Macroeconomic and Financial Mixed Frequency Factors in a Big Data Environment*

Weijia Peng[1], Norman R. Swanson[2], Xiye Yang[2], Chun Yao[3]

[1]Sacred Heart University, [2]Rutgers University and [3]Barclays

March 2023

Abstract

In this paper, we evaluate the predictive content of three new business condition indexes and uncertainty measures that are estimated using high frequency financial and low frequency macroeconomic time series data. More specifically, our measures are defined as latent factors that are extracted from a state space model that includes multiple different frequencies of non-parametrically estimated components of quadratic variation, as well as mixed frequency macroeconomic variables. When forecasting growth rates of various monthly financial and macroeconomic variables, use of our new mixed frequency factors is shown to result in significant improvement in predictive performance, relative to a number of benchmark models. Additionally, when used to forecast corporate yields, predictive gains associated with use of our measures are shown to be monotonically increasing, as one moves from predicting higher to lower rated bonds. This is consistent with the existence of a natural pricing channel wherein financial risk (as measured using our volatility factors) contains more predictive information for lower grade bonds. We also find that a variety of extant risk factors, including the Aruoba et al. (2009a) business conditions index also contain marginal predictive content for the variables that we examine, although their inclusion does not reduce the usefulness of our measures.

Keywords: Latent factor, high-frequency financial data, mixed-frequency data, state space model.

* Weijia Peng, Department of Finance, Jack Welch College of Business & Technology, Sacred Heart University, 5151 Park Avenue Fairfield, CT 06825, USA, pengw@sacredheart.edu; Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@economics.rutgers.edu; Xiye Yang, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, xiyeyang@economics.rutgers.edu; Chun Yao, US Economics Research, Barclays, 745 Seventh Avenue, New York, NY 10019, USA, chunyao.economics@gmail.com. We are grateful to the editor, Janine Illian, the associate editor, two anonymous referees, Mingmian Cheng, Valentina Corradi, Frank Diebold, Xu Jiang, Hyun Hak Kim, Yuan Liao, and Greg Tkacz for useful comments and suggestions on the topics explored in this paper. The views expressed here are those of the authors and not necessarily those of any affiliated institution. Declarations of interest: none.

1 Introduction

Uncertainty concerning the future plays an important role in the decisions of households and firms, and influences macroeconomic variables such as employment and income, as well as financial variables such as interest rates and stock market returns. For this reason, a large literature has evolved that specifies and analyzes new measures of business cycle conditions and uncertainty. Many papers in this literature focus on estimating measures of uncertainty by utilizing what often amounts to estimates of "forecast" errors (see e.g., Bloom et al. (2018), Baker et al. (2016), Carriero et al. (2016), Jo & Sekkel (2017), and Jurado et al. (2015)). In a related strand of the literature, uncertainty is instead measured directly, by examining volatility of so-called "business conditions indexes" constructed using mixed frequency models (see Aruoba et al. (2009a)) or by directly estimating volatility using high frequency financial times series variables (see e.g. Chauvet et al. (2015) and Cheng et al. (2021)).[1] Finally, a number of authors define business conditions indexes and uncertainty measures as latent factors that can be extracted from datasets of varying size, ranging from those that are low-dimensional and low-frequency to those that are high-dimensional and high-frequency (i.e., big data).

This paper adds to this large and nascent literature by specifying new mixed frequency latent factors that are designed to be useful in the context of financial and macroeconomic forecasting. In particular, we estimate latent factors from a state space model that includes multiple different frequencies of non-parametrically estimated components of volatility (i.e. quadratic variation extracted from high frequency financial returns), as well as mixed frequency macroeconomic variables. Notably, several previous studies have also utilized variables of multiple different frequencies for forecasting and for the estimation of latent factors. For example, Aruoba et al. (2009a) extract business conditions indexes from four key macroeconomic variables of different observational frequencies.[2] Additionally, Marcellino et al. (2016) use a mixed-frequency dynamic factor model to investigate business cycles in the euro area, and Andreou et al. (2013) use the Mixed Data Sampling (MIDAS) framework developed in Ghysels et al. (2007) to include daily data when forecasting lower frequency macroeconomic

[1] Needless to say, measures of uncertainty have long been estimated directly from low dimensional data, as well. However, in this paper, we focus on the use of recently available high frequency financial datasets for our analysis.

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

variables.[3]

More specifically, we introduce multi-frequency macroeconomic and financial factors that are purported to contain useful predictive content for macroeconomic and financial variables. Our latent mixed frequency factors are extracted from state space models that include multiple different frequencies of these variables, as well as multiple different frequencies of non-parametrically estimated components of quadratic variation. Our models include data frequencies ranging from 5-minutes to quarterly, and our latent factors specified in one of two ways. First, they are specified solely using latent components of quadratic variation, including continuous and jump component variation measures extracted from high frequency S&P500 data. These are our volatility type factors, called mixed frequency financial volatility (MFV) factors.[4] Alternatively, they are specified using mixed-frequency observed variables, including macroeconomic indicators such as interest rates, employment, and production. These are called our mixed frequency macroeconomic (MAC) factors. Finally, we extract factors that include both quadratic variation and macroeconomic components. These are called our mixed frequency financial volatility and macroeconomic (MIX) factors. Related papers that utilize mixed-frequency state space models include Mariano & Murasawa (2003), Frale et al. (2008), Aruoba et al. (2009a) and Marcellino et al. (2016). None of these papers, however, include multiple frequencies of the same latent variable, as is done in this paper.[5]

[3]There is now a large literature establishing the relative forecasting (and nowcasting) gains associated with using mixed frequency models for macroeconomic forecasting. For further discussion of nowcasting as well as of mixed frequency based forecasts, refer to Aastveit et al. (2014), McAlinn (2021), Andreou et al. (2013), and the references cited therein. Additionally, it should be noted that there are many interesting modeling approaches that lend themselves to estimating factors, and that constitute viable alternatives to the approach taken in this paper (see e.g. West & Harrison (1997), Ghysels et al. (2007), Aguilar & West (2000), and McAlinn (2021).

[4]As briefly mentioned above, various papers utilize approaches that differ from ours for measuring volatility (some use forecast errors and some directly estimate a variable's volatility, for example) when incorporating financial market information in uncertainty measures. For instance, Bloom (2009) and Basu & 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 et al. (2014) construct realized volatility measures using a micro-level firm-specific asset returns dataset. Finally, Carriero et al. (2016) propose a VAR model with stochastic volatility driven by common factors, Jo & Sekkel (2017) extract common factors using a stochastic volatility model, and Carriero et al. (2015) build a Bayesian model in order to estimate latent uncertainty measures.

[5]An interesting alternative method to the state space modeling approach used in this paper is the mixed data sampling (MIDAS) approach proposed by Ghysels et al. (2007). The idea underlying this method is to establish a regression relation between a low-frequency variable and a set of higher-frequency variables that are aggregated by dynamic weighting functions. Following this idea, Andreou et al. (2013) demonstrate how daily financial data can be incorporated into a forecasting model for quarterly GDP.

Before continuing, it is worth stressing that the latent factors examined in this paper are designed to contain useful economic information pertaining to both economic conditions and financial uncertainty. It is posited that said information may improve the predictive accuracy of model-based forecasts of macroeconomic and financial variables. 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, both leading to increased profitability. In our analysis, we show that corporate bond yields, for example, can be more precisely forecasted when our mixed frequency financial volatility (uncertainty) measures are added to a variety of forecasting models that are widely utilized in the industry. Another key real-world example where our results prove useful concerns government policy setting behavior. 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. Given these forecasts, they proceed to decide how and when to change the federal funds rate.[6] This works because the federal funds rate is closely linked to market interest rates, which are in turn closely linked to inflation and economic growth. More precise predictions of the variables that they rely on will lead to policy decisions concerning future federal funds rate increases and decreases that more precisely deliver the desired effects on inflation and growth. In our analysis, we show that the economic conditions and financial uncertainty measures that we develop deliver statistically significant predictive accuracy improvements when they are added to various models that are widely utilized by the federal government for forecasting a variety of variables ranging from housing starts to industrial production.[7]

An important paper in this area which is closely related to ours is Chauvet et al. (2015). These authors also implement a state space model to extract common components from realized volatilities

[6]As stated on the Federal Reserve Bank of St. Louis website: "The federal funds rate is the interest rate at which depository institutions trade federal funds (balances held at Federal Reserve Banks) with each other overnight. When a depository institution has surplus balances in its reserve account, it lends to other banks in need of larger balances. In simpler terms, a bank with excess cash, which is often referred to as liquidity, will lend to another bank that needs to quickly raise liquidity." See https://fred.stlouisfed.org/series/FEDFUNDS

[7]These variables are in turn often used by the federal government when forecasting the impact of possible interest rate changes on variables like inflation and GDP growth.

of stocks and bonds. A key difference between our approach and theirs is that while they include high-frequency based measures of volatility in their analysis, all of their measures are estimated using data of the same frequency. Our multi-frequency approach instead builds on the work of Corsi (2009), where the use of heterogeneous autoregressive realized volatility models is motivated by arguing that agents with different decision horizons react to, and cause, different volatility dynamics. In this paper it is argued that there are short-term traders with daily (or higher) trading frequencies, medium-term investors who typically re-balance their positions weekly, and long-term investors who induce low frequency volatility dynamics. Our approach mirrors this logic and considers volatility frequencies of daily, bi-daily, tri-daily, and weekly, in order to capture effects associated with short-term and medium-term agent decisions. Finally, we would be remiss if we did not stress that our paper builds on the key paper by Aruoba et al. (2009a), in which a business conditions index is constructed by extracting a latent factor from macroeconomic variables of different observational frequencies. A key difference between our approach and that of Aruoba et al. (2009a) is that they do not include nonparametric measures of uncertainty constructed using high frequency data. Instead, they analyze a model that includes macroeconomic indicators. In order to compare our results with theirs, we include a variant of our model which nests their model.

Our key findings can be summarized as follows. First, we provide broad empirical evidence that mixed frequency factors extracted from models that include both integrated volatility measures nonparametrically constructed using high frequency S&P500 returns data and mixed frequency macroeconomic data (our MIX factors) contain substantially more predictive content that mixed frequency factors constructed solely using mixed frequency macroeconomic data (our MAC factors). This is particularly apparent when observing the usefulness of our factors after the recession of 2008. Second, models augmented with our new mixed frequency factors represent the majority of the "best"-performing models in out-of-sample forecasting experiments, when comparing point mean square forecast error measures; when comparing point directional forecast accuracy measures; and when applying Giacomini-White predictive accuracy tests. Our forecasts are statistically superior to those constructed using various benchmark models including simple autoregressive models, as well as models that directly incorporate nonparametric measures of integrated volatility. Indeed, large forecasting accuracy gains are observed for a number of macroeconomic "target" variables that we forecast, including housing starts, industrial production, and payroll employment. Third, and as stated above, some of our very best mixed frequency factors are those extracted from state space models that include both integrated volatility measures and mixed frequency macroeconomic variables. Moreover, these measures are often in model confidence sets that include the "best" model. Fourth, an interesting pattern emerges when using mixed frequency factors extracted from state space models that include only financial variables (our MFV factors) to forecast corporate bond yields. Namely, our MFV factors deliver monotonically increasing predictive accuracy gains (as measured by mean square forecast error (MSFE)), as one moves from predicting bonds with higher ratings to predicting bonds with lower ratings. This is consistent with the existence of a natural pricing channel wherein financial risk is more important, predictively, for lower grade bonds. For example, models that include MFV factors are generally associated with 30% to 40% MSFE drops for bonds with ratings lower than BB; are associated with 10% to 20% drops for A and BBB rated bonds; and are associated with no drops for AAA and AA rated bonds. Summarizing, the highest rated investment grade bonds seem to act as "safe haven", in the sense that they show little dependence on volatility. Finally, we analyze four commonly used integrated volatility measures, including realized volatility (RV_t), truncated realized volatility (TRV_t), bi-power variation (BPV_t), and jump variation ($JV_t = RV_t - BPV_t$), and find that continuous quadratic variation measures (i.e., RV_t, TRV_t, and BPV_t), are the most useful in our context. This is perhaps not surprising, given the difficulties noted in the financial econometrics literature associated with extracting useful predictive content from jump variation measures. In summary, we offer compelling evidence of the "predictive" usefulness of the mixed frequency factors developed in this paper.

The rest of this paper is organized as follows. In Section 2, we outline the methodology used in the construction of the mixed frequency factors analyzed in the sequel. In Section 3, we outline the experimental setup used in order to examine the predictive content of our mixed frequency factors. Section 4 contains a description of the data used in our empirical analysis, and Section 5 contains the results of our forecasting experiments. Finally, Section 6 concludes.

2 Mixed Frequency Factors

To describe the methodology used in constructing the mixed frequency factors analyzed in the subsequent sections, we provide a detailed overview of our approach. We first summarize our method for addressing temporal aggregation and missing observations. Next, we review the high frequency measures of volatility used, followed by a detailed explanation of the state space modeling framework implemented to estimate latent mixed frequency volatility factors, macroeconomic factor, and financial volatility and macroeconomic factors. These factors are estimated using state space models that incorporate both high frequency nonparametric integrated volatility and mixed frequency macroeconomic variables.[8]

2.1 Temporal Aggregation

The temporal aggregation of variables of different frequencies and the stock and flow features of the variables that we examine have been discussed in previous research, such as Aruoba et al. (2009a) and Aruoba et al. (2009b). It is important to address these issues in the specification of our state space models.

We first differentiate between flow and stock variables and consider their impact on our state space models. Flow variables are accumulated within each period, while stock variables reflect quantities measured at a particular point in time. When the state space model is evolving at a higher frequency than the flow variable, accumulated values can result in regular shocks to the state variable. For instance, if the state space model is evolving at a daily frequency, and we observe a monthly flow variable, its accumulated values over the past 30 days will result in a shock to the state variable each time a new value is updated. To account for this, proper incorporation of flow variable observations into the system is necessary for accurate estimation of the latent factors.

[8] In a not for publication appendix that accompanies this paper, we include further discussion of the state-space modeling framework employed in this paper, along with key references to the literature, including Anderson & Moore (2012), Cargnoni et al. (1997), and Prado & West (2010). The appendix is published at https://econweb.rutgers.edu/nswanson/papers.htm.

For stock variables, this complication does not arise. The observed value for a stock variable can simply be expressed as a function of the current state variable and the stochastic disturbance term. As an example, let F_t denote the state variable at time t, and let the stock variable be y_t^s, then:

$$y_t^s = \beta F_t + u_t, \tag{2.1}$$

where F_t is state variable, and u_t is a stochastic disturbance term.

