Investor Confidence and Forecastability of US Stock Market Realized Volatility : Evidence from Machine Learning

Rangan Gupta[a], Jacobus Nel[b] and Christian Pierdzioch[c]

February 2021

Abstract

Using a machine-learning technique known as random forests, we analyze the role of investor confidence in forecasting monthly aggregate realized stock-market volatility of the United States (US), over and above a wide-array of macroeconomic and financial variables. We estimate random forests on data for a period from 2001 to 2020, and study horizons up to one year by computing forecasts for recursive and a rolling estimation window. We find that investor confidence, and especially investor confidence uncertainty has out-of-sample predictive value for overall realized volatility, as well as its "good" and "bad" variants. Our results have important implications for investors and policymakers.

JEL classification: C22; C53; G10; G17

Keywords: Investor Confidence; Realized Volatility; Macroeconomic and Financial Predictors; Forecasting; Machine Learning

[a] Department of Economics, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa; Email address: rangan.gupta@up.ac.za.
[b] **Corresponding author**. Department of Economics, University of Pretoria, Private Bag X20, Hatfield, 0028, South Africa; Email address: neljaco380@gmail.com.
[c] Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany; Email address: macroeconomics@hh.de.

1 Introduction

As market agents tend to make overly optimistic or pessimistic judgments and choices, following the seminal contributions of Baker and Wurgler (2006, 2007), many other studies (see for example, Bathia and Bredin (2013), Gebka (2014), Da et al., (2015), Huang et al. (2015), Jiang et al. (2019), Chen et al. (forthcoming)) have highlighted the role of investor sentiment in predicting (in- and out-of-sample) aggregate and firm-level (excess) stock returns of the United States (US). Building on this area of behavioral finance, recent works (see for example, Gupta and Kyei (2016), Balcilar et al. (2018), Gupta (2019), Olson and Nowak (2019)) have also highlighted the role of investor sentiment in driving US stock market volatility. This second-moment effect is not surprising due to the existence of so-called “noise traders" in the market, who in turn, are investors whose trading decisions are based on what they perceive to be an informative signal but which, to a rational agent, does not convey any information (Black (1986)). In the presence of such noise trading, equity prices tend to drift further from the fundamentals (Zhang (2006)), which then results in higher liquidity in terms of trading volume (Greene and Smart (1999)) and consequently higher risk, i.e., volatility, via the Mixture of Distribution Hypothesis (MDH) introduced by Clark (1973), or Sequential Information Arrival Hypothesis (SIAH) developed by Copeland (1976). Recall that, the MDH postulates that the innovation on returns is a linear combination of the intraday return movements. The intraday return increment incorporates the number of information flows arrived into the market in a given day. Since the intraday price movement is random, daily returns follow a mixture of normally distributed random variables with the information flow into the market as a mixing variable. To sum up, daily price changes are driven by a set of information flow and the arrival of unexpected news is accompanied by the above average trading activity. On the other hand, the SIAH questions the instantaneous relationship as predicted by MDH and provides a different explanation. It argues that each trader observes the information signal differently at times and may not receive the information simultaneously, thereby generating a series of incomplete equilibria. Market equilibrium can be established provided that all traders receive the same set of information simultaneously. Thus, the shift of new information is not immediate as considered in the MDH. Nevertheless, both MDH and SIAH believe that the price volatility of the market can be potentially predictable through the knowledge of trading volume, and that the relationship of volume and volatility is positive.

While the studies showing a causal impact of investor sentiment on stock market volatility are indeed insightful, they are based on models of in-sample predictability, which provides no guarantee of out-of-sample forecasting gains, and the ultimate test of any predictive model (in terms of the econometric methodologies and the predictors used) is in its forecasting performance (Campbell (2008)). This is precisely what we aim to do in this paper by forecasting stock market volatility of the US, using five (survey-based) measures of (individual and institutional) investor sentiment and confidence, over and above a wide-array of macroeconomic and financial variables (the importance of which for volatility has been highlighted by Engle and Rangel (2008), Rangel and Engle (2011), Asgharian et al. (2013), Engle et al. (2013), Conrad and Loch (2015) among others) summarized by (eight) latent factors (derived from a large data set), (six) metrics of associated uncertainties of these variables (as suggested to have predictive content for volatility, for example, by Bekaert and Hoerova (2014), Liu and Zhang (2015), Kaminska and Roberts-Sklar (2018), Baker et al. (2020)). Notably, we not only look at the mean values of the investor sentiment and confidence indexes, but also their corresponding time-varying volatility capturing investors' uncertainty, which in turn, can also lead to noise and stock market risk by stimulating investors' psychological biases (Escobari and Jafarinejad (2019), Liu and Gupta (2021)). One must realize that forecasting of volatility is of importance due to several reasons (as outlined in Poon and Granger (2003), and Rapach et al. (2008)): Firstly, when volatility is interpreted as uncertainty, it becomes a key input to investment decisions and portfolio choices. Secondly, volatility is the most important variable in the pricing of derivative securities. To price an option, one needs reliable estimates of the volatility of the underlying assets. Thirdly, financial risk management according to the Basle Accord as established in 1996 also requires modeling and forecasting of volatility as a compulsory input to risk-management for financial institutions around the world. Finally, financial market volatility, as witnessed during the Global Financial Crisis of 2007-2008 and more recently during the outbreak of the COVID-19 pandemic, can have wide repercussions on the economy as a whole. Hence, forecasts of market volatility can serve as a measure for the vulnerability of financial markets and the economy, and can help policymakers design appropriate policies. Evidently, accurate forecasting of the process of volatility has ample implications for portfolio selection, the pricing of derivative securities, risk management, and policy decisions.

