.967	0.922
	CLAIMS	0.946	0.971	0.966	0.941	0.979	0.911	0.984	0.965	0.967	0.933
	HOUST	0.928	0.922	0.932	0.937	0.849 **	0.781 ***	0.945 *	0.925 *	0.936	0.981
	PERMIT	0.885 *	0.902	0.902	0.909	0.853	0.784 ***	0.964	0.921	0.928	0.982 *
	RCON	0.99	0.917	0.884	0.969	0.962	0.967	0.958	0.978	0.925	0.766
	MTS	0.984	0.953	0.954	0.978	0.973	0.984	0.965	0.978	0.934	0.913
	M2	0.986 *	0.952 *	0.944	0.979 *	0.971 *	0.976 **	0.97	0.97 **	0.955	0.940
	R10	0.993	0.981	0.975	0.979	0.961	0.985	0.987	0.979	0.99	0.982
	PPI	0.985 ***	0.964 **	0.963 ***	0.979	0.966 **	0.985	0.916	0.97 **	0.98 ***	0.974 *
	CPI	0.986	0.964	0.967 *	0.978 *	0.964 *	0.981	0.865	0.972 **	0.984 ***	0.979 *
h=12	PCECI	0.987 *	0.964	0.967 **	0.978 **	0.967 *	0.978	0.872	0.973 **	0.985 ***	0.979
	S&P500	1.007	0.965	0.972	1.021	0.967	1.005	0.779	0.983	0.995 **	0.996
	RPI	0.97	0.909	0.901	0.987	0.979	0.984 *	0.949	0.96	0.933	0.996
	INDPRO	0.953	0.912	0.906	1.011	0.98	0.952	0.888	0.989	0.95	0.981
	UNRATE	0.962 **	0.879	0.875	0.988	0.991	0.946	0.938	0.982	0.915	0.974
	CLAIMS	0.987	0.873	0.872	0.992	0.996	0.958	0.983	0.972	0.943	0.969
	HOUST	0.787 **	0.974	0.965	0.906 *	0.917	0.917	0.908	0.941	0.949	0.954 **
	PERMIT	0.762 **	0.973	0.967	0.869 **	0.906	0.911	0.904	0.94	0.943	0.956
	RCON	0.945	0.835	0.836	0.982	0.99	0.838	0.977	0.988 *	0.859	0.904
	MTS	0.947	0.816	0.814	0.99	0.988	0.947	0.95	0.983	0.873	0.982
h=36	M2	0.977 *	0.903	0.900	0.992 *	0.969 **	0.942	0.953	0.967 *	0.927	0.977
	R10	0.986	0.953	0.95	0.984	0.957	1.02	0.882	0.978	0.971	0.984
	PPI	0.99 **	0.971 **	0.966	0.986	0.952	0.982	0.906	0.969 *	0.975 **	0.964
	CPI	0.992 ***	0.971 ***	0.963 **	0.99	0.949	0.992	0.907	0.967 *	0.98 ***	0.983
	PCECI	0.991 ***	0.969 ***	0.965 ***	0.988	0.954	0.99	0.905	0.971 **	0.972 ***	0.98
	S&P500	0.985	0.981	0.965	0.983	0.967	0.973	0.753	0.976	0.992	1.006

* Notes: See notes to Table 2. Tabulated entries are relative mean squared forecast error (MSFEs) for our 14 target variables, and for forecast horizons of h=1,3,6, 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 2004:6-2021:6, 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 & 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 3 for complete details.

Table 4: Comparison of the Predictive Accuracy of AR(SIC) Models with AR(SIC) + BCI Models*