On the other hand, and as discussed above, the value of flow variables reflects the aggregated value through each time period. Thus, flow variable y_t^f can be defined as follows:

$$y_t^f = \sum_{i=0}^{K_j - 1} y_{t-i}^f, \tag{2.2}$$

where indices i and j denote the i^{th} time point within the j^{th} observational interval, and K_j is the length of the interval between two observational time points (i.e., time points for which observations are available - namely, between the $(j-1)^{th}$ and j^{th} time points). Since the value of flow variable is inter-temporally accumulated over a given period of time, one straightforward way to handle inter-temporal aggregation is by defining a state vector that sums all lags of states within each period. For example, a monthly flow variable in a daily state space model can be specified as:

$$y_t^f = \beta(F_t + F_{t-1} + F_{t-2} + \cdots + F_{t-m}) + u_t, \tag{2.3}$$

where F_t, F_{t-1},$\cdots$, F_{t-m} are state variable components, and u_t is a stochastic disturbance term.

However, given that our state space is evolving at a daily frequency, and the lowest frequency flow variable is quarterly real GDP, this approach will result in the specification of a very large state variable with more than 90 lag terms, and a large number of parameters to be estimated, causing excessive calculation and convergence issues. For this reason, we instead implement the aggregated states approach of Aruoba et al. (2009b) in order to account for flow variables in our system. Namely, we define

$$y_t^f = \beta C_t + \gamma y_{t-M}^f + w_t, \tag{2.4}$$

where C_t is a latent state variable defined specifically for flow variables, M is the observational lag length, and the w_t are serially uncorrelated error terms. Here, C_t sums over its past values within the observational period of the flow variables. Namely,

$$C_{t+1} = \psi_{t+1} C_t + \rho F_t, \tag{2.5}$$

where

$$\psi_t = \begin{cases} 0, & \text{if } t \text{ is the first observation of the period} \\ 1, & \text{otherwise,} \end{cases}$$

and where ψ_t is an indicator that controls for the observational frequency of the flow variable. Hence, if a flow variable is updated at time t, then the value of C_{t+1} will be refreshed to be $0 + \rho F_t$, while if a flow variable is not updated at time t, then $C_{t+1} = C_t + \rho F_t$, which includes its past value in the sum.

2.2 High frequency measures of volatility and jump variation

Let X_t be the log-price of an asset at time t. Assume that the log-price process follows a jump-diffusion model (hence, almost surely, its paths are right continuous with left limits). Namely,

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dB_s + \sum_{s \le t} \Delta X_s. \tag{2.6}$$

In the above expression, B is a standard Brownian motion and $\Delta X_s := X_s - X_{s-}$, where $X_{s-} := \lim_{u \uparrow s} X_u$, represents the possible jump of the process X, at time s.

Consider a finite time horizon, $[0, t]$ that contains n high-frequency observations of the log-price process. A typical time horizon is one day. Let $\Delta_n = t/n$ be the sampling frequency. Then intra-daily returns can be expressed as $r_{i,n} = X_{i\Delta_n} - X_{(i-1)\Delta_n}$.

A well-established result in the high frequency econometrics literature is that realized volatility is a consistent estimator of the total quadratic variation. Namely,

$$RV_t = \sum_{i=1}^{n} r_{i,n}^2 \xrightarrow{\text{u.c.p.}} \int_0^t \sigma_s^2 ds + \sum_{s \leq T} (\Delta X_s)^2 = QV_t = IV_t + JV_t, \tag{2.7}$$

where $\xrightarrow{\text{u.c.p.}}$ denotes convergence in probability, uniformly in time. There are many estimators of integrated volatility (IV_t), which is the variation due to the continuous component of quadratic variation (QV_t). For example, multipower variations are defined as follows:

$$V_t = \sum_{i=j+1}^{n} |r_{i,n}|^{\gamma_1} |r_{i-1,n}|^{\gamma_2} ... |r_{i-j,n}|^{\gamma_j}, \tag{2.8}$$

where $\gamma_1, \gamma_2, ..., \gamma_j$ are positive such that $\sum_{i=1}^{j} \gamma_i = k$. An important special case of this estimator is bipower variation (BPV_t), which was introduced by Barndorff-Nielsen & Shephard (2004). Namely,

$$BPV_t = (\mu_1)^{-2} \sum_{i=2}^{n} |r_{i,n}||r_{i-1,n}|, \tag{2.9}$$

where $\mu_1 = E(|Z|) = 2^{1/2}\Gamma(1)/\Gamma(1/2) = \sqrt{2/\pi}$, with Z a standard normal random variable, and where $\Gamma(\cdot)$ denotes the gamma function. Another useful estimator is truncated bipower variation ($TBPV_t$), which combines the truncation method proposed by Mancini (2009) and the bipower variation (BPV_t) estimator discussed above. Namely,

$$TBPV_t = (\mu_1)^{-2} \sum_{i=2}^{n} |\overline{r}_{i,n}||\overline{r}_{i-1,n}|, \quad \overline{r}_{i,n} = r_{i,n} 1_{\{|r_{i,n}| < \alpha_n\}}, \tag{2.10}$$

where $\alpha_n = \alpha \Delta_n^{\varpi}, \varpi \in (0, \frac{1}{2})$. Similarly, truncated realized variance (TRV_t) is defined as

$$TRV_t = \sum_{i=1}^{n} \overline{r}_{i,n}^2. \tag{2.11}$$

Finally, jump variation (JV_t) can be estimated as $JV_t = RV_t - BPV_t$ or $JV_t = RV_t - TBPV_t$, for example. In the sequel, we shall utilize RV_t, TRV_t, BPV_t and $JV_t = RV_t - BPV_t$ when specifying state space models in order to construct mixed frequency factors, as well as when directly including estimates of quadratic variation in the factor augmented regression models used in our forecasting experiments.

Under certain regularity conditions (refer to the above cited papers, Jacod & Protter (2011), and Aït-Sahalia & Jacod (2014) for details), BPV_t, $TBPV_t$ and TRV_t are all consistent estimators of the integrated volatility $\mathrm{IV}_t := \int_0^t \sigma_s^2 ds$. Hence, the corresponding JV_t estimators are also consistent. Moreover, it is also well-established that these estimators converge stably in law at the rate $\sqrt{1/\Delta_n}$, or equivalently, $\sqrt{n}$. Let T be the total number of such representative finite time horizons $[0, t]$ (e.g., day, week, month or quarter). If $\Delta_n T \to 0$, then the impact of estimating the mixed frequency volatility factor based on jump variation are asymptotically negligible, since the parameters in our state space model converge at rate $\sqrt{T}$.

2.3 Mixed frequency financial volatility (MFV) factors

We utilize standard state space model to extract our volatility mixed frequency financial volatility (MFV) factors. The model that we implement is closest to that used in Chauvet et al. (2015), and also follows Aruoba et al. (2009a), although the latter authors do not consider volatility measures in their analysis. While Chauvet et al. (2015) implement a very interesting strategy for extracting a latent volatility factor from various different realized stock and bond volatility measures, we instead focus solely on S&P500 returns in our analysis and consider a model incorporating different frequencies of volatility. In this sense, the structure of our model resembles that of a heterogeneous autoregressive realized volatility type model of the variety introduced in Corsi (2009) and Corsi & Renò (2012). Our models, thus, are meant to capture the heavy persistence in volatility. Moreover, we consider different volatility estimators, including RV_t, TRV_t, BPV_t, and JV_t.

To be more specific, let the dependent variable $y_t = \{y_t^1, y_t^2, y_t^3, y_t^4\}$ in our observation equation represent data measured at 4 different time horizons, including daily (denoted by d), bi-daily (denoted by $2d$), tri-daily (denoted by $3d$), and weekly (denoted by w). Using our mixed frequency approach, we construct four MFV factors, for each of RV_t, TRV_t, BPV_t, and JV_t, respectively. For instance, when build the MFV factor based on RV_t, dependent variable y_t in the observation equation would be: $\{RV_t^d, RV_t^{2d}, RV_t^{3d}, RV_t^w\}$. In model setup, we denote our mixed frequency factor as MFV_t, which is the latent factor to be extracted from the state space model. Lastly, we include three aggregated state variables, i.e., C_t^1, C_t^2 and C_t^3, to address the inter-temporal aggregation issues as discussed in Section 2.1.[9] The state space model is:

Observation Equation:

$$\begin{pmatrix} y_t^d \\ y_t^{2d} \\ y_t^{3d} \\ y_t^w \end{pmatrix} = \begin{pmatrix} \beta_1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \beta_2 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \beta_3 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \beta_4 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} MFV_t \\ C_t^1 \\ C_t^2 \\ C_t^3 \\ u_t^1 \\ u_t^2 \\ u_t^3 \\ u_t^4 \end{pmatrix} \tag{2.12}$$

[9]It is worth stressing that our "C" state variables "handle" flow variables at lower frequencies. More specifically, when the state space framework evolves at higher frequencies (such as at a daily frequency as in our model), the two adjacent releases of the lower frequency flow variable represent the cumulative change over the period between the releases. To ensure that this change is smoothed out and distributed evenly across the days between the two release dates, we use "C" state variables, as suggested in Aruoba et al. (2009b), which is an alternatively approach for including lagged terms. These cumulative "C" state variables represent cumulative sums of the daily latent factor values over the period between two low-frequency variable releases, such as weekly and quarterly. Thus, they share the same error term, e_t^1, as the latent factor. For more detailed information on cumulative state variables, refer to the technical note published by the Philadelphia Federal Reserve Bank: https://www.philadelphiafed.org/-/media/frbp/assets/surveys-and-data/ads/ads-technical-documentation.pdf

State Equation:

$$
\begin{pmatrix} MFV_{t+1} \\ C^1_{t+1} \\ C^2_{t+1} \\ C^3_{t+1} \\ u^1_{t+1} \\ u^2_{t+1} \\ u^3_{t+1} \\ u^4_{t+1} \end{pmatrix} = \begin{pmatrix} \rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho & \psi^1_{t+1} & 0 & 0 & 0 & 0 & 0 & 0 \\ \rho & 0 & \psi^2_{t+1} & 0 & 0 & 0 & 0 & 0 \\ \rho & 0 & 0 & \psi^3_{t+1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \eta_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \eta_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \eta_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \eta_4 \end{pmatrix} \begin{pmatrix} MFV_t \\ C^1_t \\ C^2_t \\ C^3_t \\ u^1_t \\ u^2_t \\ u^3_t \\ u^4_t \end{pmatrix}
$$

$$
+ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^1_t \\ e^2_t \\ e^3_t \\ e^4_t \\ e^5_t \end{pmatrix},
$$

where the error terms follow a multivariate normal distribution:

$$
\begin{pmatrix} e^1_t \\ e^2_t \\ e^3_t \\ e^4_t \\ e^5_t \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_1 & 0 & 0 & 0 \\ 0 & 0 & \sigma^2_2 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_3 & 0 \\ 0 & 0 & 0 & 0 & \sigma^2_4 \end{pmatrix} \right]
$$

As mentioned above, the three aggregated variables in the state vector, C^1_t, C^2_t and C^3_t, are designed to handle bi-daily, tri-daily and weekly updating of our volatility series, respectively. Also, ψ_1, ψ_2 and ψ_3 are binary-valued parameters for the aggregated state variables, and are defined as follows:

$$\psi_t^1 = \begin{cases} 0, & \text{if } t \text{ is an odd number} \\ 1, & \text{otherwise,} \end{cases},$$

for the bi-daily updating series;

$$\psi_t^2 = \begin{cases} 0, & \text{if } t \text{ is the first day of every three days} \\ 1, & \text{otherwise,} \end{cases},$$

for the tri-daily updating series; and

$$\psi_t^3 = \begin{cases} 0, & \text{if } t \text{ is the first day of every week} \\ 1, & \text{otherwise,} \end{cases}$$

for the weekly series.

In the above observation equation, only the highest frequency variable, y_t^d, is directly connected with the factor via β_1. The three other volatility variables are connected with MFV_t via the aggregated state variables (i.e, C_t^1, C_t^2 and C_t^3) and via the parameters β_2, β_3 and β_4. Coupled with the setup of the binary-valued parameters (i.e., ψ_1, ψ_2 and ψ_3) in the state equation, this ensures the proper inter-temporal aggregation of the flow variables in the system and refreshes the quantity at the beginning of each period. Finally, the u_t are stochastic disturbance terms, and are assumed to follow autoregressive processes, as in Aruoba et al. (2009a). In the state equation, the first four state variables are connected with MFV_t via ρ. Of these four state variables, the last three (i.e., C_t^1, C_t^2 and C_t^3) are defined such that their previous values are added to ρMFV_t whenever flow aggregation is required.

2.4 Mixed frequency macroeconomic (MAC) factors

We again begin with $y_t = (y_t^1, y_t^2, y_t^3, y_t^4)$. In this section, these variables, however, are measured at daily (denoted by d), weekly (denoted by w), monthly (denoted by m), and quarterly (denoted by q) frequencies. This setup immediately allows us to construct a "benchmark" mixed frequency macroeconomic (MAC) factor corresponding to the business conditions index analyzed by Aruoba et al. (2009a). In particular, following Aruoba et al. (2009a), we use four macroeconomic variables with different sampling frequencies, including: (1) the daily yield curve spread (y_t^1), defined as the difference between the 10-year U.S. Treasury note yield and the 3-month Treasury bill yield; (2)

weekly initial claims for unemployment insurance (y_t^2); (3) nonfarm payroll employment (y_t^3); and (4) quarterly gross domestic product (y_t^4). The corresponding state space model used to extract MAC factor is:

Observation equation:

$$\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \end{pmatrix} = \begin{pmatrix} \beta_1 & 0 & 0 & 1 \\ 0 & \beta_2 & 0 & 0 \\ \beta_3 & 0 & 0 & 0 \\ 0 & 0 & \beta_4 & 0 \end{pmatrix} \begin{pmatrix} MAC_t \\ C_t^1 \\ C_t^2 \\ u_t^1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ \gamma_2 & 0 & 0 \\ 0 & \gamma_3 & 0 \\ 0 & 0 & \gamma_4 \end{pmatrix} \begin{pmatrix} y_{t-w}^2 \\ y_{t-m}^3 \\ y_{t-q}^4 \end{pmatrix} + \begin{pmatrix} 0 \\ w_t^2 \\ w_t^3 \\ w_t^4 \end{pmatrix}. \tag{2.13}$$

State equation:

$$\begin{pmatrix} MAC_{t+1} \\ C_{t+1}^1 \\ C_{t+1}^2 \\ u_{t+1}^1 \end{pmatrix} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ \rho & \psi_{t+1}^1 & 0 & 0 \\ \rho & 0 & \psi_{t+1}^2 & 0 \\ 0 & 0 & 0 & \gamma_1 \end{pmatrix} \begin{pmatrix} MAC_t \\ C_t^1 \\ C_t^2 \\ u_t^1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e_t^1 \\ e_t^2 \end{pmatrix}, \tag{2.14}$$

where the error terms $e_t^i \overset{i.i.d}{\sim} N(0, \sigma_i^2)$, with $i = 1, 2$.