Specifically speaking, we forecast monthly realized volatility (RV) using an extended version of the heterogeneous autoregressive (HAR)-RV model of Corsi (2009), which incorporates the role of the various monthly predictors over the period of 2001:06 to 2020:06. Note that measuring volatility using RV, which in our case is captured by the sum of squared returns of the Dow Jones Industrial Average (DJIA) over a month (following Andersen and Bollerslev (1998)), provides us an observable and unconditional metric of volatility (unlike in the case of generalized autoregressive conditional heteroscedastic (GARCH) and stochastic volatility (SV) models), which is otherwise a latent process. At the same time, the benchmark HAR-RV model, can capture long-memory and multi-scaling properties observed for stock market volatility (e.g., Mei et al. (2015)), despite having a simplistic structure. In this regard, the key feature of the HAR-RV model is that it uses volatilities from different time resolutions to forecast the realized stock-price volatility. The model, thereby, captures in a parsimonious way the key elements of the heterogeneous market hypothesis (Müller et al. (1997)), which states that different classes of market participants populate the stock market and differ in their sensitivity to information flows at different time horizons.

Econometrically, for our forecasting application, we use a machine-learning technique known as random forests (Breiman (2001)) to compute forecasts of overall RV, as well as good (sum of squared positive returns) and bad (sum of squared negative returns) RVs, given the observation by Giot et al. (2010) that financial market participants care not only about the level of volatility, but also of its nature, with all traders making the distinction between upside and downside volatilities. The use of random forests in our forecasting application has several advantages. First, random forests render it possible to analyze the links between realized volatility and a large number of predictors (as is the situation in our case, which involves as many as 34 predictors) in a fully data-driven way. Second, random forests automatically capture potential nonlinear links between realized volatility and its predictors (as depicted by the literature on investor sentiment and stock market volatility discussed above), as well as any interaction effects among the predictors. Finally, unlike the ordinary-least-squares technique commonly used to estimate HAR-RV models, random forests always yield forecasts of realized volatility that are non-negative.

To the best of our knowledge, this is the first attempt to analyze the role of investor sentiment and confidence indexes, over and above a large number of macroeconomic and financial control variables, in forecasting the RV of the DJIA using machine-learning. In the process, we add to earlier studies using machine-learning techniques to forecast stock returns (e.g., Gu et al. (2020)) and financial-market volatility (e.g., Mittnik et al. (2015), Bouri et al. (2020)) as well as to the already existing large literature on the forecastability of US stock market volatility (see for example, Ben Nasr et al. (2014, 2016), Salisu et al. (2020) for detailed reviews) by considering the role of behavioral variables. We structure the remainder of this research as follows. In Section 2, we briefly describe how a random forest is grown. In Section 3, we describe our data and report our empirical results. Finally, in Section 4, we conclude with final remarks.

2 Random forests

For forecasting realized stock-market volatility, we use a popular machine-learning technique known as random forests. Intuitively, a random forest is an ensemble of a large number of individual regression trees (see Hastie et al., (2009) for a for a textbook exposition; see also the exposition in Bouri et al. (2020); our exposition closely follows theirs). A regression tree, T, is made up of branches that subdivide the space of predictors, $\mathbf{x} = (x_1, x_2, ...)$, of realized stock-market volatility, RV, into l non-overlapping regions, R_l. In order to compute these regions, a recursive search-and-split algorithm is applied in a top-down way.

In order to illustrate how this search-and-split algorithm works, we start at the top level of a tree, and then iterate over the various predictors, s, and all potential splitting points, p, that can be formed using the data on a predictor. For every combination of a predictor and a splitting point, the search-and-split algorithm identifies two half-planes, $R_1(s,p) = \{x_s | x_s \leq p\}$ and $R_2(s,p) = \{x_s | x_s > p\}$, in such a way so as to minimize the following squared-error loss criterion:

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

where i denotes observations on RV, that belong to a half-plane, and $\bar{RV}_k = \text{mean}\{RV_i \mid x_s \in R_k(s,p)\}, k = 1,2$ is the half-plane-specific mean of RV. This minimization problem requires (i) searching over all combinations of s and p, and, (ii) given a combination of s and p, minimizing the half-plane-specific loss by an optimal choice of the half-plane-specific means of RV. Solving the minimization problem gives the top-level (because we started at the top level of a tree) optimal splitting predictor and optimal splitting point, and the two region-specific means, $\bar{RV}_k$ (because our regression tree has two terminal nodes).

The region-specific means could already be used for forecasting RV, but it is clear that growing a somewhat larger tree should give a more nuanced description of the link between realized volatility and its predictors. In order to grow a larger tree, the natural next step is to apply the search-and-split algorithm (and, thus, the minimization problem given in Equation (1)) to the two top-level half-planes, $R_1(s,p)$ and $R_2(s,p)$. The result of such an application of the search-and-split algorithm are two second-level optimal splitting predictors and optimal splitting points, and four second-level region-specific means of RV.

Naturally, solving the minimization problem over and over again results in an increasingly complex regression tree. Tree building stops when a regression tree has a preset maximum number of terminal nodes or every terminal node has a minimum number of observations. In our empirical application, we use a cross-validation technique to determine the optimal minimum number of observations per terminal node (see Section 3.2 for details).