Target Variable	Factor Extraction Method used to Construct Business Conditions Indexes										
	PCA	Lasso	LPCA	AP	GP1	GP2	GF1	GF2	EPCA	JLN	
RPI	0.699	0.756	0.756	0.732	0.738	0.713	0.756	0.72	0.61	0.777	
INDPRO	0.631	0.55 *	0.924 *	0.608	0.66 *	0.66	0.606 *	0.718	0.581 *	0.55 *	
UNRATE	0.551	0.524 **	0.545 *	0.548	0.547 *	0.556	0.542	0.68	0.54 *	0.573	
CLAIMS	0.554	0.648	0.648	0.679 *	0.677	0.621	0.56 *	0.534 *	0.532 *	0.688	
h=1	HOUTST	0.593 ***	0.543 ***	0.543 ***	0.526 ***	0.599 ***	0.572 ***	0.702 **	0.598 ***	0.524 ***	0.66 ***
	PERMIT	0.67 ***	0.888	0.888	0.705 **	0.946	0.674 **	0.753 *	0.765 **	0.658 ***	0.84 **
	RCON	0.575 *	0.96	0.96	0.676 *	0.956	0.711 *	0.953	0.874	0.532 *	0.92 **
	MTS	0.549 *	0.895	0.895	0.551	0.669 *	0.552	0.725	0.725	0.531 *	0.737
	M2	0.741	0.734	0.734	0.731	0.855	0.749	0.889	0.768	0.719	0.959
	R10	0.846	0.843	0.905	0.904	0.878	0.883	0.918	0.926	0.846	0.875
	PPI	0.708 **	0.986	0.986	0.915	0.979	0.732 **	0.936	0.814	0.699 **	0.805
	CPI	0.787 *	0.869 **	0.869 **	0.735 **	0.93	0.748 *	0.856 **	0.82 **	0.635 ***	0.92
	PCECI	0.744 **	0.807 ***	0.807 ***	0.83 *	0.7 **	0.742 **	0.813 **	0.803 **	0.643 ***	0.921
	S&P500	0.845	0.976	0.976	0.877	0.991	0.861 *	0.946	0.946	0.742	0.96
	RPI	0.801 **	0.83 *	0.83 *	0.802 **	0.831 *	0.82 **	0.844	0.806 **	0.714 **	0.891
	INDPRO	0.776	0.677	0.96	0.757	0.858	0.802	0.75	0.874	0.742	0.855
	UNRATE	0.672	0.578	0.676	0.658	0.642	0.714	0.673	0.776	0.585	0.716
h=3	CLAIMS	0.803	0.873	0.873	0.801	0.784 *	0.82	0.742	0.726	0.685	0.892
	HOUTST	0.598 ***	0.625 ***	0.625 ***	0.586 ***	0.695 ***	0.614 ***	0.689 ***	0.63 ***	0.58 ***	0.713 ***
	PERMIT	0.666 ***	0.795 **	0.795 **	0.695 **	0.88 *	0.683 **	0.726 **	0.716 ***	0.653 ***	0.857
	RCON	0.793	0.912	0.912	0.801	0.91	0.827	0.916	0.853	0.593	0.909
	MTS	0.833	0.967	0.967	0.765	0.926	0.841	0.889	0.889	0.8	0.934
	M2	0.803	0.823	0.823	0.819	0.93	0.814	0.944 *	0.824	0.791	0.964
	R10	0.919	0.842	0.992	0.941	0.896	0.922	0.903	0.993	0.919	0.925
	PPI	0.851 ***	1.011	1.011	0.976	0.99	0.906 **	0.979	0.94	0.861 **	0.929 **
	CPI	0.856	0.946 *	0.946 *	0.89	0.972	0.888	0.936 **	0.924 **	0.825 *	0.965
	PCECI	0.864 **	0.949	0.949	0.896 **	0.881	0.907	0.921	0.908	0.805 **	0.972
	S&P500	0.818	0.971	0.971	0.869 *	1.016	0.848	0.94	0.94	0.763 *	0.99
	RPI	0.866 **	0.877 **	0.877 **	0.865	0.856 *	0.889	0.853 **	0.895	0.784	0.96
	INDPRO	0.839	0.694 *	0.992	0.796	0.8	0.881	0.82 *	0.961	0.751 *	0.907
	UNRATE	0.832	0.671	0.836	0.818	0.8	0.859	0.87	0.926	0.733	0.92
h=6	CLAIMS	0.882	0.894	0.894	0.889	0.89	0.908	0.873	0.881	0.819	0.964
	HOUTST	0.619 ***	0.691 ***	0.691 ***	0.594 **	0.807 ***	0.631 ***	0.703 ***	0.625 ***	0.543 ***	0.725 **
	PERMIT	0.637 ***	0.734 *	0.734 *	0.656 **	0.889	0.735 *	0.707 **	0.721 **	0.608 ***	0.869 *
	RCON	0.751	0.843	0.843	0.789	0.853	0.829	0.84	0.791	0.562	0.876
	MTS	0.862 *	0.986	0.986	0.878	0.878	0.92	0.91 *	0.91 *	0.799	0.992 *
	M2	0.878	0.95	0.95	0.904	0.999	0.893	0.95	0.918	0.86 *	0.972
	R10	0.943	0.919	0.992	0.962	0.948	0.943	0.945	0.994	0.943	0.98
	PPI	0.861 **	0.997	0.997	0.967	1.008	0.907 *	0.98	0.903	0.852 **	0.918
	CPI	0.879 ***	0.936	0.936	0.938	0.98	0.872	0.935 *	0.918 *	0.832 *	0.966
	PCECI	0.87 **	0.977	0.977	0.897	0.852	0.901	0.927	0.917	0.837 **	0.965 **
	S&P500	0.853	0.98	0.98	0.905	0.995	0.912	0.965	0.965	0.776	0.957
	RPI	0.855 *	0.849 **	0.849 **	0.893	0.836 *	0.875	0.822 *	0.892	0.766	0.997
	INDPRO	0.829	0.654	0.991	0.741	0.831	0.787	0.849	0.966	0.752	0.924
	UNRATE	0.821	0.711	0.837	0.789	0.807	0.863	0.888	0.954	0.691	0.979
h=12	CLAIMS	0.875	0.851	0.851	0.841	0.873	0.893	0.865	0.884	0.833	0.936
	HOUTST	0.575 ***	0.654 *	0.654 *	0.586 **	0.894 ***	0.605 **	0.766	0.635 **	0.561 ***	0.762
	PERMIT	0.582 ***	0.662	0.662	0.569 ***	0.815	0.847	0.676	0.732	0.567 ***	0.869 ***
	RCON	0.743	0.804	0.804	0.718	0.807	0.796	0.799	0.761	0.657	0.885
	MTS	0.837	0.98	0.98	0.771	0.844	0.88	0.968	0.968	0.798	0.986
	M2	0.933	0.938	0.938	0.953	1.001	0.979	0.935	0.931	0.906	0.962
	R10	0.935	0.859	0.982	0.97	0.94	0.918	0.955	0.995	0.935	0.979
	PPI	0.892 *	0.995	0.995	0.982	0.989	0.938 ***	0.984	0.915 *	0.889 **	0.951
	CPI	0.903	0.975	0.975	0.887 **	0.978	0.892 *	0.97	0.965	0.874	0.982
	PCECI	0.9	0.993	0.993	0.907 **	0.89 **	0.893 ***	0.949	0.943 *	0.886	0.987
	S&P500	0.869	0.946	0.946	0.927	0.938	0.937	0.926	0.926	0.828 *	0.98

* 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 indexes (BCIs), constructed using the methods outlined in Section 2.3.