The variables in this model include: the observed vector y_t; our MAC factor MAC_t; aggregate state variables, C_t^1 and C_t^2; and stochastic disturbance terms, u_t^1, w_t^2, w_t^3, and w_t^4. Note that in this model, only y_t^2 and y_t^4 are flow variables, and hence there are only two aggregate state variables. Accordingly, we also define two binary-valued variables, ψ_1 and ψ_2, for these aggregated state variables. Namely,

$$\psi_t^1 = \begin{cases} 0, & \text{if } t \text{ is the first day of the week} \\ 1, & \text{otherwise,} \end{cases}$$

and

$$\psi_t^2 = \begin{cases} 0, & \text{if } t \text{ is the first day of the quarter} \\ 1, & \text{otherwise.} \end{cases}$$

2.5 Mixed frequency financial volatility and macroeconomic (MIX and SMIX) factors

In order to construct our third type of mixed frequency factor, we use both macroeconomic and volatility variables in the factor construction. The basic notion behind this mixed frequency factor is that "mixing" both types of data (i.e., high frequency financial and mixed frequency macroeconomic data) may yield a more complete picture of the interaction between risks directly affecting macroeconomic variables, and risks that are transmitted through financial market volatility. Namely, we are interested in ascertaining the usefulness of combining mixed frequency factors of the variety analyzed by Bloom (2009) with those analyzed by Chauvet et al. (2015), as well as Aruoba et al. (2009a).

We begin with $y_t = (y_t^1, y_t^2, y_t^3, y_t^4, y_t^5)$. Here, y_t^1 is alternatively set equal to daily RV_t, TRV_t, BPV_t, or JV_t. The rest of the observed variables in our model are the same as those used when constructing MAC. The MIX factor extracted in this setup depends on the definition of y_t^1. Namely, we first extract MIX factors MIX_t for each of y_t^1 equal to RV_t, TRV_t, BPV_t or JV_t, respectively; and second, square root mixed frequency financial volatility and macroeconomic factors $SMIX_t$ for each of y_t^1 equal to the square root of RV_t, TRV_t, BPV_t, or JV_t, respectively. The state space model is:

Observation equation:

$$\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \\ y_t^5 \end{pmatrix} = \begin{pmatrix} \beta_0 & 0 & 0 & 0 & 1 \\ \beta_1 & 0 & 0 & 1 & 0 \\ 0 & \beta_2 & 0 & 0 & 0 \\ \beta_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_4 & 0 & 0 \end{pmatrix} \begin{pmatrix} MIX_t \\ C_t^1 \\ C_t^2 \\ u_t^1 \\ u_t^0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \gamma_2 & 0 & 0 \\ 0 & \gamma_3 & 0 \\ 0 & 0 & \gamma_4 \end{pmatrix} \begin{pmatrix} y_{t-w}^3 \\ y_{t-m}^4 \\ y_{t-q}^5 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ w_t^2 \\ w_t^3 \\ w_t^4 \end{pmatrix}. \tag{2.15}$$

State equation:

$$\begin{pmatrix} MIX_{t+1} \\ C_{t+1}^1 \\ C_{t+1}^2 \\ u_{t+1}^1 \\ u_{t+1}^0 \end{pmatrix} = \begin{pmatrix} \rho & 0 & 0 & 0 & 0 \\ \rho & \psi_{t+1}^1 & 0 & 0 & 0 \\ \rho & 0 & \psi_{t+1}^2 & 0 & 0 \\ 0 & 0 & 0 & \gamma_1 & 0 \\ 0 & 0 & 0 & 0 & \gamma_0 \end{pmatrix} \begin{pmatrix} MIX_t \\ C_t^1 \\ C_t^2 \\ u_t^1 \\ u_t^0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e_t^1 \\ e_t^2 \\ e_t^3 \end{pmatrix}, \tag{2.16}$$

where the error terms $e_t^i \overset{i.i.d}{\sim} N(0, \sigma_i^2)$, with $i = 1, 2, 3$.

The variables in this model include observed variables, the y_t; our latent mixed frequency factor, MIX_t; aggregate state variables, C_t^1 and C_t^2; and stochastic disturbance terms, u_t^1, u_t^0, w_t^2, w_t^3, and w_t^4. As above, only y_t^2 and y_t^4 are flow variables in this model, and hence there are only two aggregate state variables. Accordingly, we also define two binary-valued variables ψ_1 and ψ_2 for these aggregated state variables. Namely,

$$\psi_t^1 = \begin{cases} 0, & \text{if t is the first day of the week} \\ 1, & \text{otherwise} \end{cases}$$

and

$$\psi_t^2 = \begin{cases} 0, & \text{if t is the first day of the quarter} \\ 1, & \text{otherwise} \end{cases} .$$

3 Experimental Setup

All of our prediction experiments are based on sample periods from January 2006 - December 2018 (*Sample* 1) and January 2009 - December 2018 (*Sample* 2). When constructing forecasting models, in-sample model estimation is carried out using a rolling window estimation scheme, with window lengths of $w = 36$ and $w = 72$.[10,11] Monthly forecasts for 6 macroeconomic and 7 corporate bond yield variables (see Section 4) are constructed for $h = 1, 2, 3, 4, 5, 6$ months ahead, with ex-ante prediction periods beginning in January 2012 (under *Sample* 1) and January 2015 (under *Sample* 2). In the remaining sub-sections, we describe the forecasting models as well as evaluation metrics used in our prediction experiments.

[10]For a discussion of the use of alternative windowing schemes in the context of forecasting, see Clark & McCracken (2009) and Hansen & Timmermann (2012) and Rossi & Inoue (2012).

[11]Recall that our latent uncertainy measures are extracted from state space models. These models are estimated using all data up until the period prior to the construction of each prediction, because of this the mixed frequency factors appearing under *Sample* 1 and *Sample* 2 in our prediction experiments are different, regardless of the fact that rolling windows are used in the estimation of the autoregressive type forecasting models described in this section.

3.1 Forecasting Models

Autoregressive benchmark model

The benchmark is an autoregression of order p (i.e., an AR(p) model) specified as follows:

$$y_{t+h} = c + \alpha' W_t + \epsilon_{t+h}, \tag{3.1}$$

where y_t is the "target" forecast variable of interest, h denotes the forecast horizon, W_t contains lags of y_t, and α is a conformably defined coefficient vector. Lag orders are chosen anew, prior to the construction of each monthly forecast, using either the Akaike Information Criterion (AIC) or the Schwarz Information Criterion (SIC). Tabulated mean square forecast errors (MSFEs) reported in the sequel are for the case where lag orders are selected using the AIC. Results for the SIC case are qualitatively the same and available upon request.[12]

Autoregressive models with one mixed frequency factor

Let F_t denote one of the latent mixed frequency factors (i.e., MAC, MFV, MIX, or SMIX), and estimate the following model:

$$y_{t+h} = c + \alpha' W_t + \rho' F_t + \epsilon_{t+h}. \tag{3.2}$$

All terms in this model are as defined above.

Autoregressive models with two mixed frequency factors

Let F_t^a denote MAC and F_t^b denote one of the each type of MFV factors. As these factors do not contain MIX they can be directly compared with models that only include MIX, in order to ascertain whether combination of high frequency financial and mixed frequency macroeconomic data is preferred to the use of mixed frequency factors constructed separately using each variety of dataset. Interestingly, as shall be seen later, models with MIX are always preferred, based on our predictive

[12] A referee rightly suggested that our AR benchmark is only one of many alternative benchmarks of potential interest. In light of this fact, we also carried out the analysis in the sequel using distributed lag models (DLMs), each of which contained three variables selected using the SIC. Our findings based on DLMs were qualitatively the same as those reported here, and are available upon request from the authors.

accuracy assessments. This model is specified as follows:

$$y_{t+h} = c + \alpha' W_t + \rho' F_t^a + \rho'' F_t^b + \epsilon_{t+h}. \tag{3.3}$$

All terms in this model are as defined above.

Autoregressive models with daily volatility measures

In addition to estimating models with our mixed frequency factors, we also estimated models using standard daily quadratic variation component measures, including RV_t, TRV_t, BPV_t, and JV_t. Let X_t denote one of these volatility measures, as well as lags thereof. We estimate the following volatility augmented forecasting model:

$$y_{t+h} = c + \alpha' W_t + \gamma' X_t + \epsilon_{t+h}. \tag{3.4}$$

All terms in this model are as defined above.

Autoregressive models with mixed frequency factors and daily volatility measures

We also combine latent volatility mixed frequency factors and quadratic variation terms using the following model:

$$y_{t+h} = c + \alpha' W_t + \rho' F_t + \gamma' X_t + \epsilon_{t+h}, \tag{3.5}$$

All terms are as defined above, other than F_t, which includes only MFV in this model.

3.2 Predictive Accuracy Assessment

As mentioned above, although some components in the vector process y are estimates obtained from high frequency data, as long as the sampling interval length Δ_n shrinks to zero fast enough, their associated estimation errors have asymptotically negligible effects on the parameters of interest in our setup. Alternatively, we can view our high frequency estimators as observed quantities associated with the latent factors in the state equation of our state-space models. Then, high frequency estimation errors are naturally embedded in the residuals of the observation equation. Since these high frequency estimation errors converge to zero, they are bounded in probability, and hence satisfy our standard assumption on the residuals of the processes that we have specified.

Forecasts made in the sequel are analyzed using point mean square forecast error (MSFE) and directional predictive accuracy (DPAR) measures, as well as associated tests including: (i) the Giacomini and White (GW) test (see Giacomini & White (2006)), which is a conditional Diebold-Mariano (DM) predictive accuracy test (see Diebold & Mariano (1995)); (ii) the well-known chi-square test of independence; and (iii) model confidence sets, which are groups of models that contain the "best" model, with a given level of confidence (in the sequel, we discuss 90% confidence sets - see Hansen et al. (2011) for further details).

Recall that the null hypothesis of the DM test is: $H_0 : \mathrm{E}[L(\epsilon_{t+h}^{(1)})] - \mathrm{E}[L(\epsilon_{t+h}^{(2)})] = 0$, where the $\epsilon_{t+h}^{(i)}$ are prediction errors associated with model i, for $i = 1, 2$. In our analysis, $L(\cdot)$ is a quadratic loss function, and the test statistic that we utilize is:

$$\mathrm{DM}_P = P^{-1} \sum_{t=1}^{P} \frac{d_{t+h}}{\hat{\sigma}_{\bar{d}}}, \tag{3.6}$$

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.[13] If the DM_P statistic is significantly negative, then Model 1 is preferred to Model 2. In the sequel, we assume that the test statistic is asymptotically normal following Giacomini & White (2006). For an interesting discussion of alternative approaches to assessing forecasting performance, see Rossi & Sekhposyan (2011).

Additionally, in the sequel we construct DPAR rates using contingency tables, as in Swanson & White (1995). In particular, we utilize contingency table associated with directional predictions, as follows:

[13] Giacomini & White (2006) also discuss a Wald version of this test statistic, which we do not utilize in this paper.

		actual	
		down	up
predicted	down	d_1	d_2
	up	d_3	d_4

Here d_1 (d_4) is the number of correct forecasts of downward (upward) movements and d_2 (d_3) is the number of incorrect forecasts of downward (upward) movements. Define $P_1 = d_1 + d_3$, $P_2 = d_2 + d_4$, and $P = P_1 + P_2$. Also, DPAR $= (d_1 + d_4)/N$. The null hypothesis is that there is independence between actual and predicted directions, which can be expressed as the hypothesis that a given model is of no value in forecasting the direction of change of a given target variable. To test this hypothesis, we use the chi-square test of independence (see e.g., Pesaran & Timmermann (1994)).[14]

4 Data

Our data span the period from January 03, 2006 to December 31, 2018, and include financial asset transaction prices, macroeconomic variables, and bond yields.

A number of our models utilize daily nonparametric volatility estimators of components of the quadratic variation of the S&P500, in addition to various daily macroeconomic and financial variables. These volatility estimators are constructed using 5-minute SPY (SPDR S&P 500 ETF Trust) transaction prices, which are collected from the NYSE Trade and Quote (TAQ) database[15].

Our macroeconomic variables and bond yields are obtained from the FRED-MD database maintained by the Federal Reserve Bank of St. Louis. More specifically, the following macroeconomic variables were collected for use in our state space models: (1) daily yield curve spread, defined as the difference between the 10-year U.S. Treasury note yield and the 3-month Treasury bill yield; (2) weekly initial claims for unemployment insurance; (3) monthly number of nonfarm payroll employees;

[14] In this paper, we take a frequentist approach to assessing the accuracy of our forecasts. However, it is important to note that there are other interesting approaches for assessing the predictions made in our experiments, as discussed in Aastveit et al. (2014), where density nowcasts from various model classes are combined to construct density nowcasts for US GDP, as well as Nakajima & West (2013), McAlinn & West (2019), McAlinn et al. (2020), and the references cited therein.

[15] https://wrds-www.wharton.upenn.edu/pages/about/data-vendors/nyse-trade-and-quote-taq/.

and (4) quarterly gross domestic product. All of these variables are log differenced in all calculations in order to induce stationarity, and then standardized, with the exception of yield spreads.

Additionally, the following monthly macroeconomic variables are used as "target" variables in our forecasting experiments: industrial production (IP), the monthly number of total nonfarm payroll employees (PAY), housing starts (HS), personal consumption expenditures (PCE), the University of Michigan consumer sentiment index (SI), and core consumption price index (CPI), which excludes food and energy. The first three of these variables are most closely related to firm level business spending and residential investment activities, while the latter three most closely reflect consumer spending activity. All of these variables are also log differenced in all calculations in order to induce stationarity, except for the housing starts and sentiment index as suggested by the FRED-MD database appendix to respectively take log and perform first order difference.

Finally, we construct a corporate bond yield dataset and extract a second set of "target" variables, which we forecast. These variables include monthly bond yields for Fitch-rated AAA, AA, A, BBB, BB, B, and CCC bonds (see Table 1 for details). The source of this dataset is the ICE Benchmark Administration Limited (IBA), and they are available from FRED database. Details and transformations of macroeconomic and financial variables used in our experiments are given in Table 1.

Before turning to a discussion of our empirical findings, it should be stressed that our subsequent analysis in which we forecast bond yields is "real-time", in the sense that the data that we employ are never revised. On the other hand, our experiments in which we forecast macroeconomic variables are only pseudo real-time, in the sense that we use "final-revised data" which in many cases were not the data available at the time that our forecasts were made, as initially available data macroeconomic are generally subject to revisions over time. Thus, our macroeconomic forecasting results should be viewed with caution. Further research with real-time datasets, once they become available for the variables that we model will yield further insights into the usefulness of our mixed frequency factors.