Equipped with a regression tree, we send the data on the predictors from the top level of the tree down to its leaves along the nodes and branches of the tree, and compute the forecast of RV by its region-specific mean. For a regression tree made up of L regions, this forecast is formed as follows ($\mathbf{1}$ denotes the indicator function):

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

A problem when using a large regression tree for forecasting RV is that its complex hierarchical structure gives rise to an overfitting and data-sensitivity problem. A random forest stabilize forecasting performance by solving the overfitting problem in three steps. (1) A large number of bootstrap samples (sampling with replacement) is drawn from the data. (2) A random regression tree is fitted to every bootstrap sample. A random regression tree uses for every splitting step only a random subset of the predictors and, thereby, reduces the effect of influential predictors on tree building. In addition, growing a large number of such random trees lowers the correlation of forecasts from the individual trees. (3) A random forecast averages the forecasts of RV obtained from the individual random regression trees stabilizes the forecasts of RV.

3 Empirical analysis

3.1 Data and empirical models

Since the construction of the confidence indexes, details of which we will discuss below, is based on the survey conducted on investors in the Dow Jones, we consider the daily market returns of the DJIA to compute the realized market volatility estimates (RV) for each month from log daily returns (r_t) as follows:

$$RV_t = \sum_{i=1}^{N} r_i^2, \tag{3}$$

where N denotes the number of data available for the month. In addition to total aggregate RV, we examine "good" and "bad" realized volatility. The "good" and "bad" components of realized volatility are defined as the upside and downside realized semi-variances (RV^B and RV^G), respectively, computed from positive and negative returns (see Barndorff-Nielsen et al. (2010)) as follows (I denotes the indicator function):

$$RV_t^B = \sum_{i=1}^{T} r_i^2 \, I_{[(r_i)<0]}, \tag{4}$$

$$RV_t^G = \sum_{i=1}^{T} r_i^2 \, I_{[(r_i)>0]}. \tag{5}$$

When studying daily realized volatility, it is common practice among researchers to consider daily, weekly, and monthly realized volatilities as predictors of subsequent realized volatility. In our case, because we study monthly data, in the benchmark HAR-RV framework, we forecast the average realized volatility, RV_{t+h} from month $t+1$ to month $t+h$ using the current realized volatility, RV_t, the quarterly realized volatility, $RV_{t,q}$, computed as the average realized volatility from month $t-3$ to month $t-1$, and the yearly realized volatility, $RV_{t,y}$, computed as the average realized volatility from month $t-12$ to month $t-1$. We compute these quarterly and yearly average realized volatilities for the standard measure of realized volatility and for good and bad realized volatilities. The baseline HAR-RV model defined in this way forms our Model 1.

Besides the current, quarterly and yearly realized volatilities, we now turn our attention to the detailed discussion of our other predictors. In this regard, we use 8 factors derived from the 134 macroeconomic variables of Ludvigson and Ng (2009, 2011).[1] Including these series gives us the advantage of capturing broad categories of overall and regional macroeconomic time series namely, real output and income, employment and hours, real retail, manufacturing and sales data, international trade, consumer spending, housing starts, housing building permits, inventories and inventory sales ratios, orders and unfilled orders, compensation and labor costs, capacity utilization measures, price indexes, interest rates and interest rate spreads, stock market indicators, and foreign exchange measures. In addition, we use the macroeconomic uncertainty (MU) and financial uncertainty (FU) measures of Jurado et al. (2015) and Ludvigson et al. (forthcoming), which, in turn, is the average time-varying variance in the unpredictable component of 134 macroeconomic and 148 financial time-series respectively, i.e., it attempts to capture the average volatility in the shocks to the factors that summarize real and financial conditions.[2] In other words, it is derived based on the second approach outlined above. Note that the same 134 variables are used in computing the factors used as predictors and the metric of macroeconomic uncertainty. The metrics that we use are the broadest measures of macroeconomic and financial uncertainties available for the US. The uncertainty indexes are available for three forecasting horizons of 1-, 3-,and 12-month-ahead. So the benchmark HAR-RV model is extended with these 14 (8 factors and 3 MUs and FUs each) additional predictors to constitute Model 2.

[1]The factors are available for download from: `https://www.sydneyludvigson.com/data-and-appendixes`.

[2]The MU and FU indexes are available for download from: `https://www.sydneyludvigson.com/data-and-appendixes`.

As far as the investor sentiment index is concerned, we rely on the American Association of Individual Investors (AAII) Sentiment Survey, whereby the sentiment survey measures the percentage of individual investors who are bullish, bearish, and neutral on the stock market in the short-term, with individuals polled from the AAII website on a weekly basis, and only one vote per member is accepted in each weekly voting period.[3] Since 1987, AAII members have been answering the same simple question each week, i.e., What direction do AAII members feel the stock market will be in the next 6 months? The results are compiled into the AAII investor sentiment index, measured as the difference between the bullish and bearish percentages,[4] which offers insight into the mood of individual investors. The AAII Investor Sentiment Survey has become a widely followed measure of the mood of individual investors. The weekly survey results are published in financial publications including Barron's and Bloomberg and are widely followed by market strategists, investment newsletter writers and other financial professionals. Since our analysis is at the monthly-level, we take the weekly averages over a month, as well as the standard deviation of the bull minus bear opinions to capture monthly investor sentiment and uncertainty around the same. The investor sentiment index is added to the predictors in Model 2, to give us Model 3, while the standard deviation of the index is included in Model 3 to yield Model 4. At this stage, it must be realized that there are multiple ways, such as combining information on market-based proxies and internet searches besides surveys (see Zhou (2018) for a detailed review in this regard), in measuring investor sentiment, which is a latent variable. We chose to use the AAII index as compared to the other publicly available indexes like that of Baker and Wurgler (2006, 2007) and Da et al. (2015), since data is available until recent months (unlike December, 2018 and December, 2011 respectively), and also with the unobservable investor confidence indexes, which we discuss next, being survey-based, makes the comparison in terms of their creation to some degree similar.