Table 5: Comparison of the Predictive Accuracy of AR(SIC) Models with AR(SIC) + EUI + BCI Models*

		Factor Extraction Method used to Construct Business Conditions Indexes									
Target Variable	PCA	Lasso	LPCA	AP	GP1	GP2	GF1	GF2	EPCA	JLN	
RPI	0.649	0.706	0.708	0.698	0.684	0.675	0.712	0.672	0.587	0.718	
INDPRO	0.583 *	0.51 **	0.812	0.559 *	0.605 **	0.618	0.563 *	0.651	0.535 **	0.52 *	
UNRATE	0.542 *	0.528 **	0.535 *	0.537	0.531 *	0.544 *	0.533	0.614	0.53 *	0.534	
CLAIMS	0.544	0.556	0.55	0.63	0.632	0.559	0.549	0.53	0.527 *	0.644	
h=1	HOUST	0.579 ***	0.502 ***	0.503 ***	0.52 ***	0.555 ***	0.553 ***	0.679 ***	0.569 ***	0.5 ***	0.642 ***
	PERMIT	0.646 ***	0.797 *	0.797 *	0.67 **	0.872	0.656 **	0.711 *	0.746 *	0.621 ***	0.813 **
	RCON	0.541	0.646	0.69	0.544	0.666	0.564	0.726	0.74	0.527 *	0.703
	MTS	0.533 **	0.713	0.745	0.536 *	0.565 **	0.543 *	0.687	0.698	0.526 **	0.627
	M2	0.673	0.643	0.65	0.673	0.731	0.687	0.8	0.717	0.659	0.902
	R10	0.853	0.806	0.869	0.878	0.832 *	0.86	0.904	0.898	0.852	0.866
	PPI	0.711 **	0.926	0.924	0.884	0.922	0.716 **	0.908	0.793	0.705 **	0.785
	CPI	0.798	0.85	0.856 *	0.743 **	0.906	0.743 *	0.847	0.822	0.645 ***	0.906
	PCECI	0.751 **	0.778 **	0.784 **	0.831 *	0.676 **	0.735 **	0.801 **	0.784 **	0.642 ***	0.908
	S&P500	0.786	0.864	0.874	0.817	0.886	0.788	0.861	0.884	0.699	0.901
	RPI	0.758 *	0.806 **	0.8	0.782 *	0.809 *	0.765	0.834	0.759	0.699 **	0.844
	INDPRO	0.683	0.639	0.862	0.712	0.827	0.661	0.743	0.731	0.702	0.81
h=3	UNRATE	0.542	0.54	0.593	0.585	0.6	0.566	0.667	0.636	0.531	0.656
	CLAIMS	0.622	0.751	0.709	0.636	0.753	0.612	0.738	0.565	0.592	0.782
	HOUST	0.536 ***	0.525***	0.566 ***	0.551 ***	0.61 ***	0.518 ***	0.678 ***	0.548 ***	0.516 ***	0.66 ***
	PERMIT	0.619 ***	0.63 ***	0.737 **	0.682 ***	0.757 *	0.555 ***	0.728 **	0.616 ***	0.639 ***	0.826
	RCON	0.584	0.829	0.768	0.647	0.784	0.601	0.893	0.579	0.578	0.738
	MTS	0.695	0.916	0.876	0.696	0.852	0.696	0.879	0.732	0.731	0.848
	M2	0.739	0.796 *	0.795 *	0.777	0.89	0.721	0.908 *	0.727	0.768	0.921
	R10	0.873 *	0.813	0.949	0.936	0.876	0.876 **	0.86	0.929	0.886	0.912
	PPI	0.789 *	0.956	0.948	0.912	0.934 *	0.837	0.934	0.852	0.827 **	0.897 **
	CPI	0.829 *	0.899 **	0.893 **	0.864 **	0.918 *	0.855 *	0.863	0.858 **	0.814 *	0.938 *
	PCECI	0.836 **	0.901 *	0.896 **	0.862 ***	0.836	0.866	0.862	0.85 *	0.796 **	0.943 **
	S&P500	0.79	0.926	0.922	0.87	0.97	0.804	0.901	0.894	0.752 *	0.934
h=6	RPI	0.871 *	0.846 **	0.849 *	0.855	0.841 **	0.87	0.842	0.865	0.76 *	0.922
	INDPRO	0.822 *	0.643	0.936	0.776	0.757	0.826	0.791 *	0.938	0.726 *	0.825
	UNRATE	0.764	0.653	0.826	0.785	0.791	0.755	0.861	0.904	0.717	0.855
	CLAIMS	0.815	0.871	0.873	0.834	0.874	0.798	0.855	0.824	0.787	0.908
	HOUST	0.579 ***	0.609 ***	0.602 ***	0.544 **	0.652 ***	0.557 ***	0.654 ***	0.616 ***	0.525 ***	0.713 **
	PERMIT	0.589 ***	0.608 ***	0.593 ***	0.633 **	0.715 **	0.665 *	0.672 **	0.688 ***	0.605 **	0.869 *
	RCON	0.738	0.744	0.756	0.752	0.832	0.761	0.829	0.761	0.523	0.711
	MTS	0.852 *	0.942	0.941	0.836 *	0.844	0.865	0.881	0.874	0.767	0.885
	M2	0.86	0.891	0.882	0.864	0.964	0.865 *	0.923	0.873	0.825 *	0.924
	R10	0.948	0.88	0.956	0.949	0.918	0.92	0.916	0.973	0.939	0.964
	PPI	0.846 **	0.958 *	0.956 *	0.951 **	0.971	0.893 **	0.905	0.878 *	0.834 **	0.9
	CPI	0.868 ***	0.9	0.898 *	0.929 *	0.937 **	0.847	0.804	0.888 **	0.817 *	0.946 *
h=12	PCECI	0.861 **	0.94	0.939	0.894	0.839	0.882	0.818	0.882 *	0.823 **	0.946 **
	S&P500	0.849	0.942	0.937	0.91	0.943	0.913	0.777	0.943	0.775	0.932
	RPI	0.839	0.745	0.745	0.867	0.818 *	0.876	0.789 **	0.844	0.694	0.992
	INDPRO	0.803	0.565	0.892	0.717	0.788	0.762	0.773	0.96	0.646	0.913
	UNRATE	0.802	0.593	0.721	0.771	0.789 *	0.848	0.857	0.94	0.625	0.953
	CLAIMS	0.86	0.738	0.74	0.811	0.866	0.881	0.848	0.855	0.765	0.922
	HOUST	0.531 ***	0.576 *	0.579 *	0.521 ***	0.611	0.595 **	0.705	0.607 **	0.52 ***	0.759
	PERMIT	0.536 ***	0.528 *	0.532 *	0.511 ***	0.598	0.808	0.649	0.684	0.509 ***	0.861 ***
	RCON	0.694	0.677	0.681	0.66	0.793	0.728	0.782	0.742	0.522	0.857
	MTS	0.801	0.787	0.792	0.734	0.815	0.853	0.936	0.955	0.605	0.968
	M2	0.916	0.843	0.847	0.935	0.964	0.95	0.889 *	0.902	0.824	0.951
	R10	0.942	0.815	0.927	0.955	0.898 *	0.926	0.855	0.973	0.903	0.976
	PPI	0.882 *	0.952	0.957	0.966	0.927 **	0.918 ***	0.896	0.884 *	0.863 **	0.935
	CPI	0.898	0.934 **	0.937 **	0.874 **	0.918 *	0.882 *	0.881	0.931 *	0.852	0.965
	PCECI	0.896	0.954 *	0.956 *	0.899 **	0.841 ***	0.887 **	0.859	0.912 ***	0.856 *	0.967
	S&P500	0.86	0.905	0.909 *	0.912	0.846	0.935	0.718	0.879 **	0.79	0.986