5 Empirical Findings

Our experimental findings are summarized using MSFE and DPAR statistics, and inference on these statistics is carried out using GW tests, chi-square tests of independence, and model confidence sets. For forecast model construction, two data sub-samples are utilized, and are called *Sample* 1 (2006:1 - 2018:12) and *Sample* 2 (2009:1 - 2018:12). Corresponding ex-ante prediction periods are 2012:1 - 2018:12 and 2015:1 - 2018:12, respectively; and forecasts are constructed for $h = 1, 2, 3, 4, 5, 6$ and $w = 36, 72$. For complete details refer to Section 3. The variables used when constructing forecast models are listed in Table 1.[16]

The target variables that we forecast include IP, PAY, HS, PCE, SI, CPI, AAA, AA, A, BBB, BB, B, and CCC (see Section 4 for complete details). Additionally, all of the forecasting models analyzed in our experiments are summarized in Table 2. Turning to our experimental findings, note that Tables 3A and 3B contain MSFE results, while Tables 3C and 3D contain DPAR results, for HS. In these 4 tables, MSFE-"best" and DPAR-"best" models are denoted in bold font, and starred entries denote rejections of the DM-test null hypothesis of equal model accuracy (for Tables 3A and 3B) and rejection of the independence null hypothesis (for Tables 3C and 3D).[17] For the sake of brevity, results for our other 5 macroeconomic target variables are gathered in the supplemental appendix, as are all model confidence set results. Table 4 contains MSFE results from forecasting our 6 corporate bond yields. Of note is that only results for models that performed amongst the top 4 when predicting our macroeconomic variables are reported on in this table - again for the sake of brevity. Finally, Tables 5A to 5D report the overall "winners" in our experiments, by listing the MSFE- and DPAR-"best" models for each of the six forecast horizons and two estimation window sizes.

Prior to discussing our tabulated results, and as an aid to understanding the difference between the different mixed frequency factors in our analysis, consider Figures 1 - 4, in which all of the mixed frequency factors utilized in our experiments are plotted using *Sample* 1. A key take-away form inspection of these figures is that the MAC factor plotted in Figure 1 is very different from

[16]High frequency S&P500 returns are also utilized in our experiments, as discussed above.

[17]For GW tests, each forecasting model listed in the table is compared with an AR benchmark model.

the other factors, which re plotted in Figures 2 - 4. Additionally, and as expected, uncertainty of our factor estimates increases during periods of economic duress. For example, note that in all 4 figures, confidence intervals around our factor estimates are widest around the 2008 financial crisis. Additionally, note that in Figure 1, we include plots of variants of our mixed frequency macroeconomic factor where we alternately drop daily and then daily and weekly data from our model. As might be expected, the factors exhibit increasing volatility as higher frequency variables are dropped.

5.1 Macroeconometric Forecasting Results

5.1.1 Variables most closely related to firm level business spending and residential investment decisions

The variables in this category include housing starts (HS), industrial production (IP), and nonfarm payroll employment (PAY), all of which are related to a firm's decisions on business spending and residential investment. As discussed above, MSFE, DPAR, and model confidence set results for these variables are summarized in Tables 3A - 3D, Tables 5A - 5D, and in the supplemental appendix. Our findings can be summarized as follows.

First, for HS and IP, out-of-sample MSFEs of various factor augmented models show significant reductions, relative to the AR benchmark model. Examples of models performing notably well include RV, TRV, BPV, and JV, for example. All of these models utilize nonparametric quadratic variation measures when estimating the latent factors. Consider the results for HS in Tables 3A to 3D). For *Sample* 1, the use of the BPV model results in MSFE decreases (relative to the AR benchmark) of 10.3% when h=2 and w=36, and 22.1%, when h=2 and w=72. Notice that the longer rolling window yields substantially lower MSFEs for the BPV model, and that this is our MSFE-best model. This finding characterizes many of our target variables, as evidenced upon inspection of Tables 5A and 5B, in which MSFE-best models are summarized, across all forecast horizons, windows, and sample periods. Interestingly, the maximum MSFE reduction is very high, at 53.7%, and is achieved by the VTRV model, when h=6. For Sample 2, VRV, VTRV, and VBPV mixed frequency factor augmented models again appreciably reduce MSFEs, for h=1 or 2, and for w=36 or 72. Here, RV-based MFV factor augmented models reduce MFSE the most (15.8% when h=2 and w=36). Interestingly, our mixed frequency financial volatility and macroeconomic factor augmented models that utilize MIX also yield forecasting improvement, when w=72. For instance, use of the CMTRV1 model decreases the MSFE by 7.3%, relative to the AR benchmark. Results for IP (see supplemental appendix), are qualitatively the same as those reported for HS, although MSFE reductions are appreciably less, across all horizons, windows, and sample periods. These results provide initial evidence of the usefulness of our proposed mixed frequency factors, when forecasting a variety of macroeconomic variables.

Second, for PAY, use of our macroeconomic mixed frequency factor (i.e., MAC) as well as square root mixed frequency financial volatility and macroeconomic factors (i.e., $SMIX$ - see models CMJV2, CMTRV2, CMBPV2, and CMJV2 in Table 2), lead to substantial MSFE reductions for both our shorter and longer sample periods. For example, for *Sample* 1, the use of CMRV2, CMTRV2, and CMBPV2 results in MSFE reductions of 8.8%, 8%, and 3.2%, respectively, when h=1 and w=36; and the use of CMTRV2 and CMBPV2 models result in MSFE decreases of 9.2% and 8.4%, respectively, when h=1 and w=72. These results carry over to the case where *Sample* 2 is used, in which case MSFE reductions are all greater than 10% (i.e., MSFE decreases are 17.2%, 15.6%, and 9.8%, for w=36, and MSFE decreases are 7.8%, 13.6%, and 11.6% for w=72, respectively). This result suggests that MAC or MFV factors can be improved upon, in certain cases, by utilizing our factors that include both high frequency financial data as well as mixed frequency macroeconomic data.

Third, DPAR and model confidence set results are also promising, indicating significant predictive accuracy gains. Results are again particularly promising when using MAC and MIX/SMIX factors (i.e., see models CMJV1, CMTRV1, CMBPV1, CMJV1, CMJV2, CMTRV2, CMBPV2, and CMJV2). Still, it is worth noting that models associated with the largest directional accuracy rates are not always the same as those associated with the smallest relative MSFEs. For example, for HS, the DPAR-"best" model is CMTRV2, in which the directional forecasting accuracy rate is 76.7% (for h=1, w=72, and *Sample* 2). On the other hand, the analogous "MSFE-best" model is CMJV1. Of course, all of these models still include MIX and SMIX factors.

Finally, inspection of the results gathered in Tables 3 and 4 (and the supplemental appendix) indicate a great number of cases for which our latent factor augmented models dominate the AR benchmark, both based on GW tests and based on chi-square tests of independence. Moreover, the model confidence set results indicate that models that include MFV as well as MIX and SMIX factors contain useful predictive information. For example, MRV, MTRV, MBPV, MJV (all of which include MIX factors) appear in various model confidence sets for HS, IP, and PAY, as do models with MIX factors, such as CMRV1, MCTRV1, and CMBPV1.

Summarizing, our new latent mixed frequency factors are very useful in reducing MSFE and increasing directional predictive accuracy. Moreover, our MIX and SMIX factors that from state space models that include both mixed frequency macroeconomic variables as well as nonparametric quadratic variation measures based on high frequency financial data appear to perform the best.

5.1.2 Variables most closely related to consumer spending decisions

The variables in this category include the consumer sentiment index (SI), the consumer price index (CPI), and personal consumption expenditures (PCE). Results from our prediction experiments using these variables are gathered in Tables 5A - 5D, as well as in the supplemental appendix. Our findings can be summarized as follows.

First, the "MSFE-best" models often include factor augmented models. However, unlike the case of HS, IP, and PAY, where a large number of augmented models yield lower MSFEs than our benchmark AR model, we only observe occasional MSFE improvements for CPI, PCE, and SI. Still, even in the worst performing scenarios, such as in the case of CPI, there is some indication that mixed frequency factors may be useful. For example, for CPI at the h=5 horizon, the MRV model results in MSFE reductions of 9.6% (for w=36) and 5.1% (for w=72) in *Sample* 1 and MSFE reductions of 4.4% (w=36) and 5.4% (w=72) in *Sample* 2. Additionally, for SI at the h=1 horizon, the MBPV model results in MSFE reductions of 6.6% and 14.7%, for *Sample* 1 and *Sample* 2, respectively, when w=72.

Second, directional forecast accuracy rates are comparable to rates achieved for our business spending and residential investment variables, when factor augmented models are utilized for directional prediction. Still, as evidenced in Tables 5C and 5D, AR models do sometimes yield the highest directional forecast accuracy rates. Given that this also occurs when predicting the direction of change for HS, IP, and PAY, we have evidence that mixed frequency factors are more useful for predicting absolute magnitudes of our variables than turning points. Drilling down into our findings more deeply, note that for CPI, our RV, TRV, and BPV models, as well as our MRV, MTRV, and MBPV models generally result in around 5% to 6% increases in directional predictive accuracy, relative to the AR benchmark, when h=5 and w=72, for both *Sample* 1 and *Sample* 2. For PCE, the MRV, MTRV, and MBPV models yield increases in directional accuracy of comparable (and greater) magnitudes, when h=5 and w=72, in both sample periods.

Finally, it is worth noting that model confidence sets still include a variety models with our latent mixed frequency factors. For example, the confidence set for PCE only includes MVJV, while the confidence set for CSI also only includes MVJV (when $w = 72$), but includes virtually all of our models with mixed frequency financial volatility and macroeconomic factors when $w = 36$. Thus, although a little weaker, we again have evidence of the usefulness of new mixed frequency factors introduced in this paper.

5.2 Corporate Bond Yield Forecasting Results

In order to ascertain whether our above findings apply to other datasets, we investigated the importance of our new measures in the context of predicting corporate bond yields. In particular, we used a subset of 4 of the very best models from our macroeconomic variables prediction experiment for forecasting yields on AAA, AA, A, BBB, BB, and B rated bonds. Results from this experiment are gathered in Table 4 and Tables 5A - 5D. Our findings can be summarized as follows.

First, it is very clear upon inspection of the MSFEs in Table 4 that predictive accuracy associated with the use of our mixed frequency factors increases as the quality of the bond deteriorates. The greatest gains are associated with junk bonds, while there is little to gain by using mixed frequency factors when predicting AAA rated bonds. Take the case where $w = 36$ as an example, which is reported in Table 4. Bonds with B and CCC ratings show the largest MSFE reductions from amongst all bonds, when the MFV factor augmented models are utilized. For example, for CCC-rated bond yield forecasting, the TRV model results in MSFE reductions of 9.3%, 22.4%, and 13.9%, for h=4 to 6, respectively, in *Sample* 1; and results in MSFE reductions of 12.9%, 35.6%, and 35.8% in *Sample* 2. For B-rated bonds, the TRV model results in MSFE reductions of 22.5%, 25.3%, and 12.4%, for h=4 to 6, respectively, in *Sample* 2. However, predictive gains deteriorate as the investment quality of the bond increases. For example, for BB rated bonds in *Sample* 2, the TRV model results in MSFE reductions of 16.8%, 11.1%, and 3.3%, for h=4, 5, and 6, respectively. All of these percentages are lower than the corresponding ones for CCC and B-rated bonds. The same result holds when comparing BBB versus BB-rated bonds, and A versus BBB-rated bonds, etc. Thus, we have strong evidence of the usefulness of our mixed frequency financial volatility factors (i.e. MFV) for predicting corporate bond yields that involve substantial financial risk, as might be expected.

Second, notice that the fourth row of entries in each panel of Table 4 summarize results based on the JV model. In this model, the MFV factor augmented is the jump variation type of MFV factor. Results are less than starling in these case, as jump-based MFV factors are of little use when predicting corporate bond yields. Instead, our mixed frequency factors that capture the continuous components of quadratic variation yield the most promising results.

Broadly speaking, the above illustration based on bond forecasting again suggests that our new mixed frequency financial volatility factors are useful for reducing MSFE and increasing directional predictive accuracy for a variety of economic variables.

6 Concluding Remarks

In this paper, we analyze three types of mixed frequency factors, and explore their usefulness in a series of forecasting experiments. The new mixed frequency factors are latent variables extracted from state space models that include multiple different frequencies of macroeconomic and financial variables, as well as non-parametrically estimated components of quadratic variation. The state space models are specified in one of two ways. First, they are specified solely using the latent components of quadratic variation, including continuous and jump component variation measures extracted from high frequency S&P500 data. Alternatively, they are specified using quadratic variation components as well as additional observed variables, including macroeconomic indicators such as interest rates, employment, and production, which are measured at multiple different frequencies. Finally, three types of mixed frequency factors are constructed using (i) high frequency financial data; (ii) mixed frequency macroeconomic data; and (iii) both types of data. Our key findings can be summarized as follows. First, our multi-frequency financial (MFV) and financial-macroeconomic volatility (MIX and SMIX) mixed frequency factors yield significantly improved predictions for a number of variables including housing starts, industrial production and nonfarm payroll, relative to benchmark models including simple autoregressive models, as well as mixed frequency models driven solely by macroeconomic indicators (MAC). Second, the same mixed frequency factors are useful for predicting low-grade corporate bond yields; but not high-grade corporate bond yields, underscores the importance of the investment grade of bonds for withstanding turbulent market conditions, as might be expected. Third, four different measures of volatility are used in our analysis, including realized volatility (RV_t), truncated realized volatility (TRV_t), bi-power variation (BPV_t), and jump variation ($JV_t = RV_t - BPV_t$). In our forecasting experiments, TRV_t is clearly the most effective measure to use when constructing volatility mixed frequency factors. Moreover, factors constructed using JV_t perform quite poorly in our prediction experiments.

In closing, it is worth stressing that much work remains to be done. For example, on the theoretical side it remains to construct a unified model that combines key results from the continuous time finance literature used in this paper with results from the discrete time mixed frequency modeling literature. Additionally, while some of the empirical results presented in this paper are truly “real-time”, other findings are only pseudo real-time, in the sense that information used in rolling forecasts was not available in real time, hence only fully revised data were used. It remains, thus, to carry out some of the analysis done herein with real-time datasets, as they become available.

References

Aastveit, K. A., Gerdrup, K. R., Jore, A. S. & Thorsrud, L. A. (2014), 'Nowcasting gdp in real time: A density combination approach', *Journal of Business & Economic Statistics* **32**(1), 48–68.

Aguilar, O. & West, M. (2000), 'Bayesian dynamic factor models and portfolio allocation', *Journal of Business & Economic Statistics* **18**(3), 338–357.

Aït-Sahalia, Y. & Jacod, J. (2014), *High Frequency Financial Econometrics*, Princeton University Press.

Anderson, B. D. & Moore, J. B. (2012), *Optimal filtering*, Courier Corporation.

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.