Based on surveys conducted on both large individual and institutional investors by the International Center for Finance at the Yale School of Management, we consider four investor confi-

[3]Note that, the University of Michigan survey-based consumer confidence index is included in the 8 macro and financial factors included in Model 2.

[4]The data is publicly available for download from: `https://www.aaii.com/sentimentsurvey/sent_results`, and is published every Wednesday, based on the poll conducted the previous Friday.

dence indexes, which are the only known measures for the US that are available publicly.[5] The first index is called the US one-year confidence index, which captures the percentage of the population expecting an increase (a number strictly greater than zero) in the DJIA in the coming year; the second one is the US buy-on-dips confidence index, and measures the percentage of the population expecting a rebound (increase) the next day should the market ever drop 3% in one day; the third index is called the US crash confidence index and is associated with the percentage of the population who attach little probability (strictly less than 10%) to a stock market crash in the next six months, and finally, the US the valuation confidence index which gives the percentage of the population who think that the market is not too high, i.e., the number of respondents who choose "too low" or "about right" as a percentage of those who choose "too low", "too high", and "about right". As indicated earlier, to capture investor confidence we use the index values, and the standard errors associated with these indexes to provide a proxy for investor confidence uncertainty, giving us a total of 16 additional predictors. Given this, Model 5 involves the predictors in Model 4 plus the 4 each investor confidence indexes for individual and institutional investors, and Model 6 include the predictors of Model 5 and the 8 investor confidence uncertainty measures.

Though the surveys for the calculation of the confidence indexes are conducted since September, 1989, continuous monthly data begins only from June, 2001, which forms our start date, and with the macroeconomic and financial factors and associated uncertainty measures available until June, 2020, our analysis ends at that point.

3.2 Forecasting results

We use a recursively expanding estimation window to estimate our Models 1 to 6 by random forests. To this end, we start with a training period of 100 observations, compute out-of-sample forecasts for $h = 1, 3, 6, 12$, add one month's data to reestimate the models and to compute another set of out-of-sample forecasts, and continue in this way until we reach the end of our

[5] The indexes are downloadable from: https://som.yale.edu/faculty-research-centers/centers-initiatives/international-center-for-finance/data/stock-market-confidence-indices/united-states-stock-market-confidence-indices.

sample period. We also consider somewhat longer training periods of 120 and 140 observations (and correspondingly shorter out-of-sample periods). For estimation of random forests and evaluation of the out-of-sample forecasts, we use the statistical computing program R (R Core Team 2019) and the add-on package "grf" (Tibshirani et al. 2020). While recursively expanding the estimation window, we use cross-validation for every model to optimize the number of predictors randomly selected for splitting, the minimum node size of a tree, and the parameter that governs the maximum imbalance of a node. We use 2,000 random regression trees to grow a random forest.[6]

Table 1 summarizes ratios of root-mean-squared-forecasting errors (RMSFE). We compute these ratios for various model configurations. For example, the column entitled "Model 1 / Model 2" summarizes the ratios of the RMSFE computed for Model 1 divided by the RMSFE that we obtain for Model 2. A RMSFE ratio that exceedss unity, thereby, indicates that the second model outperforms the first one. We compute the RMSFE ratios for both RV and $\sqrt{RV}$, that is, for the realized variance and the realized standard deviation (the latter is often being used in the empirical literature as a metric of realized volatility).

– Please include Table 1 about here. –

As one would have expected, the results that we summarize in Table 1 clearly show that Model 2 outperforms Model 1 in terms of the RMSFE for all trainng periods and forecast horizons, and for both RV (Panel A) and $\sqrt{RV}$ (Panel B). The differences between Models 2, 3, and 4 in terms of the RMSFE criterion, in turn, are less clear-cut. While the RMSFE ratio exceeds unity for some configurations of training periods and forecast horizon, it falls short of unity for others, where Model 3 (Model 4) tends to perform somewhat better than Model 2 (Model 3) when we study $\sqrt{RV}$. Furthermore, Model 5 tends to perform better than Model 4 in the majority of cases, and Model 6 tends to show a better performance in terms of the RMSFE criterion than Model 5, notably for the intermediate training period of (120 observations). In other words, a main

[6]It should be noted that the "grf" package offers the option to use different subsamples for constructing a tree and for making predictions. We deactivate this option, as in a classic random forest.

result is that adding the investor confidence indexes for individual and institutional investors, and especially the 8 investor confidence uncertainty measures improves forecast performance. A comparison of Model 6 and Model 2 (that is, the model that features the 8 factors and 3 MUs and FUs each).[7]