* 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 both economic uncertainty indexes (EUIs) and business conditions indexes (BCIs), constructed using the methods outlined in Section 2.3.

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

Target Variable	Rank	h=1	h=3	h=6	h=12
RPI	1	AR+EUI+BCI (EPCA)	AR+EUI+BCI (EPCA) **	AR+EUI+BCI (EPCA) *	AR+EUI+BCI (EPCA)
	2	AR+BCI (EPCA)	AR+BCI (EPCA) **	AR+BCI (EPCA)	AR+EUI+BCI (Lasso)
	3	AR+EUI+BCI (PCA)	AR+EUI+BCI (PCA) *	AR+EUI+BCI (GP1) **	AR+EUI+BCI (LPCA)
INDPRO	1	AR+EUI+BCI (Lasso) **	AR+EUI+BCI (Lasso)	AR+EUI+BCI (Lasso)	AR+EUI+BCI (Lasso)
	2	AR+EUI+BCI (JLN)*	AR+EUI+BCI (GP2)	AR+BCI (Lasso) *	AR+EUI+BCI (EPCA)
	3	AR+EUI+BCI (EPCA) **	AR+BCI (Lasso)	AR+EUI+BCI (EPCA) *	AR+BCI (Lasso)
UNRATE	1	AR+EUI+BCI (Lasso) **	AR+EUI+BCI (EPCA)	AR+EUI+BCI (Lasso)	AR+EUI+BCI (Lasso)
	2	AR+BCI (Lasso) **	AR+EUI+BCI (Lasso)	AR+BCI (Lasso)	AR+EUI+BCI (EPCA)
	3	AR+EUI+BCI (EPCA) *	AR+EUI+BCI (PCA)	AR+EUI+BCI (EPCA)	AR+BCI (EPCA)
CLAIMS	1	AR+EUI+BCI (EPCA) *	AR+EUI+BCI (GF2)	AR+EUI+BCI (EPCA)	AR+EUI+BCI (Lasso)
	2	AR+EUI+BCI (GF2)	AR+EUI+BCI (EPCA)	AR+EUI+BCI (GP2)	AR+EUI+BCI (LPCA)
	3	AR+BCI (EPCA) *	AR+EUI+BCI (GP2)	AR+EUI+BCI (PCA)	AR+EUI+BCI (EPCA)
HOUST	1	AR+EUI+BCI (EPCA) ***	AR+EUI+BCI (EPCA) ***	AR+EUI+BCI (EPCA) ***	AR+EUI+BCI (EPCA) ***
	2	AR+EUI+BCI (Lasso) ***	AR+EUI+BCI (GP2) ***	AR+BCI (EPCA) ***	AR+EUI+BCI (AP) ***
	3	AR+EUI+BCI (LPCA) ***	AR+EUI+BCI (Lasso) ***	AR+EUI+BCI (AP) **	AR+EUI+BCI (PCA) ***
PERMIT	1	AR+EUI+BCI (EPCA) ***	AR+EUI+BCI (GP2) ***	AR+EUI+BCI (PCA) ***	AR+EUI+BCI (EPCA) ***
	2	AR+EUI+BCI (PCA) ***	AR+EUI+BCI (GF2) ***	AR+EUI+BCI (LPCA) ***	AR+EUI+BCI (AP) ***
	3	AR+EUI+BCI (GP2) **	AR+EUI+BCI (PCA) ***	AR+EUI+BCI (EPCA) **	AR+EUI+BCI (Lasso) *
RCON	1	AR+EUI+BCI (EPCA) *	AR+EUI+BCI (EPCA)	AR+EUI+BCI (EPCA)	AR+EUI+BCI (EPCA)
	2	AR+BCI (EPCA) *	AR+EUI+BCI (GF2)	AR+BCI (EPCA)	AR+BCI (EPCA)
	3	AR+EUI+BCI (PCA)	AR+EUI+BCI (PCA)	AR+EUI+BCI (JLN)	AR+EUI+BCI (AP)
MTS	1	AR+EUI+BCI (EPCA) **	AR+EUI+BCI (PCA)	AR+EUI+BCI (EPCA)	AR+EUI+BCI (EPCA)
	2	AR+BCI (EPCA) *	AR+EUI+BCI (AP)	AR+BCI (EPCA)	AR+EUI+BCI (AP)
	3	AR+EUI+BCI (PCA) **	AR+EUI+BCI (GP2)	AR+EUI+BCI (AP) *	AR+BCI (AP)
M2	1	AR+EUI+BCI (Lasso)	AR+EUI+BCI (PCA)	AR+EUI+BCI (EPCA) *	AR+EUI+BCI (EPCA)
	2	AR+EUI+BCI (LPCA)	AR+EUI+BCI (GP2)	AR+BCI (PCA)	AR+EUI+BCI (Lasso)
	3	AR+EUI+BCI (AP)	AR+EUI+BCI (GF2)	AR+BCI (EPCA) *	AR+EUI+BCI (LPCA)
R10	1	AR+EUI+BCI (Lasso)	AR+EUI+BCI (Lasso)	AR+EUI+BCI (Lasso)	AR+EUI+BCI (Lasso)
	2	AR+EUI+BCI (GP1) *	AR+BCI (Lasso)	AR+EUI+BCI (GF1)	AR+EUI+BCI (GF1)
	3	AR+BCI (Lasso)	AR+EUI+BCI (GF1)	AR+EUI+BCI (GP1)	AR+BCI (Lasso)
PPI	1	AR+BCI (PCA) **	AR+EUI+BCI (PCA) *	AR+EUI+BCI (EPCA) **	AR+EUI+BCI (EPCA) **
	2	AR+BCI (EPCA) **	AR+EUI+BCI (GP2)	AR+BCI (EPCA) **	AR+EUI+BCI (PCA) *
	3	AR+EUI+BCI (PCA) **	AR+EUI+BCI (EPCA) **	AR+EUI+BCI (PCA) **	AR+BCI (PCA) *
CPI	1	AR+BCI (EPCA) ***	AR+EUI (GF1)	AR+EUI (GF1)	AR+EUI (GP1) *
	2	AR+EUI+BCI (EPCA) ***	AR+BCI (PCA)	AR+BCI (PCA) ***	AR+EUI (GF1)
	3	AR+BCI (PCA) *	AR+BCI (AP)	AR+BCI (GP2)	AR+BCI (PCA)
PCECI	1	AR+BCI (PCA) **	AR+EUI (PCA)	AR+EUI (Lasso)	AR
	2	AR+BCI (Lasso) ***	AR+EUI (Lasso)	AR+EUI (LPCA)	AR+EUI (PCA) *
	3	AR+BCI (LPCA) ***	AR+EUI (LPCA) *	AR+EUI (GF1)	AR+EUI (Lasso) ***
S&P500	1	AR+EUI+BCI (EPCA)	AR+EUI+BCI (EPCA) *	AR+EUI+BCI (EPCA)	AR+EUI+BCI (GF1)
	2	AR+BCI (EPCA)	AR+BCI (EPCA) *	AR+BCI (EPCA)	AR+EUI (GF1)
	3	AR+EUI+BCI (PCA)	AR+EUI+BCI (PCA)	AR+EUI+BCI (GF1)	AR+EUI+BCI (EPCA)