Aruoba, S., Diebold, F. & Scotti, C. (2009b), 'Updates on ads index calculation', *Manuscript, www. philadelphiafed. org/research-and-data/real-time-center/business-conditions-index* .

Baker, S. R., Bloom, N. & Davis, S. J. (2016), 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics* **131**(4), 1593–1636.

Barndorff-Nielsen, O. E. & Shephard, N. (2004), 'Power and bipower variation with stochastic volatility and jumps', *Journal of Financial Econometrics* **2**(1), 1–37.

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.

Cargnoni, C., Müller, P. & West, M. (1997), 'Bayesian forecasting of multinomial time series through conditionally gaussian dynamic models', *Journal of the American Statistical Association* **92**(438), 640–647.

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.

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.

Clark, T. E. & McCracken, M. W. (2009), 'Improving forecast accuracy by combining recursive and rolling forecasts', *International Economic Review* **50**(2), 363–395.

Corsi, F. (2009), 'A simple approxiamte long memory model of realized volatility', *Journal of Financial Econometrics* **7**(2), 174–196.

Corsi, F. & Renò, R. (2012), 'Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling', *Journal of Business & Economic Statistics* **30**(3), 368–380.

Diebold, F. X. & Mariano, R. S. (1995), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics* **13**(3), 253–263.

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. L. (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* .

Hansen, P. R., Lunde, A. & Nason, J. M. (2011), 'The model confidence set', *Econometrica* **79**(2), 453–497.

Hansen, P. R. & Timmermann, A. (2012), 'Choice of sample split in out-of-sample forecast evaluation', *European University Institute; Series/Number: EUI ECO; 2012/10* .

Jacod, J. & Protter, P. (2011), *Discretization of Processes*, Springer.

Jo, S. & Sekkel, R. (2017), 'Macroeconomic uncertainty through the lens of professional forecasters', *Journal of Business & Economic Statistics* pp. 1–11.

Jurado, K., Ludvigson, S. C. & Ng, S. (2015), 'Measuring uncertainty', *American Economic Review* **105**(3), 1177–1216.

Mancini, C. (2009), 'Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps', *Scandinavian Journal of Statistics* **36**(2), 270–296.

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.

McAlinn, K. (2021), 'Mixed-frequency bayesian predictive synthesis for economic nowcasting', *Journal of the Royal Statistical Society Series C: Applied Statistics* **70**(5), 1143–1163.

McAlinn, K., Aastveit, K. A., Nakajima, J. & West, M. (2020), 'Multivariate bayesian predictive synthesis in macroeconomic forecasting', *Journal of the American Statistical Association* **115**(531), 1092–1110.

McAlinn, K. & West, M. (2019), 'Dynamic bayesian predictive synthesis in time series forecasting', *Journal of econometrics* **210**(1), 155–169.

Nakajima, J. & West, M. (2013), 'Bayesian analysis of latent threshold dynamic models', *Journal of Business & Economic Statistics* **31**(2), 151–164.

Pesaran, M. H. & Timmermann, A. G. (1994), 'A generalization of the non-parametric henriksson-merton test of market timing', *Economics Letters* **44**(1-2), 1–7.

Prado, R. & West, M. (2010), *Time series: modeling, computation, and inference*, Chapman and Hall/CRC.

Rossi, B. & Inoue, A. (2012), 'Out-of-sample forecast tests robust to the choice of window size', *Journal of Business & Economic Statistics* **30**(3), 432–453.

Rossi, B. & Sekhposyan, T. (2011), 'Understanding models' forecasting performance', *Journal of Econometrics* **164**(1), 158–172.

Swanson, N. R. & White, H. (1995), 'A model-selection approach to assessing the information in the term structure using linear models and artificial neural networks', *Journal of Business & Economic Statistics* **13**(3), 265–275.

West, M. & Harrison, J. (1997), *Bayesian forecasting and dynamic models*, Springer Verlag.

Table 1: Macroeconomic and Financial Variables Used in Factor Construction and in Forecasting Experiments[1]

Name	Frequency	Description	Treatment
SPY	5-Minute	SPDR S&P 500 ETF Trust Price	$\Delta log(x_t)$
SPR	Daily	Yield Curve Spread (10-year Treasury Note Yield Minus 3-month Yield)	no transformation
IC	Weekly	Initial Claims for Unemployment Insurance	$\Delta log(x_t)$
PAY	Monthly	Number of Employees on Non-agricultural Payrolls	$\Delta log(x_t)$
GDP	Quarterly	Real Gross Domestic Product	$\Delta log(x_t)$
IP	Monthly	Industrial Production Index	$\Delta log(x_t)$
HS	Monthly	Housing Starts	$log(x_t)$
PCE	Monthly	Personal Consumption Expenditures	$\Delta log(x_t)$
SI	Monthly	University of Michigan Consumer Sentiment Index	Δx_t
CPI	Monthly	Consumer Price Index Less Food and Energy	$\Delta log(x_t)$
AAA	Monthly	US Corporate AAA Effective Yield	Δx_t
AA	Monthly	US Corporate AA Effective Yield	Δx_t
A	Monthly	US Corporate A Effective Yield	Δx_t
BBB	Monthly	US Corporate BBB Effective Yield	Δx_t
BB	Monthly	US High Yield BB Effective Yield	Δx_t
B	Monthly	US High Yield B Effective Yield	Δx_t
CCC	Monthly	US High Yield CCC or Below Effective Yield	Δx_t

[1] The SPY data used in this study were obtained from the WRDS Trade and Quotes (TAQ) database, while all other series were sourced from the FRED-MD database of the St. Louis Federal Reserve Bank, and have been seasonally adjusted, with bond classifications ranging from AAA to CCC based on standards set by S&P500 and Fitch.

Table 2: Forecasting Models[1]

Model	Description
Benchmark Model:	
AR: Autoregression	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \epsilon_t$
Mixed Frequency Macroeconomic (MAC) Factors Augmented Model:	
MMF: $AR + MAC$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \epsilon_t$
Mixed Frequency Financial Volatility (MFV) Factors Augmented Models:	
RV: $AR + MFV^{RV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_3 MFV_t^{RV} + \epsilon_t$
TRV: $AR + MFV^{TRV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_3 MFV_t^{TRV} + \epsilon_t$
BPV: $AR + MFV^{BPV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_3 MFV_t^{BPV} + \epsilon_t$
JV: $AR + MFV^{JV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_3 MFV_t^{JV} + \epsilon_t$
Mixed Frequency Financial Volatility and Macroeconomic (MIX or SMIX) Factors Augmented Models:	
CMRV1: $AR + MIX^{RV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 MIX_t^{RV} + \epsilon_t$
CMTRV1: $AR + MIX^{TRV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 MIX_t^{TRV} + \epsilon_t$
CMBPV1: $AR + MIX^{BPV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 MIX_t^{BPV} + \epsilon_t$
CMJV1: $AR + MIX^{JV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 MIX_t^{JV} + \epsilon_t$
CMRV2: $AR + SMIX^{\sqrt{RV}}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 SMIX_t^{\sqrt{RV}} + \epsilon_t$
CMTRV2: $AR + SMIX^{\sqrt{TRV}}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 SMIX_t^{\sqrt{TRV}} + \epsilon_t$
CMBPV2: $AR + SMIX^{\sqrt{BPV}}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 SMIX_t^{\sqrt{BPV}} + \epsilon_t$
CMJV2: $AR + SMIX^{\sqrt{JV}}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_2 SMIX_t^{\sqrt{JV}} + \epsilon_t$
Volatility Augmented Models:	
VRV: $AR + RV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \gamma RV_t + \epsilon_t$
VTRV: $AR + TRV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \gamma TRV_t + \epsilon_t$
VBPV: $AR + BPV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \gamma BPV_t + \epsilon_t$
VJV: $AR + JV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \gamma JV_t + \epsilon_t$
MAC and MFV Augmented Models:	
MRV: $AR + MAC + MFV^{RV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \rho_2 MFV_t^{RV} + \epsilon_t$
MTRV: $AR + MAC + MFV^{TRV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \rho_2 MFV_t^{TRV} + \epsilon_t$
MBPV: $AR + MAC + MFV^{BPV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \rho_2 MFV_t^{BPV} + \epsilon_t$
MJV: $AR + MAC + MFV^{JV}$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \rho_2 MFV_t^{JV} + \epsilon_t$
MAC and Volatility Augmented Models:	
MVRV: $AR + MAC + RV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \gamma RV_t + \epsilon_t$
MVTRV: $AR + MAC + TRV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \gamma TRV_t + \epsilon_t$
MVBPV: $AR + MAC + BPV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \gamma BPV_t + \epsilon_t$
MVJV: $AR + MAC + JV$	$y_{t+h} = c + \sum_{i=1}^{r} \alpha_i y_{t-i} + \rho_1 MAC_t + \gamma JV_t + \epsilon_t$

[1] We determine the number of lags, r, for the AR model using the AIC (results for the case where the SIC is instead used are qualitatively the same). We constructed mixed frequency factors following the procedure outlined in Section 2. The mixed frequency macroeconomic factors are denoted as MAC, and the mixed frequency financial volatility factors are denoted as MFV, with the upper corner denoting the type of volatility measurement on which the factor is based. For instance, MFV^{RV}, MFV^{TRV}, MFV^{BPV}, and MFV^{JV} represent mixed frequency financial volatility factors based on realized volatility, truncated realized volatility, bi-power variation, and the jump component of quadratic variation, respectively. Similarly, we use MIX to denote mixed frequency financial volatility and macroeconomic factors with upper corner coding for RV, TRV, BPV, and JV, and SMIX to denote square root mixed frequency financial volatility and macroeconomic factors.

Table 3A: Ex-Ante Relative MSFEs for Housing Starts (Sample 1: 2006:1 - 2018:12)[1]

Model	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
rolling window size = 36						
AR	1.000	1.000	1.000	1.000	1.000	1.000
MMF	1.173***	1.092***	1.104***	1.027	0.981***	**0.938*** **
RV	**0.930*** **	**0.897*** **	0.999***	1.006	1.037	1.038***
TRV	0.930***	0.902***	1.015	1.013	1.043	1.040***
BPV	0.931***	0.897***	1.012	1.002**	1.036	1.035***
JV	2.277***	2.178***	2.010***	1.726***	1.747***	1.602***
CMRV1	0.999***	1.096***	1.161***	1.028***	1.004*	0.968***
CMTRV1	1.019	1.003*	**0.998*** **	1.004*	1.058***	1.047
CMBPV1	0.997***	1.096**	1.158***	1.033***	1.032***	0.967***
CMJV1	1.011	1.133***	1.078***	1.022	0.979***	0.952***
CMRV2	0.992***	0.974***	1.011	**0.982*** **	1.079***	1.031
CMTRV2	0.991***	0.989***	1.055***	1.004*	1.062***	1.044
CMBPV2	1.007	1.020	1.016	0.992***	1.076***	1.059
CMJV2	1.014	0.991***	1.048***	1.031***	**0.962*** **	0.963***
VRV	1.040***	1.022	1.052	1.041***	1.037***	0.983***
VTRV	1.037***	1.021	1.053	1.041***	1.038***	0.985***
VBPV	1.043***	1.023	1.058	1.049***	1.043***	0.982***
VJV	1.044*	0.952***	0.999***	0.998***	0.962***	0.964***
MRV	1.152***	1.090***	1.165***	1.321***	1.194***	1.208***
MTRV	1.151***	1.091***	1.181***	1.336***	1.199***	1.218***
MBPV	1.146***	1.075***	1.160***	1.326***	1.194***	1.193***
MJV	1.253***	1.254***	1.096	1.174***	1.252***	1.286***
MVRV	1.272***	1.197***	1.116***	1.129***	1.005*	1.037
MVTRV	1.272***	1.200***	1.115***	1.132***	1.016	1.040
MVBPV	1.272***	1.196***	1.119***	1.129***	1.011	1.032
MVJV	1.253***	1.156***	1.103***	1.016	1.029	1.085
rolling window size = 72						
AR	1.000	1.000	1.000	1.000	1.000	1.000
MMF	0.925***	0.851***	0.841***	0.700***	0.623***	0.591***
RV	**0.849*** **	0.780***	0.711***	0.569***	0.484***	0.479***
TRV	0.852***	0.798***	0.724***	**0.561*** **	**0.477*** **	**0.463*** **
BPV	0.855***	**0.779*** **	**0.710*** **	0.567***	0.482***	0.477***
JV	1.081	1.061	1.034	0.796***	0.630***	0.559***
CMRV1	0.920***	0.882***	0.841***	0.760***	0.709***	0.741***
CMTRV1	0.889***	0.883***	0.884***	0.832***	0.872***	0.881***
CMBPV1	0.916***	0.877***	0.833***	0.754***	0.704***	0.749***
CMJV1	0.906***	0.892***	0.880***	0.812***	0.756***	0.792***
CMRV2	0.910***	0.922***	0.919***	0.853***	0.875***	0.857***
CMTRV2	0.873***	0.868***	0.856***	0.834***	0.875***	0.812***
CMBPV2	0.866***	0.877***	0.899***	0.866***	0.872***	0.832***
CMJV2	0.922***	0.915***	0.911***	0.805***	0.721***	0.770***
VRV	0.922***	0.836***	0.766***	0.753***	0.642***	0.664***
VTRV	0.917***	0.837***	0.764***	0.753***	0.629***	0.660***
VBPV	0.923***	0.837***	0.764***	0.754***	0.637***	0.667***
VJV	0.938***	0.863***	0.883***	0.794***	0.698***	0.723***
MRV	0.886***	0.806***	0.785***	0.660***	0.607***	0.615***
MTRV	0.887***	0.813***	0.789***	0.668***	0.607***	0.616***
MBPV	0.884***	0.804***	0.782***	0.655***	0.605***	0.610***
MJV	1.011	0.894***	0.801***	0.712***	0.631***	0.658***
MVRV	0.930***	0.823***	0.838***	0.706***	0.627***	0.639***
MVTRV	0.928***	0.825***	0.842***	0.706***	0.627***	0.633***
MVBPV	0.930***	0.820***	0.834***	0.699***	0.626***	0.640***
MVJV	0.920***	0.880***	0.839***	0.727***	0.629***	0.632***

[1] This table displays the mean square forecast errors (MSFEs) relative to the AR benchmark model. The first column indicates the forecasting model (refer to Table 2 for a description of the models). Entries marked with stars denote rejections of the Giacomini & White (2006) test of conditional predictive accuracy, where ***, **, and * indicate rejection at the 1%, 5%, and 10% significance levels, respectively. The entire sample period used in the forecasting experiment is 2006:1-2018:12, and ex-ante rolling window MSFEs correspond to predictions made for the period 2012:1 to 2018:12.