– Please include Table 2 about here. –

Table 2 summarizes the results (p-values) of the test suggested by Clark and West (2007) of equal mean-squared prediction errors. The test results corroborate the results of the RMSFE analysis. While, unsurprisingly, Model 2 outperforms Model 1, the test results are largely insignificant for the comparisons of Models 2 and 3 as well as Models 3 and 4. In sharp contrast, the test results become increasingly significant when we compare Models 4 and 5 (especially so when we study $\sqrt{RV}$), and they are significant in the overwhelming majority of configurations of training periods and forecast horizons when we test Model 5 against Model 6, and now for both metrics of realized volatility, RV (Panel A) and $\sqrt{RV}$ (Panel B). Similarly, Model 6 tends to outperform Model 2, especially for the intermediate and longer training period. Again, we conclude that investor confidence indexes for individual and institutional investors, and especially the investor confidence uncertainty measures, have predictive value for realized stock-market volatility.[8]

– Please include Table 3 about here. –

The results for good and bad realized volatility that we summarize in Table 3 lend further support to our main conclusion, with one qualification. Model 5 outperforms Model 4 to a lesser extent than in the case of standard aggregate realized stock-market volatility.The results for Model 6, in turn, clearly show that the investor confidence uncertainty measures can help investors to forecast out-of-sample realized good and bad stock-market volatility.

[7]Table A1 given at the end of the paper (Appendix) shows that using the mean-absolute-forecast error (MAFE) rather than the RMSFE gives qualitatively similar results.

[8]As a robustness check, we consider alternative rolling-estimation windows. Results are summarized at the end of the paper (Table 2) and confirm the results for the recursive-estimation windows.

4 Concluding Remarks

In several earlier empirical studies, machine-learning techniques have been used to study the determinants of the returns and the volatility of important financial time series. Our empirical study contributes to this mushrooming literature in that we use random forests to demonstrate that investor confidence in general, and investor confidence uncertainty in particular have predictive value for aggregate monthly realized US stock-market volatility (and its good and bad variants). Importantly, we observe this predictive value in out-of-sample forecasting experiments based on various recursive and rolling estimation windows, and for different forecast horizons of up to one year. Moreover, we observe the predictive value of investor confidence and especially investor confidence uncertainty in models that feature, in addition, a wide-array of macroeconomic and financial variables that have been extensively studied in earlier empirical research and which have been shown to contain important value for modeling volatility. Our empirical results, thereby, demonstrate that not only conventional macroeconomic and financial fundamentals, but also behavioral factors, and investor uncertainty specifically, matter for forecasting realized stock-market volatility.

Given that volatility forecasts are used as inputs for optimal asset-allocation decisions, our findings suggest that incorporating the role of investor confidence, over and above other macroeconomic and financial predictors, in forecasting models of market agents can help to improve the design of portfolios across various investment horizons. Given that stock-market volatility has historically impacted the real economy of the US (Pierdzioch and Gupta (2020)), policymakers would need to monitor investor confidence closely, and design policies accordingly to ensure that investor confidence uncertainty does not lead to volatile stock markets.

As part of future research, contingent on availability of data on investor confidence, it would be interesting to extend our analysis to other developed and emerging markets, and also to other asset markets, given evidence of volatility-connectedness (see, for example, Tiwari et al. (2018)).

References

Andersen, T.G., and Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. International Economic Review, 39 (4), 885−905.

Asgharian, H., Hou , A.J., and Javed , F. (2013). The importance of the macroeconomic variables in forecasting stock return variance: A GARCH MIDAS approach. Journal of Forecasting, 32: 600−612.

Barndorff-Nielsen, O.E., Kinnebrouk, S., and Shephard, N. (2010). Measuring downside risk: realised semivariance. In T. Bollerslev, J. Russell and M. Watson (eds.), Volatility and time series econometrics: Essays in honor of Robert F. Engle, 117−136. Oxford University Press.

Baker, S., Bloom, N., Davis, S. J., Kost, K., Sammon, M., and Viratyosin, T. (2020). The unprecedented stock market reaction to COVID-19. Review of Asset Pricing Studies, 10(4): 742−758.

Baker, M., and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. Journal of Finance, 61: 1645−1680.

Baker, M., and Wurgler, J. (2007). Investor sentiment in the stock market. Journal of Economic Perspectives, 21: 129−152.

Balcilar, M., Gupta, R., and Kyei, C. (2018). Predicting stock returns and volatility with investor sentiment indices: A reconsideration using a nonparametric causality-in-quantiles test. Bulletin of Economic Research, 70(1): 74−87.

Bathia, D., and Bredin, D. (2013). An examination of investor sentiment effect on G7 stock market returns. European Journal of Finance, 19(9): 909−937.

Bekaert, G., and Hoerova, M. (2014). The VIX, the variance premium and stock market volatility. Journal of Econometrics, 183(2): 181−192.

Ben Nasr, A., and Ajmi, A.N., and Gupta, R. (2014). Modelling the volatility of the Dow Jones Islamic Market World Index using a fractionally integrated time-varying GARCH (FITVGARCH) model. Applied Financial Economics, 24(14): 993−1004.

Ben Nasr, A. Lux, T., Ajmi, A.N., and Gupta, R. (2016). Forecasting the volatility of the Dow Jones Islamic stock market index: Long memory vs. regime switching. International Review of Economics and Finance, 45(1), 559−571.

Black, F. (1986). Noise. Journal of Finance, 41: 529−43.

Bouri, E., Gkillas, K., Gupta, R., and Pierdzioch, C. (2020). Forecasting realized volatility of Bitcoin: The role of the trade war. Computational Economics (forthcoming).

Breiman, L. (2001). Random forests. Machine Learning, 45: 5−32.