* 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)+BCI, and AR(SIC)+EUI+BCI models. For each of these models, the associated factor extraction method used when constructing the EUIs and BCIs 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 BCIs are yield more accurate predictions than our AR benchmark model, using the (Giacomini & 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” Factor Extraction Methods I*
 Comparison of AR(SIC) Models with AR(SIC) + EUI Models

Target Variable	Rank	h=1	h=3	h=6	h=12
RPI	1	GP1	GF2	JLN	LPCA
	2	Lasso	GP2	GP2	Lasso
	3	LPCA	PCA	EPCA	EPCA
INDPRO	1	Lasso	PCA	JLN	GF1
	2	GP1	GF2	Lasso	LPCA
	3	LPCA	GP2	EPCA	Lasso
UNRATE	1	GP1	PCA	PCA	LPCA
	2	Lasso	GF2	JLN	Lasso
	3	LPCA	GP2	GP2	EPCA
CLAIMS	1	Lasso	GF2	GP2	LPCA
	2	GP1	PCA	JLN	Lasso
	3	LPCA	AP	AP	EPCA
HOUST	1	Lasso	Lasso ***	GP2 ***	PCA **
	2	LPCA	GP2 **	GP1 **	AP *
	3	GP1	GF2 *	Lasso	GF1
PERMIT	1	Lasso	Lasso **	GP2 ***	PCA **
	2	LPCA	GP2 ***	GP1	AP **
	3	GF1	GF2 *	PCA *	GF1
RCON	1	Lasso	PCA	JLN	Lasso
	2	GP1	GF2	LPCA	LPCA
	3	LPCA	AP	Lasso	GP2
MTS	1	Lasso	PCA	JLN	LPCA
	2	GP1	GF2	EPCA	Lasso
	3	LPCA	GP2	Lasso	EPCA
M2	1	Lasso	GF2	JLN	LPCA
	2	GP1	PCA	LPCA	Lasso
	3	LPCA	AP	Lasso *	EPCA
R10	1	Lasso	GF2	GP1	GF1
	2	LPCA	GF1	LPCA	LPCA
	3	GP1	GP1	AP	Lasso
PPI	1	LPCA	GF2	GF1	GF1
	2	Lasso	PCA	LPCA ***	GP1
	3	GP1	AP **	Lasso **	JLN
CPI	1	Lasso	GF1	GF1	GF1
	2	LPCA	GF2 *	Lasso	GP1
	3	GP1	LPCA *	GP1 *	LPCA **
PCECI	1	Lasso	GF1	GF1	GF1
	2	LPCA	GF2 *	Lasso	GP1
	3	GP1	LPCA *	GP1 *	LPCA ***
S&P500	1	Lasso	Lasso	GF1	GF1
	2	GP1	GF1	Lasso	LPCA
	3	LPCA	GP2	GP1	GP1

* Notes: See notes to Table 6. This table lists the top 3 factor 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” Factor Extraction Methods II*
 Comparison of AR(SIC) Models with AR(SIC) + BCI Models