Table 3B: Ex-Ante Relative MSFEs for Housing Starts (Sample 2: 2009:1 - 2018:12)[1]

Model	Forecast horizon					
	1-month	2-month	3-month	4-month	5-month	6-month
			rolling window size = 36			
AR	1.000	1.000	1.000	1.000	1.000	1.000
MMF	1.231***	1.114**	1.166	0.970***	0.967***	0.969***
RV	0.954***	0.842***	0.944***	0.924***	1.193***	1.064
TRV	**0.950*****	0.846***	0.981***	0.938***	1.205***	1.061
BPV	0.958***	0.847***	0.952***	**0.910*****	1.187***	1.050
JV	2.977***	3.567***	3.028***	1.923***	2.313***	2.173***
CMRV1	1.121***	1.191***	1.304***	1.056***	1.147***	0.924***
CMTRV1	1.030	0.993***	**0.927*****	1.074***	1.050***	0.990***
CMBPV1	1.122***	1.181***	1.296***	1.074***	1.149***	**0.920*****
CMJV1	1.121***	1.267***	1.148***	0.988***	1.083***	0.936***
CMRV2	0.958***	0.995***	0.944***	1.023	1.114***	0.986***
CMTRV2	0.994***	0.984***	1.081	1.052***	1.061***	1.009
CMBPV2	1.003*	0.987***	0.970***	1.030	1.098***	0.988***
CMJV2	1.058**	0.969***	1.189***	1.008	0.997***	0.940***
VRV	1.117***	1.145***	1.262***	1.023	1.047	1.007
VTRV	1.107***	1.136***	1.259***	1.021	1.045	1.008
VBPV	1.125***	1.153***	1.279***	1.032	1.061	1.012
VJV	1.137***	**0.818*****	0.989***	1.058***	**0.966*****	0.972***
MRV	1.231***	1.123	1.259***	1.479***	1.474***	1.637***
MTRV	1.232***	1.121	1.316***	1.512***	1.476***	1.652***
MBPV	1.231***	1.099	1.243***	1.489***	1.473***	1.588***
MJV	1.127	0.993***	1.048	1.024	1.212	1.513***
MVRV	1.383***	1.393***	1.226***	1.219**	1.105	1.155
MVTRV	1.379***	1.399***	1.224***	1.224**	1.107	1.156
MVBPV	1.389***	1.400***	1.225***	1.217*	1.102	1.148
MVJV	1.235***	1.245***	1.296***	1.028	1.095	1.345***
			rolling window size = 72			
AR	1.000	1.000	1.000	1.000	1.000	1.000
MMF	1.064	1.043	1.030	1.020	1.050	**0.949*****
RV	0.989***	0.963***	0.998***	1.001**	0.994***	0.971***
TRV	0.985***	0.970***	1.001**	0.974***	0.997***	0.972***
BPV	1.014	0.958***	0.994***	0.996***	0.992***	0.971***
JV	1.420***	1.447***	1.356***	1.239***	1.192***	1.117***
CMRV1	1.037	0.950***	0.978***	0.965***	0.958***	0.980***
CMTRV1	1.033	**0.927*****	1.022***	**0.951*****	0.975***	1.033
CMBPV1	1.043	0.952***	0.980***	0.963***	0.956***	0.962***
CMJV1	**0.938*****	0.958***	1.009	1.049	0.991***	1.079***
CMRV2	0.989***	0.994***	1.028***	0.999***	1.004	1.048
CMTRV2	0.971***	0.974***	1.036***	0.979***	1.004	0.979***
CMBPV2	0.988***	0.947***	1.038***	0.963***	0.984***	1.028
CMJV2	0.955***	0.943***	1.024***	1.036*	0.983***	1.083***
VRV	0.999***	1.008	0.956***	1.017	0.949***	0.984***
VTRV	0.997***	1.002**	0.958***	1.021	0.952***	1.001**
VBPV	1.002*	1.014	**0.955*****	1.016	**0.947*****	0.993***
VJV	1.019	0.961***	1.026	0.998***	1.006	1.032*
MRV	1.106*	1.027	1.049	1.053	1.087	0.982***
MTRV	1.105*	1.032	1.053	1.056	1.091	0.987***
MBPV	1.106*	1.022	1.046	1.049	1.085	0.981***
MJV	1.267***	1.332***	1.188***	1.222***	1.141*	1.144
MVRV	1.147***	0.991***	1.040	0.984***	1.060	0.986***
MVTRV	1.141***	0.990***	1.042	0.985***	1.058	1.009*
MVBPV	1.153***	0.993***	1.040	0.984***	1.057	0.988***
MVJV	1.008*	1.087	0.976***	1.026	1.008*	0.983***

[1] Please refer to the notes in Table 3A. The forecasting experiment covers the entire sample period from 2006:1 to 2018:12, and the ex-ante rolling window MSFEs pertain to predictions made for the period from 2015:1 to 2018:12.

Table 3C: Ex-ante Directional Accuracy Rates for Housing Starts (Sample 1: 2006:1 - 2018:12)[1]

Model	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
			rolling window size = 36			
AR	65.8%***	70.9%***	67.1%***	73.4%***	65.8%***	59.5%*
MMF	62.0%***	65.8%***	65.8%***	69.6%***	67.1%***	64.6%***
RV	65.8%***	**73.4%*****	60.8%*	68.4%***	69.6%***	55.7%
TRV	65.8%***	73.4%***	62.0%**	69.6%***	**70.9%*****	54.4%
BPV	65.8%***	73.4%***	60.8%*	69.6%***	70.9%***	55.7%
JV	48.1%	49.4%	55.7%	57.0%	55.7%	55.7%
CMRV1	64.6%***	69.6%***	59.5%**	70.9%***	63.3%***	64.6%***
CMTRV1	69.6%***	70.9%***	65.8%***	**75.9%*****	67.1%***	58.2%
CMBPV1	64.6%***	69.6%***	59.5%**	70.9%***	63.3%***	64.6%***
CMJV1	67.1%***	68.4%***	67.1%***	70.9%***	63.3%***	62.0%**
CMRV2	68.4%***	69.6%***	**68.4%*****	75.9%***	67.1%***	60.8%**
CMTRV2	**70.9%*****	69.6%***	67.1%***	75.9%***	67.1%***	57.0%
CMBPV2	69.6%***	69.6%***	67.1%***	75.9%***	67.1%***	58.2%
CMJV2	68.4%***	67.1%***	62.0%**	67.1%***	65.8%***	64.6%***
VRV	65.8%***	67.1%***	60.8%**	67.1%***	69.6%***	57.0%
VTRV	63.3%***	67.1%***	60.8%**	67.1%***	69.6%***	55.7%
VBPV	64.6%***	67.1%***	62%**	67.1%***	69.6%***	57.0%
VJV	68.4%***	68.4%***	63.3%**	74.7%***	68.4%***	60.8%**
MRV	59.5%**	63.3%***	62.0%**	68.4%***	67.1%***	62.0%**
MTRV	58.2%*	63.3%***	62.0%**	67.1%***	67.1%***	62.0%**
MBPV	60.8%**	63.3%***	63.3%**	67.1%***	67.1%***	62.0%**
MJV	60.8%**	55.7%	62.0%**	64.6%***	64.6%***	**68.4%*****
MVRV	57.0%*	65.8%***	65.8%***	67.1%***	67.1%***	67.1%***
MVTRV	57.0%*	65.8%***	65.8%***	65.8%***	67.1%***	67.1%***
MVBPV	57.0%*	65.8%***	65.8%***	67.1%***	65.8%***	67.1%***
MVJV	62.0%**	65.8%***	65.8%***	74.7%***	68.4%***	63.3%**
			rolling window size = 72			
AR	72.2%***	69.6%***	63.3%***	58.2%**	48.1%	48.1%
MMF	70.9%***	73.4%***	70.9%***	70.9%***	63.3%***	64.6%***
RV	70.9%***	74.7%***	67.1%***	**79.7%*****	**73.4%*****	68.4%***
TRV	72.2%***	74.7%***	68.4%***	79.7%***	73.4%***	**70.9%*****
BPV	70.9%***	74.7%***	67.1%***	79.7%***	73.4%***	67.1%***
JV	64.6%***	60.8%**	65.8%***	74.7%***	63.3%	68.4%***
CMRV1	72.2%***	73.4%***	**75.9%*****	64.6%***	57.0%*	53.2%
CMTRV1	**74.7%*****	72.2%***	69.6%***	63.3%***	53.2%	57.0%*
CMBPV1	72.2%***	73.4%***	73.4%***	64.6%***	57.0%*	50.6%
CMJV1	74.7%***	72.2%***	69.6%***	58.2%**	53.2%	51.9%
CMRV2	73.4%***	69.6%***	65.8%***	65.8%***	54.4%	58.2%*
CMTRV2	73.4%***	72.2%***	69.6%***	64.6%***	50.6%	55.7%
CMBPV2	73.4%***	72.2%***	69.6%***	65.8%***	51.9%	57.0%*
CMJV2	72.2%***	70.9%***	67.1%***	59.5%***	58.2%**	55.7%
VRV	74.7%***	**78.5%*****	70.9%***	60.8%***	62.0%**	60.8%**
VTRV	74.7%***	78.5%***	70.9%***	60.8%***	62.0%**	60.8%**
VBPV	74.7%***	77.2%***	70.9%***	60.8%***	62.0%**	58.2%*
VJV	73.4%***	75.9%***	68.4%***	60.8%***	60.8%**	60.8%**
MRV	68.4%***	72.2%***	70.9%***	73.4%***	62.0%**	62.0%***
MTRV	68.4%***	74.7%***	70.9%***	73.4%***	63.3%***	60.8%**
MBPV	68.4%***	73.4%***	70.9%***	73.4%***	64.6%***	62.0%***
MJV	63.3%***	72.2%***	67.1%***	70.9%***	62.0%***	58.2%**
MVRV	65.8%***	74.7%***	70.9%***	70.9%***	62.0%**	62.0%***
MVTRV	65.8%***	74.7%***	69.6%***	70.9%***	62.0%***	60.8%**
MVBPV	65.8%***	74.7%***	70.9%***	72.2%***	62.0%**	62.0%***
MVJV	67.1%***	75.9%***	70.9%***	72.2%***	62.0%**	62.0%***

[1] This table reports direction accuracy rates, and starred entries indicate rejections of the directional accuracy test based on the contingency tables discussed in Section 3 and Pesaran & Timmermann (1994). Please refer to the notes in Table 3A for details the sample period and significance levels.

Table 3D: Ex-ante Directional Accuracy Rates for Housing Starts (Sample 2: 2009:1 - 2018:12)[1]

Model	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
			rolling window size = 36			
AR	62.8%	76.7%***	79.1%***	72.1%***	65.1%**	65.1%**
MMF	65.1%	69.8%***	72.1%***	76.7%***	72.1%***	69.8%***
RV	62.8%	**79.1%*****	69.8%***	69.8%***	74.4%***	62.8%**
TRV	62.8%	79.1%***	72.1%***	69.8%***	**76.7%*****	60.5%*
BPV	62.8%**	79.1%***	69.8%***	69.8%***	76.7%***	62.8%**
JV	41.9%	46.5%	53.5%	58.1%	62.8%*	55.8%
CMRV1	62.8%**	69.8%***	67.4%***	74.4%***	60.5%*	74.4%***
CMTRV1	69.8%***	72.1%***	76.7%***	76.7%***	67.4%***	67.4%***
CMBPV1	62.8%**	69.8%***	67.4%***	74.4%***	60.5%*	74.4%***
CMJV1	65.1%**	67.4%***	76.7%***	72.1%***	60.5%*	69.8%***
CMRV2	69.8%***	69.8%***	79.1%***	76.7%***	67.4%***	67.4%***
CMTRV2	**72.1%*****	69.8%***	76.7%***	76.7%***	67.4%***	65.1%**
CMBPV2	69.8%***	72.1%***	**81.4%*****	76.7%***	67.4%***	67.4%***
CMJV2	67.4%**	72.1%***	69.8%***	72.1%***	65.1%**	69.8%***
VRV	65.1%**	69.8%***	69.8%***	72.1%***	72.1%***	62.8%**
VTRV	65.1%**	69.8%***	69.8%***	72.1%***	72.1%***	60.5%*
VBPV	62.8%*	69.8%***	69.8%***	72.1%***	72.1%***	62.8%**
VJV	69.8%***	72.1%***	72.1%***	74.4%***	69.8%***	67.4%***
MRV	58.1%	62.8%**	67.4%***	72.1%***	67.4%***	69.8%***
MTRV	58.1%	62.8%**	67.4%***	69.8%***	67.4%***	69.8%***
MBPV	60.5%*	62.8%**	69.8%***	69.8%***	67.4%***	69.8%***
MJV	60.5%*	60.5%*	67.4%***	67.4%***	69.8%***	72.1%***
MVRV	55.8%	67.4%***	74.4%***	67.4%**	69.8%***	**76.7%*****
MVTRV	55.8%	67.4%***	74.4%***	65.1%**	69.8%***	76.7%***
MVBPV	55.8%	67.4%***	74.4%***	67.4%***	67.4%***	76.7%***
MVJV	62.8%**	67.4%***	72.1%***	**79.1%*****	72.1%***	72.1%***
			rolling window size = 72			
AR	74.4%***	74.4%***	79.1%***	74.4%***	62.8%*	62.8%**
MMF	67.4%***	69.8%***	76.7%***	74.4%***	62.8%*	**69.8%*****
RV	72.1%***	69.8%***	74.4%***	76.7%***	**67.4%*****	67.4%***
TRV	72.1%***	69.8%***	76.7%***	76.7%***	67.4%***	67.4%***
BPV	72.1%***	69.8%***	74.4%***	76.7%***	67.4%***	65.1%***
JV	58.1%**	60.5%*	72.1%***	**79.1%*****	58.1%	65.1%**
CMRV1	**76.7%*****	74.4%***	**81.4%*****	79.1%***	62.8%*	65.1%**
CMTRV1	74.4%***	72.1%***	81.4%***	76.7%***	60.5%	65.1%***
CMBPV1	76.7%***	74.4%***	76.7%***	79.1%***	62.8%*	65.1%**
CMJV1	76.7%***	69.8%***	72.1%***	72.1%***	65.1%**	65.1%***
CMRV2	76.7%***	69.8%***	81.4%***	79.1%***	62.8%*	65.1%***
CMTRV2	76.7%***	69.8%***	81.4%***	76.7%***	58.1%	65.1%***
CMBPV2	76.7%***	72.1%***	81.4%***	79.1%***	60.5%	65.1%***
CMJV2	74.4%***	67.4%***	72.1%***	74.4%***	65.1%**	65.1%***
VRV	76.7%***	**76.7%*****	79.1%***	74.4%***	67.4%***	65.1%**
VTRV	76.7%***	76.7%***	79.1%***	74.4%***	67.4%***	65.1%**
VBPV	76.7%***	74.4%***	79.1%***	74.4%***	67.4%***	60.5%*
VJV	74.4%***	74.4%***	76.7%***	76.7%***	67.4%***	69.8%***
MRV	60.5%*	72.1%***	74.4%***	76.7%***	62.8%*	69.8%***
MTRV	60.5%*	72.1%***	74.4%***	76.7%***	62.8%*	69.8%***
MBPV	60.5%*	72.1%***	74.4%***	76.7%***	65.1%**	69.8%***
MJV	60.5%**	69.8%***	74.4%***	76.7%***	67.4%***	65.1%**
MVRV	60.5%	72.1%***	76.7%***	76.7%***	62.8%*	69.8%***
MVTRV	60.5%	72.1%***	74.4%***	76.7%***	62.8%*	67.4%***
MVBPV	60.5%	72.1%***	76.7%***	79.1%***	62.8%*	69.8%***
MVJV	62.8%**	72.1%***	76.7%***	79.1%***	62.8%*	67.4%***

[1] Please refer to the notes in Table 3B and 3C.