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

Chen, J. Tang, G., Yao, J., and Zhou, G. (Forthcoming). Investor attention and stock returns. Journal of Financial and Quantitative Analysis.

Clark, P. (1973). A subordinated stochastic process model with finite variances for speculative prices. Econometrica, 41(1): 135−155.

Conrad, C., and Loch, K. (2015). Anticipating long term stock market volatility. Journal of Applied Econometrics, 30: 1090−1114.

Copeland, T.E. (1976). A model for asset trading under the assumption of sequential information arrival. Journal of Finance, 31(4), 1149−1168.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. Journal of Financial Econometrics, 7: 174−196.

Da, Z., Engelberg, J., and Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. Review of Financial Studies, 28(1): 1−32.

Escobari, D., and Jafarinejad, M. (2019). Investors? Uncertainty and stock narket risk. Journal of Behavioral Finance, 20(3): 304−315.

Engle, R.F., and Rangel, J.G. (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. Review of Financial Studies, 21(3), 1187−1222.

Engle, R.F., Ghysels, E., and Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. Review of Economics and Statistics, 95(3): 776−797.

Gebka, B. (2014). The non-linear and linear impact of investor sentiment on stock returns: An empirical analysis of the US market. In: Ma, J., Wohar, M, ed. Recent advances in estimating nonlinear models: With applications in economics and finance. New York: Springer, 281−299.

Giot, P., Laurent, S., and Petitjean, M. (2010). Trading activity, realized volatility and jumps. Journal of Empirical Finance, 17(1), 168−175.

Greene, J., and Smart, S. (1999). Liquidity provision and noise trading: Evidence from the “Investment Dartboard" column. Journal of Finance, 54(5): 1885−1899.

Gu, S., Kelly, B., and Xiou, D. (2020). Empirical asset pricing via machine learning. Review of Financial Studies 33(5): 2223−2273.

Gupta, R. (2019). Manager sentiment and stock market volatility. Journal of Management Information and Decision Sciences, 22(1): 11−21.

Gupta, R., and Kyei, C. (2016). A nonlinear approach for predicting stock returns and volatility with the use of investor sentiment indices. Applied Economics, 48(31): 2895−2898.

Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor sentiment aligned: a powerful predictor of stock returns. Review of Financial Studies, 28: 791−837.

Jiang, F., Lee, J., Martin, X., and Zhou, G. (2019). Manager sentiment and stock returns. Journal of Financial Economics, 132: 126−149.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. American Economic Review, 105(3): 1177−1216.

Kaminska, I., and Roberts-Sklar, M. (2018). Volatility in equity markets and monetary policy rate uncertainty. Journal of Empirical Finance, 45: 68−83.

Liu, R., and Gupta, R. (2021). Investors' uncertainty and forecasting stock market volatility. Journal of Behavioral Finance. DOI: `https://doi.org/10.1080/15427560.2020.1867551`.

Liu, L., and Zhang, T. (2015). Economic policy uncertainty and stock market volatility. Finance Research Letters, 15: 99−105.

Ludvigson, S., Ma, S., and Ng, S. (Forthcoming). Uncertainty and business cycles: Exogenous impulse or endogenous response? American Economic Journal: Macroeconomics.

Ludvigson, S.C., and Ng, S. (2009). Macro factors in bond risk premia. Review of Financial Studies, 22(12): 5027−5067.

Ludvisgon, S.C., and Ng, S. (2011). A factor analysis of bond risk premia. In A. Ulah, and D. Giles (eds.), Handbook of Empirical Economics and Finance, 313−372. London: Chapman and Hall.

Mei, D., Liu, J., Ma, F., and Chen, W. (2017). Forecasting stock market volatility: Do realized skewness and kurtosis help?, Physica A: Statistical Mechanics and its Applications, 481: 153?159.

Mittnik, S., Robinzonov, N., and Spindler, M. (2015). Stock market volatility: Identifying major drivers and the nature of their impact. Journal of Banking and Finance 58: 1−14.

Müller, U. A., Dacorogna, M.M., Davé, R. D., Olsen, R. B., and Pictet, O. V. (1997). Volatilities of different time resolutions − Analyzing the dynamics of market components. Journal of Empirical Finance, 4: 213−239.

Olson, E., and Nowak, A. (2019). Sentiment?s effect on the variance of stock returns. Applied Economics Letters, 27(18): 1469−1473.

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

Poon, S-H., and Granger, C.W.J. (2003). Forecasting volatility in financial markets: A review. Journal of Economic Literature, 41(2): 478−539.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URLhttps://www.R-project.org/.

Rangel, J.G., and Engle, R.F. (2011). The factor-spline-GARCH model for high and low frequency correlations. Journal of Business and Economic Statistics 30(1), 109−124.

Rapach, D.E., Strauss, J.K., and Wohar, M.E. (2008). Forecasting stock return volatility in the presence of structural breaks, in Forecasting in the Presence of Structural Breaks and Model Uncertainty, in David E. Rapach and Mark E. Wohar (Eds.), Vol. 3 of Frontiers of Economics and Globalization, Bingley, United Kingdom: Emerald (May 2008), 381−416.

Salisu, A.A., Gupta, R., and Ogbonna, A.E. (2020). A moving average heterogeneous autoregressive model for forecasting the realized volatility of the US stock market: Evidence from over a century of data. International Journal of Finance and Economics, DOI: https://doi.org/10.1002/ijfe.2158.