Target Variable	Rank	h=1	h=3	h=6	h=12
RPI	1	EPCA	EPCA **	EPCA	EPCA
	2	PCA	PCA **	GF2 **	GF2 *
	3	GP1	AP **	GF1 *	GF1 *
INDPRO	1	Lasso *	Lasso	Lasso *	Lasso
	2	JLN *	EPCA	EPCA *	AP
	3	EPCA *	GF2	AP	EPCA
UNRATE	1	Lasso **	Lasso	Lasso	EPCA
	2	EPCA *	EPCA	EPCA	Lasso
	3	GF1 *	GF1	GF1	AP
CLAIMS	1	EPCA *	EPCA	EPCA	EPCA
	2	GP2 *	GP2	GF2	AP
	3	PCA	GF2	GP2	Lasso
HOUST	1	EPCA ***	EPCA ***	EPCA ***	EPCA ***
	2	AP ***	AP ***	AP **	PCA ***
	3	Lasso ***	PCA ***	PCA ***	AP **
PERMIT	1	EPCA ***	EPCA ***	EPCA ***	EPCA ***
	2	PCA ***	PCA ***	PCA ***	AP ***
	3	GP1 **	GP1 **	AP **	PCA ***
RCON	1	EPCA *	EPCA	EPCA	EPCA
	2	PCA *	PCA	PCA	AP
	3	AP *	AP	AP	PCA
MTS	1	EPCA *	AP	EPCA	AP
	2	PCA *	EPCA	PCA *	EPCA
	3	AP	PCA	AP	PCA
M2	1	EPCA	EPCA	EPCA *	EPCA
	2	AP	PCA	PCA	GP2
	3	Lasso	GP1	GP1	PCA
R10	1	Lasso	Lasso	Lasso	Lasso
	2	PCA	GF1	PCA	GP1
	3	EPCA	GF2	GP1	PCA
PPI	1	EPCA **	PCA ***	EPCA **	EPCA**
	2	PCA **	PCA **	PCA **	PCA *
	3	GP1 **	GP1 **	GP2	GP2 *
CPI	1	EPCA ***	EPCA *	EPCA *	EPCA
	2	AP **	PCA	GP1	AP **
	3	GP1 *	GP1	PCA ***	GP1 *
PCECI	1	EPCA ***	EPCA **	EPCA **	EPCA
	2	GF1 **	PCA **	GF1	GF1 **
	3	GP1 **	GF1	PCA **	GP1 ***
S&P500	1	EPCA	EPCA *	EPCA	EPCA *
	2	PCA	PCA	PCA	PCA
	3	GP1 *	GP1	AP	GF2

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

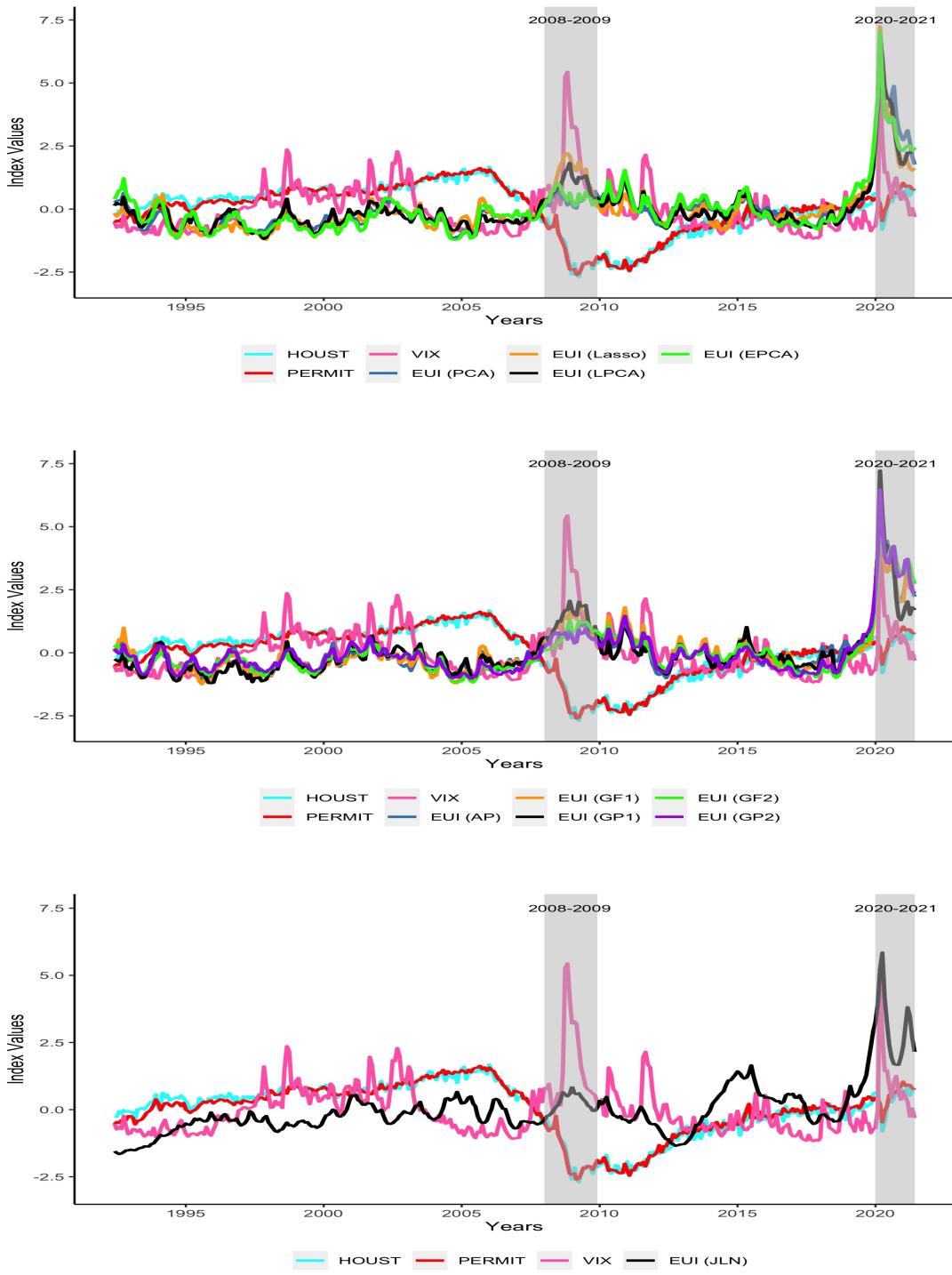
Table 9: Top 3 “MSFE-Best” Factor Extraction Methods III*