Table 4: Ex-Ante Relative MSFEs for Corporate Bond Yields using MFV Augmented Models[1]

Target	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
Sample 1: 2006:1 - 2018:12						
AR	1.000	1.000	1.000	1.000	1.000	1.000
AAA	1.029	1.028	0.873***	0.896***	0.999***	1.084***
	1.031	1.024	0.873***	0.897***	0.993***	1.085***
	1.030	1.029	0.876***	0.895***	1.000***	1.087***
	1.921***	1.922***	1.706***	1.533***	1.429***	1.601***
AA	1.039	1.001***	0.916***	0.962***	1.030	1.079
	1.039	1.000***	0.921***	0.968***	1.036	1.086
	1.039	1.000***	0.916***	0.959***	1.026	1.077
	1.796***	1.817***	1.922***	2.094***	1.892***	2.090***
A	1.080***	1.007*	0.935***	1.007*	1.054	1.093
	1.081***	1.006*	0.938***	1.010*	1.060	1.098
	1.082***	1.008*	0.914***	1.004**	1.051	1.102
	1.902***	1.877***	2.045***	2.260***	2.301***	2.217***
BBB	1.136***	1.053	0.966***	0.951***	1.001***	1.052
	1.138***	1.055	0.974***	0.959***	1.010*	1.062
	1.137***	1.077***	0.958***	0.943***	0.982***	1.044
	3.204***	2.601***	2.156***	2.441***	3.105***	3.087***
BB	1.584***	1.237***	1.051	1.102	1.154*	1.137*
	1.594***	1.244***	1.083	1.119	1.165**	1.137*
	1.602***	1.238***	1.046	1.096	1.148	1.141*
	3.016***	2.633***	2.647***	3.260***	3.587***	3.160***
B	1.794***	1.191***	0.983***	0.967***	1.064	1.170***
	1.816***	1.188***	0.991***	0.978***	0.985***	1.180***
	1.805***	1.197***	0.982***	0.955***	1.056	1.168***
	4.167***	2.991***	2.461***	3.115***	3.966***	3.654***
CCC or below	1.569***	1.327***	1.059	0.909***	0.783***	0.890***
	1.591***	1.331***	0.973***	0.901***	0.794***	0.879***
	1.568***	1.336***	1.041	0.917***	0.776***	0.861***
	3.939***	4.188***	4.223***	2.619***	2.781***	3.270***
Sample 2: 2009:1 - 2018:12						
AAA	1.128	1.097	0.940***	1.016	1.049	1.065
	1.132	1.090	0.942***	1.017	1.028	1.057
	1.125	1.100	0.945***	1.019	1.058	1.076
	2.264***	2.55***	2.235***	2.013***	2.027***	2.164***
AA	1.001***	0.954***	0.861***	0.864***	0.943***	0.941***
	1.000***	0.952***	0.864***	0.868***	0.945***	0.946***
	0.999***	0.954***	0.857***	0.858***	0.939***	0.938***
	1.967***	1.972***	2.006***	2.194***	1.738***	1.988***
A	1.056	0.937***	0.867***	0.875***	0.897***	0.930***
	1.053	0.933***	0.868***	0.875***	0.897***	0.929***
	1.057	0.939***	0.827***	0.874***	0.893***	0.948***
	2.325***	2.018***	2.226***	2.153***	2.232***	2.023***
BBB	1.154***	1.031	0.981***	0.831***	0.891***	0.957***
	1.156***	1.033	0.990***	0.838***	0.899***	0.967***
	1.149***	1.059	0.975***	0.821***	0.859***	0.942***
	5.195***	3.540***	2.623***	2.685***	3.748***	3.860***
BB	1.217***	1.005*	0.930***	0.940***	0.877***	0.811***
	1.215***	1.004*	0.971***	0.961***	0.888***	0.812***
	1.227***	1.006	0.923***	0.927***	0.870***	0.814***
	2.943***	2.236***	2.518***	3.642***	3.968***	2.558***
B	1.519***	0.989***	0.880***	0.765***	0.874***	0.870***
	1.530***	0.983***	0.882***	0.775***	0.747***	0.876***
	1.526***	0.997***	0.878***	0.751***	0.865***	0.872***
	4.508***	2.832***	2.324***	3.361***	4.406***	3.452***
CCC or below	1.278***	1.127***	0.997***	0.861***	0.649***	0.672***
	1.290***	1.120*	0.888***	0.848***	0.659***	0.654***
	1.294***	1.139***	0.977***	0.871***	0.644***	0.642***
	3.667***	3.555***	4.292***	2.561***	2.617***	2.916***

[1] This table presents results for a subset of models, including those that perform the best in terms of mean square forecast errors (MSFEs) relative to the AR benchmark, for the corporate bond yield target variables analyzed in our prediction experiments. Please refer to Section 3 for a detailed discussion of these variables and Section 4 for a summary of the empirical findings. The entries in the table are arranged in blocks of four rows for each variable, displaying the MSFEs for the following models in sequence: AR+MFV^{RV}, AR+MFV^{TRV}, AR+MFV^{BPV}, and AR+MVF^{JV}, as described in Table 2.

Table 5A: MSFE-Best Models (Sample 1: 2006:1 - 2018:12)[1]

Targets	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
	rolling window size = 36					
HS	RV	RV	CMTRV1	CMRV2	CMJV2	MMF
	0.930^{***}	0.897^{***}	0.998^{***}	0.982^{***}	0.962^{***}	0.938^{***}
IP	CMTRV1	MVJV	MVRV	MMF	CMJV2	TRV
	0.987^{***}	0.939^{***}	0.952^{***}	1.010	0.973^{***}	0.997^{***}
PAY	CMRV2	VBPV	MMF	CMTRV1	MMF	VJV
	0.912^{***}	1.024	0.978^{***}	0.926^{***}	0.956^{***}	1.010^{*}
CPI	VJV	VJV	CMJV2	CMJV2	CMRV2	VBPV
	1.025^{***}	0.998^{***}	0.995^{***}	0.972^{***}	0.904^{***}	1.024
PCE	CMRV2	CMBPV1	VJV	CMJV2	VJV	CMJV2
	1.008	1.000^{***}	1.020	0.986^{***}	0.943^{***}	0.973^{***}
SI	MVRV	CMJV2	VRV	MMF	CMTRV2	CMJV2
	1.006	0.998^{***}	1.028^{***}	1.022	1.012	0.965^{***}
AAA	VJV	CMBPV2	TRV	BPV	TRV	MMF
	1.007	0.987	0.873	0.895	0.993	0.976
AA	VJV	CMTRV2	BPV	VJV	VJV	CMTRV1
	0.907	0.996	0.916	0.896	0.974	1.003
A	VJV	CMTRV1	BPV	VJV	VJV	CMTRV1
	0.983	0.970	0.914	0.857	0.993	1.016
BBB	VJV	VJV	CMJV2	VJV	BPV	CMTRV1
	0.997	1.012	0.949	0.940	0.982	0.934
BB	CMTRV1	MMF	CMJV2	CMJV2	VJV	CMJV1
	1.050	1.029	0.945	0.988	0.947	0.885
B	VJV	VJV	CMJV1	VJV	CMJV2	CMJV2
	1.054	0.996	0.911	0.830	0.964	0.940
CCC	VBPV	CMJV2	CMJV1	TRV	BPV	CMJV2
	1.008	0.997	0.956	0.901	0.776	0.830
	rolling window size = 72					
HS	RV	BPV	BPV	TRV	TRV	TRV
	0.849^{***}	0.779^{***}	0.710^{***}	0.561^{***}	0.477^{***}	0.463^{***}
IP	MMF	CMRV2	MVBPV	MVRV	MVRV	MVRV
	0.954^{***}	0.985^{***}	0.890^{***}	0.924^{***}	0.95624^{***}	0.912^{***}
PAY	MMF	VTRV	VJV	VTRV	VBPV	RV
	0.832^{***}	0.931^{***}	0.837^{***}	0.795^{***}	0.788^{***}	0.678^{***}
CPI	CMRV1	CMTRV2	CMBPV2	CMRV1	CMRV2	CMJV2
	0.980^{***}	1.000^{***}	1.007	0.993^{***}	0.949^{***}	1.001^{**}
PCE	CMBPV2	CMTRV2	MVRV	MVTRV	MVBPV	MMF
	0.977^{***}	1.008	0.991^{***}	1.025	0.948^{***}	0.952^{***}
SI	CMBPV1	CMJV2	CMBPV2	CMJV1	CMTRV2	CMTRV1
	0.946^{***}	1.001	0.991^{***}	1.008	0.984^{***}	0.985^{***}
AAA	MRV	MMF	CMJV1	MBPV	MVBPV	MVBPV
	0.862	0.942	0.942	0.908	0.888	0.862
AA	MTRV	TRV	MVJV	VJV	MVRV	VTRV
	0.912	0.859	0.876	0.872	0.892	0.902
A	MTRV	TRV	MVJV	VJV	VTRV	VBPV
	0.972	0.912	0.926	0.931	0.923	0.909
BBB	MMF	MTRV	MVJV	VJV	CMJV2	CMTRV2
	1.011	0.920	0.891	0.881	0.969	0.905
BB	MVJV	MVJV	MVJV	VJV	CMJV1	CMJV1
	1.012	1.011	0.883	0.977	0.942	0.883
B	MMF	CMJV2	CMJV2	CMBPV2	CMJV1	CMJV2
	0.977	0.986	0.920	0.981	0.946	0.901
CCC	MMF	MVRV	TRV	CMTRV1	CMJV1	CMTRV1
	0.934	0.906	0.994	0.969	0.934	0.963

[1] This table presents the MSFE-best models for all 6 macroeconomic and 7 yield target variables that were examined in the forecasting experiments. The models were selected based on their superior performance relative to other candidate models. For each variable, the entry indicates the forecasting model that exhibited the lowest relative MSFE. Please refer to Section 4 for a summary of the empirical results.

Table 5B: MSFE-Best Models (Sample 2: 2009:1 - 2018:12)

Targets	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
			rolling window size = 36			
HS	TRV	VJV	CMTRV1	BPV	VJV	CMBPV1
	0.950***	0.818***	0.927***	0.910***	0.966***	0.920***
IP	CMTRV1	MVJV	MVRV	CMRV2	CMJV2	TRV
	0.931***	0.876***	0.956***	1.018	0.975***	0.975***
PAY	CMRV2	VJV	CMRV1	CMTRV1	MMF	CMJV2
	0.828***	1.040	0.996***	0.910***	0.993***	0.994***
CPI	CMJV2	VJV	CMTRV2	CMJV1	CMRV1	VBPV
	1.015***	0.988***	0.986***	0.946***	0.954***	1.043
PCE	CMRV2	CMRV2	CMJV2	CMJV2	VRV	CMTRV1
	0.992***	0.977***	1.011	1.016	0.992***	0.962***
SI	CMJV2	CMJV2	VJV	CMJV2	CMTRV2	CMJV2
	1.036	1.013***	1.012	1.028***	0.993***	0.976***
AAA	CMJV2	CMTRV1	VJV	VJV	CMJV2	CMJV2
	0.995	0.936	0.910	0.957	0.960	0.987
AA	BPV	TRV	BPV	VJV	VJV	CMTRV2
	0.999	0.952	0.857	0.815	0.841	0.931
A	CMJV2	TRV	BPV	VJV	VJV	TRV
	1.009	0.933	0.827	0.708	0.824	0.929
BBB	CMTRV1	VJV	BPV	BPV	BPV	CMTRV1
	1.023	1.007	0.975	0.821	0.859	0.847
BB	VBPV	TRV	CMJV2	MTRV	BPV	CMJV2
	0.860	1.004	0.921	0.880	0.870	0.771
B	VRV	CMJV2	MTRV	MTRV	MTRV	CMJV2
	0.945	0.935	0.851	0.727	0.669	0.824
CCC	VBPV	CMJV2	TRV	TRV	BPV	BPV
	0.993	0.964	0.888	0.848	0.644	0.642
			rolling window size = 72			
HS	CMJV1	CMTRV1	VBPV	CMTRV1	VBPV	MMF
	0.938***	0.927***	0.955***	0.951***	0.947***	0.949***
IP	MMF	TRV	MVBPV	VRV	CMJV1	MVJV
	0.950***	0.962***	0.972***	0.963***	0.963***	0.938***
PAY	CMTRV2	CMJV2	CMRV2	CMJV2	MRV	VTRV
	0.864***	1.048*	0.975***	0.957***	0.902***	0.953***
CPI	CMRV1	CMTRV2	CMBPV2	CMRV2	CMRV2	TRV
	0.970***	0.993***	0.985***	0.983***	0.946***	0.995***
PCE	MVJV	CMTRV2	MVRV	MVBPV	MVBPV	MMF
	0.979***	0.994***	0.934***	0.962***	0.894***	0.858***
SI	CMBPV1	CMJV2	CMBPV2	MMF	CMRV2	CMTRV1
	0.853***	1.000***	0.982***	0.948***	0.987***	0.943***
AAA	MTRV	MTRV	CMBPV2	MBPV	MVBPV	MJV
	0.880	0.941	0.945	0.890	0.862	0.898
AA	VJV	MTRV	MTRV	BPV	BPV	VRV
	0.870	0.848	0.885	0.789	0.850	0.920
A	VJV	MJV	TRV	TRV	BPV	MTRV
	0.894	0.851	0.806	0.727	0.814	0.883
BBB	VJV	MJV	TRV	TRV	TRV	BPV
	0.961	0.826	0.843	0.780	0.816	0.865
BB	VBPV	CMJV2	VTRV	VJV	CMJV1	CMJV2
	0.842	0.913	0.895	0.876	0.919	0.833
B	VBPV	CMJV2	MVJV	VJV	CMJV2	CMJV2
	0.835	0.937	0.920	0.869	0.890	0.841
CCC	MMF	MVTRV	MTRV	VTRV	BPV	CMJV2
	0.947	0.892	0.977	0.929	0.924	0.968

[1] This table presents results similar to those in Table 5A, but with the difference that Sample 2 is used instead of Sample 1 in forecasting experiments.