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

Tiwari, A.K., Cunado, J., Gupta, R., and Wohar, M.E. (2018). Volatility spillovers across global asset classes: Evidence from time and frequency domains. The Quarterly Review of Economics and Finance, 70: 194−202.

Zhang, X. (2006). Information uncertainty and stock returns. Journal of Finance, 61(1): 105−137.

Zhou, G. (2018). Measuring Investor Sentiment. Annual Review of Financial Economics, 10: 239−259.

Table 1: Root-mean-squared-forecast-error ratios

Panel A: Results for *RV*

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	1.224	0.979	0.989	0.997	1.001	0.967
Window=100 / h=3	1.389	1.019	0.997	1.032	0.988	1.037
Window=100 / h=6	1.675	1.004	1.005	1.007	1.092	1.110
Window=100 / h=12	1.682	0.994	0.995	0.909	0.934	0.839
Window=120 / h=1	1.079	0.953	0.987	1.010	1.026	0.974
Window=120 / h=3	1.119	1.001	0.990	1.043	1.073	1.108
Window=120 / h=6	1.475	1.001	1.017	0.955	1.250	1.214
Window=120 / h=12	1.595	0.994	0.996	0.936	1.019	0.944
Window=140 / h=1	1.070	0.947	0.983	1.050	1.004	0.981
Window=140 / h=3	1.095	1.006	0.991	1.071	1.052	1.122
Window=140 / h=6	1.425	1.007	1.019	1.080	1.116	1.238
Window=140 / h=12	1.606	0.994	0.997	0.981	0.978	0.950

Panel B: Results for $\sqrt{RV}$

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	1.125	0.999	0.999	1.020	0.997	1.015
Window=100 / h=3	1.407	1.028	1.003	1.043	1.009	1.085
Window=100 / h=6	1.723	1.013	0.993	1.050	1.064	1.125
Window=100 / h=12	1.695	0.980	0.996	0.918	0.934	0.837
Window=120 / h=1	1.038	1.017	1.004	1.055	1.002	1.080
Window=120 / h=3	1.094	1.043	1.005	1.068	1.125	1.260
Window=120 / h=6	1.436	1.016	0.993	1.005	1.257	1.275
Window=120 / h=12	1.610	0.989	0.992	0.932	1.020	0.933
Window=140 / h=1	1.024	1.023	1.002	1.078	1.003	1.108
Window=140 / h=3	1.058	1.050	1.006	1.098	1.108	1.286
Window=140 / h=6	1.404	1.017	0.999	1.123	1.140	1.300
Window=140 / h=12	1.622	0.989	0.992	0.984	0.971	0.937

Note: This table reports RMSFE ratios, computed for out-of-sample forecasts. The first model is the baseline and the second model is the alternative model. A ratio larger than unity indicates that the alternative model outperforms the corresponding baseline model. Estimation is by a recursive window, where "window" in the first column gives the length of the training period. The parameter h denotes the forecast horizon (in months). The random forests are built using 2,000 trees.

Table 2: Comparing models by means of the Clark-West test

Panel A: Results for RV

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	0.049	0.619	0.317	0.141	0.055	0.408
Window=100 / h=3	0.057	0.084	0.226	0.053	0.314	0.069
Window=100 / h=6	0.021	0.242	0.134	0.167	0.037	0.074
Window=100 / h=12	0.001	0.765	0.783	0.803	0.140	0.740
Window=120 / h=1	0.004	0.726	0.392	0.162	0.008	0.321
Window=120 / h=3	0.009	0.297	0.330	0.077	0.022	0.064
Window=120 / h=6	0.000	0.234	0.054	0.332	0.014	0.081
Window=120 / h=12	0.002	0.747	0.737	0.947	0.167	0.698
Window=140 / h=1	0.006	0.759	0.492	0.111	0.019	0.453
Window=140 / h=3	0.015	0.221	0.187	0.032	0.073	0.048
Window=140 / h=6	0.001	0.203	0.045	0.119	0.052	0.076
Window=140 / h=12	0.001	0.809	0.691	0.862	0.495	0.695

Panel B: Results for $\sqrt{RV}$

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	0.031	0.212	0.207	0.060	0.016	0.143
Window=100 / h=3	0.067	0.064	0.225	0.060	0.187	0.112
Window=100 / h=6	0.029	0.178	0.387	0.090	0.039	0.067
Window=100 / h=12	0.001	0.907	0.586	0.835	0.199	0.764
Window=120 / h=1	0.080	0.060	0.166	0.075	0.063	0.066
Window=120 / h=3	0.005	0.100	0.222	0.095	0.043	0.086
Window=120 / h=6	0.001	0.204	0.233	0.190	0.014	0.072
Window=120 / h=12	0.003	0.780	0.862	0.942	0.193	0.731
Window=140 / h=1	0.092	0.041	0.207	0.054	0.086	0.048
Window=140 / h=3	0.011	0.083	0.214	0.056	0.083	0.081
Window=140 / h=6	0.002	0.208	0.182	0.064	0.073	0.076
Window=140 / h=12	0.002	0.770	0.884	0.741	0.591	0.725

Note: This table reports the results (p-values) of the Clark-West test of equal mean-squared prediction errors. The null hypothesis is that the alternative model has the same out-of-sample forecasting performance as the baseline model. The alternative hypothesis is that the alternative model performs better than the baseline model. Results are based on Newey-West robust standard errors. Estimation is by a recursive window, where "window" in the first column gives the length of the training period. The parameter h denotes the forecast horizon (in months). The random forests are built using 2,000 trees.