Comparison of AR(SIC) Models with AR(SIC) + EUI + BCI Models

Target Variable	Rank	h=1	h=3	h=6	h=12
RPI	1	EPCA	EPCA **	EPCA *	EPCA
	2	PCA	PCA *	GP1 **	Lasso
	3	GF2	GF2	GF1	LPCA
INDPRO	1	Lasso **	Lasso	Lasso	Lasso
	2	JLN *	GP2	EPCA *	EPCA
	3	EPCA **	PCA	GP1	AP
UNRATE	1	Lasso **	EPCA	Lasso	Lasso
	2	EPCA *	Lasso	EPCA	EPCA
	3	GP1 *	PCA	GP2	LPCA
CLAIMS	1	EPCA *	GF2	EPCA	Lasso
	2	GF2	EPCA	GP2	LPCA
	3	PCA	GP2	PCA	EPCA
HOUST	1	EPCA ***	EPCA ***	EPCA ***	EPCA ***
	2	Lasso ***	Lasso ***	AP **	PCA ***
	3	LPCA ***	GP2 ***	GP2 ***	AP ***
PERMIT	1	EPCA ***	GP2 ***	PCA ***	EPCA ***
	2	PCA ***	GF2 ***	EPCA **	PCA ***
	3	GP2 **	PCA ***	LPCA ***	AP ***
RCON	1	EPCA *	EPCA	EPCA	EPCA
	2	PCA	GF2	JLN	AP
	3	AP	PCA	PCA	Lasso
MTS	1	EPCA **	PCA	EPCA	EPCA
	2	PCA **	AP	AP *	AP
	3	AP *	GP2	GP1	LPCA
M2	1	Lasso	GP2	EPCA *	EPCA
	2	LPCA	GF2	GP2 *	LPCA
	3	EPCA	PCA	PCA	Lasso
R10	1	Lasso	Lasso	Lasso	Lasso
	2	GP1 *	GF1	GF1	GF1
	3	EPCA	PCA *	GP2	EPCA
PPI	1	EPCA **	PCA *	EPCA **	EPCA **
	2	PCA **	EPCA **	PCA **	PCA **
	3	GP2 **	GP2	GF2 *	GF2 *
CPI	1	EPCA ***	EPCA *	GF1	EPCA
	2	GP2 *	PCA *	EPCA *	AP **
	3	AP **	GP2 *	GP2	GF1
PCECI	1	EPCA ***	EPCA **	GF1	GP1 ***
	2	GP1 **	GP1	EPCA **	EPCA *
	3	GP2 **	PCA **	GP1	GF1
S&P500	1	EPCA	EPCA *	GF1	GF1
	2	PCA	PCA	EPCA	EPCA
	3	GP2	GP2	PCA	PCA

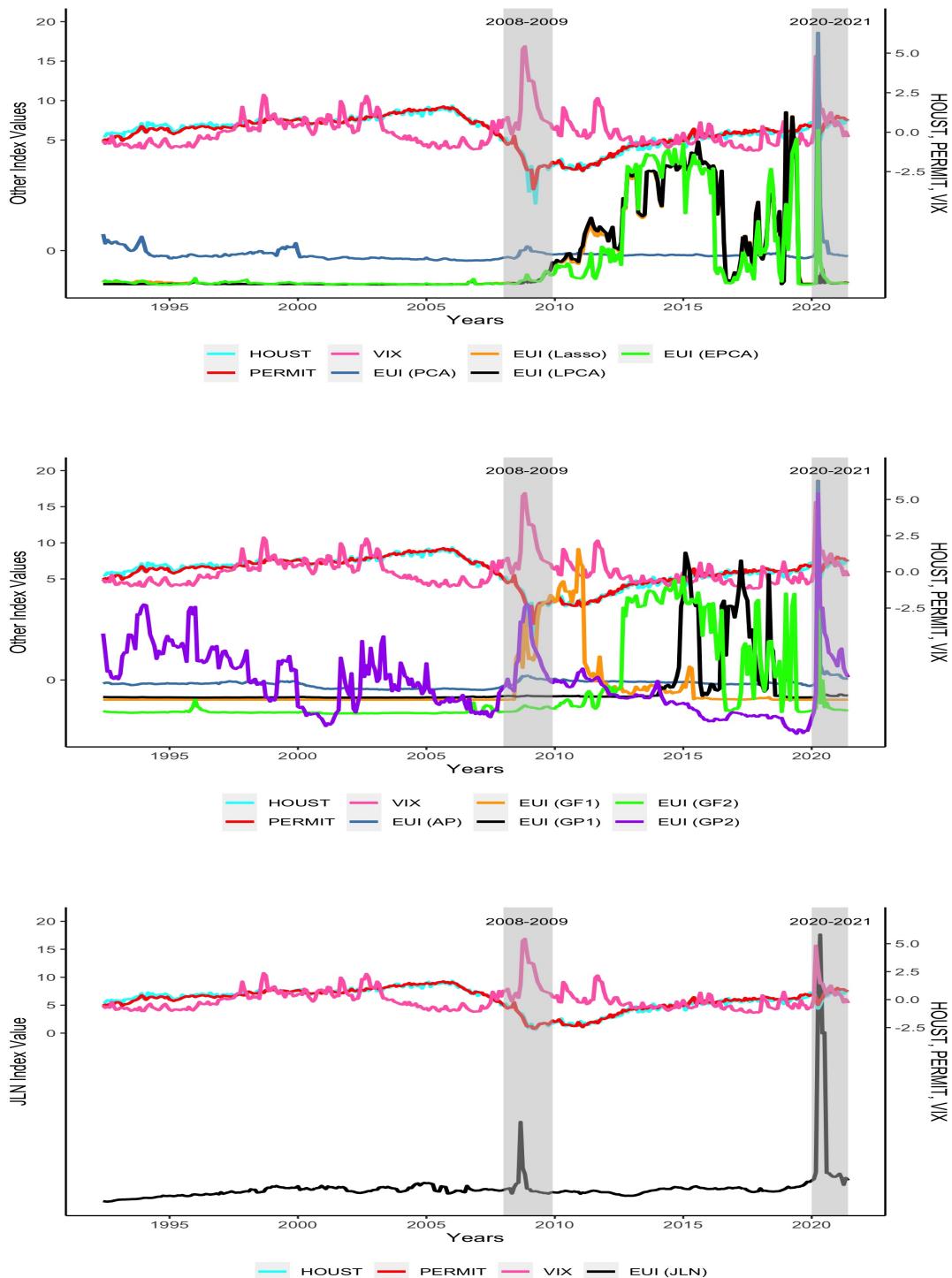
* Notes: See notes to Table 6. This table lists the top 3 factor extraction methods used in our “MSFE-best” forecasting models, when comparing our benchmark AR(SIC) model with our AR(SIC)+EUI+BCI model.