Table 5C: Ex-ante Directional Accuracy Rate Best Models (Sample 1: 2006:1 - 2018:12)[1]

Targets	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
			rolling window size = 36			
HS	CMTRV2	RV	CMRV2	CMTRV1	TRV	MJV
	70.9%***	73.4%***	68.4%***	75.9%***	70.9%***	68.4%***
IP	CMJV1	CMTRV1	CMBPV2	MJV	MVRV	MRV
	72.2% ***	73.4% ***	65.8% ***	72.2% ***	74.7% ***	64.6% **
PAY	MVJV	CMTRV2	MVJV	MMF	MMF	CMRV2
	78.5%***	81.0%***	72.2%***	72.2%***	75.9%***	75.9%***
CPI	BPV	MMF	BPV	CMBPV1	RV	BPV
	72.2%***	81.0%***	81.0%***	83.5%***	82.3%***	77.2%***
PCE	JV	MTRV	CMTRV1	JV	BPV	MMF
	75.9%***	74.7%***	74.7%***	81.0%***	77.2%***	72.2%***
SI	CMJV2	MMF	CMBPV1	MMF	CMRV2	CMRV2
	75.9%***	72.2%***	75.9%***	75.9%***	73.4%***	74.7%***
			rolling window size = 72			
HS	CMTRV1	VRV	CMRV1	RV	RV	TRV
	74.7%***	78.5%***	75.9%***	79.7%***	73.4%***	70.9%***
IP	CMTRV2	BPV	CMBPV2	MVRV	MJV	JV
	75.9%	70.9% ***	69.6% ***	78.5% ***	75.9% ***	65.8% *
PAY	VTRV	VRV	MJV	MVTRV	RV	MMF
	79.7%***	77.2%***	77.2%***	77.2%***	73.4%***	79.7%***
CPI	CMJV1	VRV	CMJV2	CMRV2	MRV	MJV
	69.6%***	81.0%***	79.7%***	78.5%***	81.0%***	78.5%***
PCE	CMTRV1	CMBPV1	VRV	VRV	MTRV	VRV
	77.2%***	70.9%***	74.7%***	78.5%***	75.9%***	72.2%***
SI	CMJV2	MRV	VBPV	MMF	CMRV2	CMJV1
	75.9%***	68.4%***	79.7%***	73.4%***	73.4%***	70.9%***

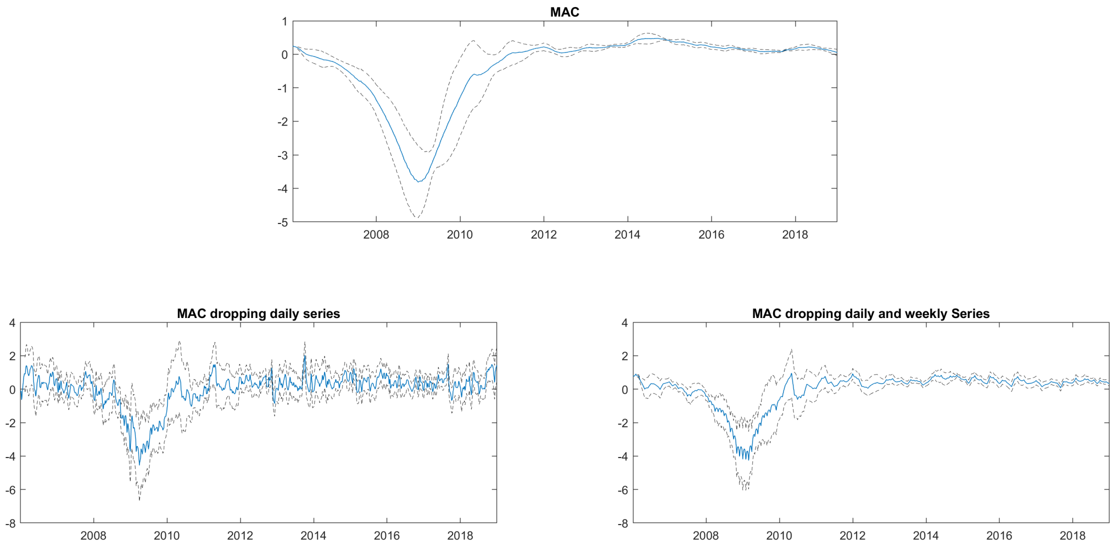
[1] This table is similar to Table 5A, but instead of mean square forecast errors, it reports directional accuracy rates for the best models. The table lists the models that have the highest directional accuracy rates for the 6 macroeconomic and 7 yield target variables examined in the forecasting experiments discussed in Section 3.

Table 5D: Ex-ante Directional Accuracy Rate Best Models (Sample 2: 2009:1 - 2018:12)[1]

Targets	Forecast horizon 1-month	2-month	3-month	4-month	5-month	6-month
	rolling window size = 36					
HS	CMTRV2	RV	CMBPV2	MVJV	TRV	MVRV
	72.1%***	79.1%***	81.4%***	79.1%***	76.7%***	76.7%***
IP	CMTRV1	CMJV2	CMRV1	MVRV	MVRV	MRV
	72.1%***	76.7%***	67.4%*	76.7%***	86.0%***	65.1%*
PAY	TRV	TRV	TRV	MMF	MVJV	CMJV2
	83.7%***	86.0%***	72.1%***	72.1%***	81.4%***	81.4%***
CPI	BPV	MMF	BPV	BPV	RV	MMF
	72.1%***	81.4%***	81.4%***	79.1%***	81.4%***	79.1%***
PCE	JV	MTRV	CMTRV1	JV	CMRV2	MMF
	81.4%***	72.1%***	72.1%***	79.1%***	79.1%***	72.1%***
SI	VJV	JV	CMRV2	MMF	MVRV	CMRV1
	79.1%***	74.4%***	74.4%***	79.1%***	76.7%***	72.1%***
	rolling window size = 72					
HS	CMRV1	VRV	CMRV1	JV	RV	MMF
	76.7%***	76.7%***	81.4%***	79.1%***	67.4%***	69.8%***
IP	CMTRV2	MMF	RV	MRV	JV	JV
	74.4%	74.4%***	67.4%***	76.7%***	81.4%***	67.4%
PAY	CMBPV2	CMJV1	CMRV1	CMJV2	MMF	RV
	86.0%***	86.0%***	81.4%***	74.4%***	79.1%***	83.7%***
CPI	VRV	CMJV2	CMJV2	MMF	JV	MJV
	74.4%***	81.4%***	83.7%***	74.4%***	81.4%***	81.4%***
PCE	CMTRV1	MMF	VRV	MMF	MRV	MMF
	81.4%***	69.8%***	69.8%***	79.1%***	76.7%***	72.1%***
SI	CMTRV1	MRV	VBPV	MMF	CMRV2	VRV
	76.7%***	74.4%***	79.1%***	79.1%***	74.4%***	72.1%***

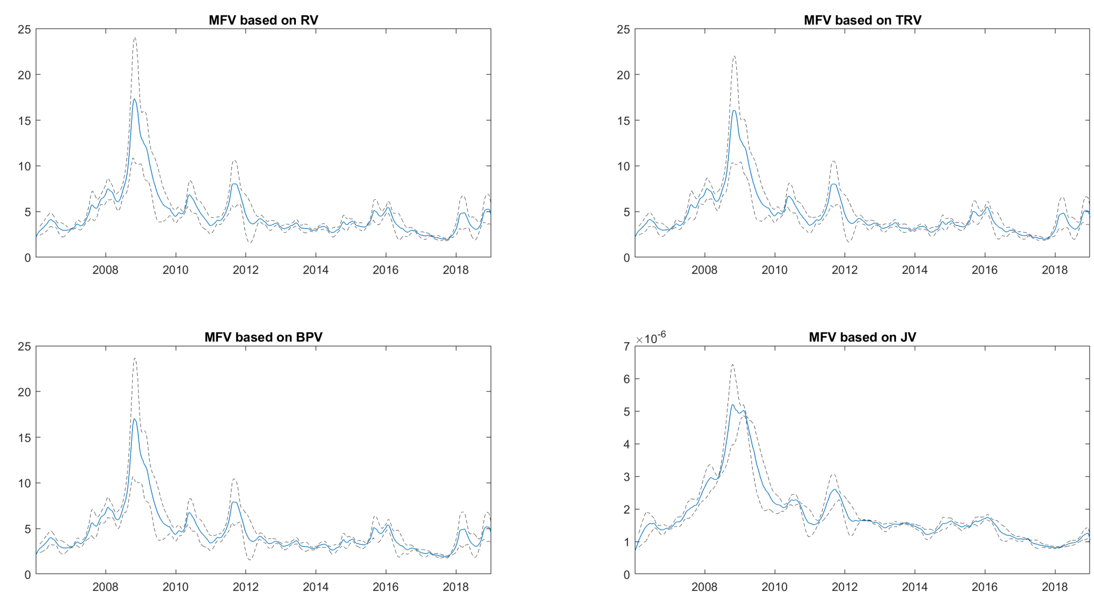
[1] This table presents results similar to those in Table 5C, but with the difference that Sample 2 is used instead of Sample 1 in forecasting experiments.

Figure 1: Mixed frequency macroeconomic (MAC) factor



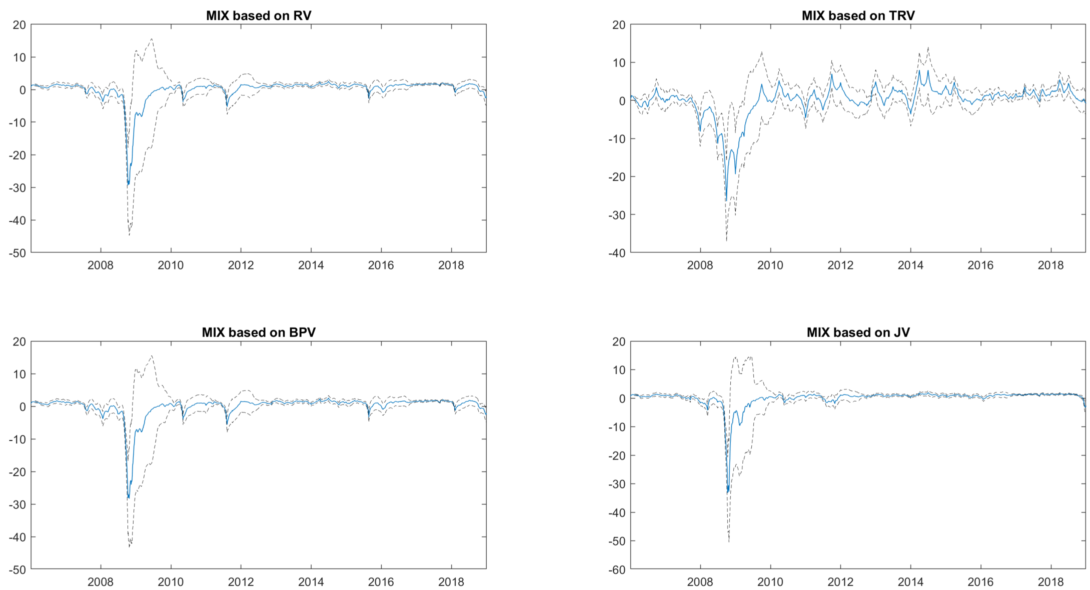
Notes: This figure contains plots of our mixed frequency macroeconomic (MAC) factor (solid line), along with a 95% confidence intervals constructed using recursively estimated variances (dashed line). The upper plot shows the main MAC factor estimated using all four component variables, while the two lower plots show two variants of the MAC factor obtained by dropping daily as well as daily and weekly variables from the model, respectively. Section 2 provided details of the model setup and methodology, while details on the macroeconomic variables used to construct the MAC factor are contained in Section 4.

Figure 2: Mixed frequency financial volatility (MFV) factors



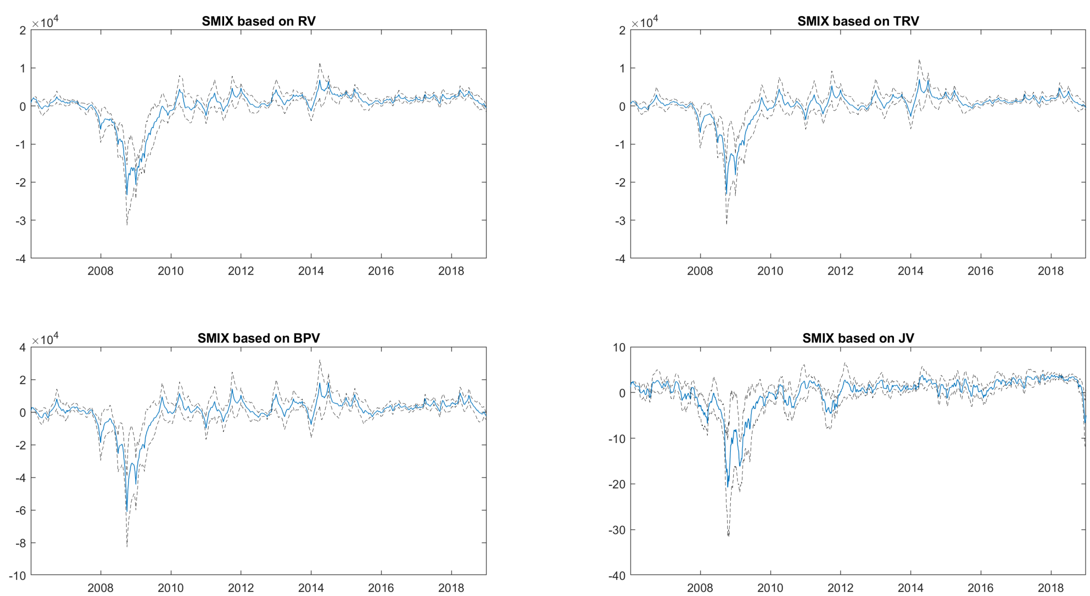
Notes: See notes to Figure 1. This figure contains plots of our four mixed frequency financial volatility (MFV) factors based on different volatility measures, including RV, TRV, BPV, and JV. For more information on the specific volatility measures plotted here, refer to Section 2.

Figure 3: Mixed frequency financial volatility and macroeconomic (MIX) factors



Notes: See notes to Figure 2. This figure contains plots of our four mixed frequency financial volatility and macroeconomic (MIX) factors. All factors use our four macroeconomic component variables and one of each of our volatility measures, including RV, TRV, BPV, and JV.

Figure 4: Square root mixed frequency financial volatility and macroeconomic (SMIX) factors



Notes: See noytes to Figure 3. This figure contains plots of our four square root mixed frequency financial volatility and macroeconomic (SMIX) factors, which are constructed using a setup similar to that used for our MIX factors, except that the volatility measures are replaced with the square root of RV, TRV, BPV, and JV.