Figure 1: Plots of Housing Starts and Permits, Various Economic Uncertainty Indexes, and the VIX (Forecasting Horizon, $h=1$)*



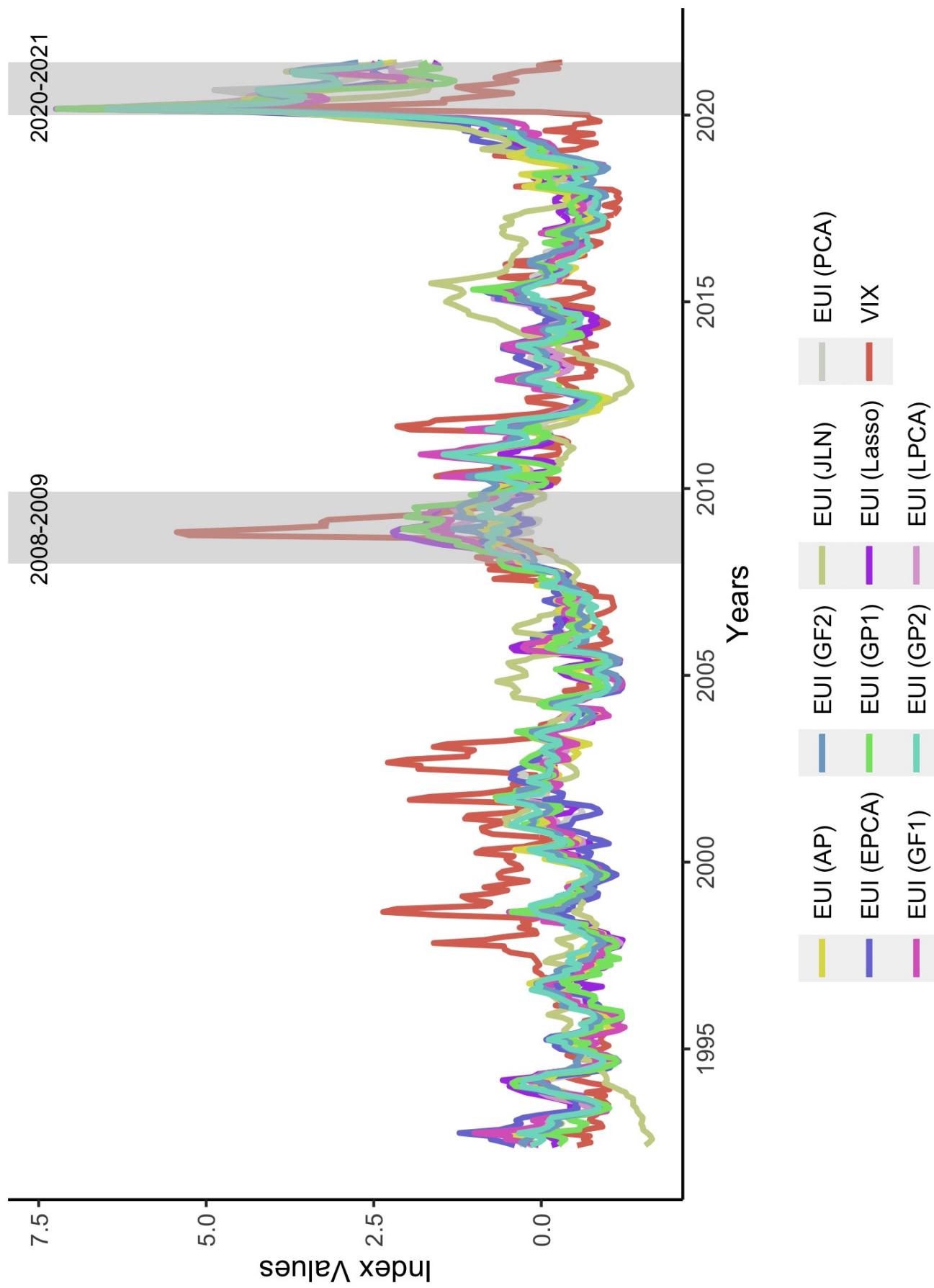
*Notes: See notes to Table 3. The above three figures include various plots of our economic uncertainty indexes (EUIs), HOUST, PERMIT, and the VIX, which is also compared with the EUIs proposed in Jurado et al. (2015). All EUI plots depict real-time measures constructed for the period 1992:6-2021:6, 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 the Covid-19 pandemic period (2020:1-2021:6).

Figure 2: Plots of Housing Starts and Permits, Various Economic Uncertainty Indexes, and the VIX (Forecasting Horizon, $h=12$)*



*Notes: See notes to Figure 1.

Figure 3: Plots of the VIX and Our 10 Economic Uncertainty Indexes that Utilize Forecasts from Business Conditions Indexes Constructed Using Machine Learning Methods (Forecasting Horizon, $h=1$)*



*Notes: See notes to Figure 1. This figure contains plots of all of our economic uncertainty indexes, as well as the VIX.