Table 3: Test results for good and bad volatility

Panel A: Results for $\sqrt{RVG}$

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	0.054	0.063	0.049	0.122	0.017	0.130
Window=100 / h=3	0.055	0.285	0.926	0.087	0.146	0.085
Window=100 / h=6	0.027	0.300	0.670	0.099	0.030	0.065
Window=100 / h=12	0.001	0.942	0.871	0.779	0.209	0.802
Window=120 / h=1	0.068	0.147	0.052	0.123	0.026	0.101
Window=120 / h=3	0.002	0.527	1.000	0.097	0.060	0.085
Window=120 / h=6	0.000	0.147	0.717	0.134	0.015	0.046
Window=120 / h=12	0.002	0.962	0.728	0.941	0.218	0.759
Window=140 / h=1	0.095	0.108	0.055	0.090	0.010	0.080
Window=140 / h=3	0.006	0.508	1.000	0.077	0.083	0.071
Window=140 / h=6	0.001	0.137	0.695	0.079	0.066	0.043
Window=140 / h=12	0.001	0.965	0.774	0.784	0.608	0.750

Panel B: Results for $\sqrt{RVB}$

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	0.034	0.515	0.042	0.148	0.020	0.184
Window=100 / h=3	0.058	0.067	0.938	0.058	0.217	0.099
Window=100 / h=6	0.019	0.122	0.414	0.110	0.040	0.064
Window=100 / h=12	0.001	0.266	0.888	0.679	0.140	0.789
Window=120 / h=1	0.011	0.557	0.064	0.203	0.022	0.071
Window=120 / h=3	0.007	0.108	0.898	0.097	0.068	0.084
Window=120 / h=6	0.001	0.121	0.161	0.164	0.012	0.065
Window=120 / h=12	0.003	0.217	0.824	0.892	0.203	0.731
Window=140 / h=1	0.016	0.626	0.091	0.124	0.071	0.062
Window=140 / h=3	0.023	0.086	0.891	0.066	0.131	0.077
Window=140 / h=6	0.001	0.108	0.164	0.039	0.054	0.053
Window=140 / h=12	0.002	0.239	0.810	0.612	0.601	0.716

Note: This table reports the results (p-values) of the Clark-West test of equal mean-squared prediction errors. The null hypothesis is that the alternative model has the same out-of-sample forecasting performance as the baseline model. The alternative hypothesis is that the alternative model performs better than the baseline model. Results are based on Newey-West robust standard errors. Estimation is by a recursive window, where “window” in the first column gives the length of the training period. The parameter h denotes the forecast horizon (in months). The random forests are built using 2,000 trees.

Appendix

Table A1: Mean-absolute-forecast-error ratios

Panel A: Results for *RV*

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	1.163	1.000	0.997	0.937	1.045	0.976
Window=100 / h=3	1.331	1.007	0.995	0.999	0.990	0.990
Window=100 / h=6	1.593	0.977	1.020	0.928	1.151	1.065
Window=100 / h=12	1.876	0.979	0.996	0.840	0.986	0.807
Window=120 / h=1	1.055	0.991	1.019	0.921	1.077	1.002
Window=120 / h=3	1.211	0.985	0.993	1.008	1.088	1.073
Window=120 / h=6	1.524	0.964	1.042	0.900	1.199	1.084
Window=120 / h=12	1.803	0.981	1.000	0.853	1.150	0.962
Window=140 / h=1	1.056	0.979	1.015	0.981	1.033	1.007
Window=140 / h=3	1.185	0.995	0.990	1.069	1.049	1.105
Window=140 / h=6	1.476	0.968	1.048	1.024	1.057	1.099
Window=140 / h=12	1.910	0.983	1.001	0.961	1.044	0.987

Panel B: Results for $\sqrt{RV}$

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	1.088	0.999	0.985	1.001	1.037	1.022
Window=100 / h=3	1.381	1.000	0.990	0.988	1.022	1.000
Window=100 / h=6	1.590	0.988	1.009	0.966	1.116	1.075
Window=100 / h=12	1.903	0.970	0.998	0.842	0.997	0.813
Window=120 / h=1	1.048	1.022	0.982	0.990	1.070	1.063
Window=120 / h=3	1.245	0.989	0.986	1.000	1.119	1.091
Window=120 / h=6	1.497	0.978	1.022	0.937	1.184	1.109
Window=120 / h=12	1.817	0.988	0.996	0.847	1.161	0.968
Window=140 / h=1	1.030	1.038	0.971	1.035	1.062	1.109
Window=140 / h=3	1.206	0.998	0.985	1.059	1.089	1.133
Window=140 / h=6	1.471	0.973	1.036	1.061	1.051	1.124
Window=140 / h=12	1.932	0.992	0.994	0.969	1.038	0.991

Note: This table reports MAFE ratios, computed for out-of-sample forecasts. The first model is the baseline and the second model is the alternative model. A ratio larger than unity indicates that the alternative model outperforms the corresponding baseline model. Estimation is by a recursive window, where "window" in the first column gives the length of the training period. The parameter h denotes the forecast horizon (in months). The random forests are built using 2,000 trees.

Table 2: Test results for a rolling-estimation window

Panel A: Results for *RV*

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	0.085	0.092	0.693	0.173	0.004	0.131
Window=100 / h=3	0.066	0.167	0.093	0.195	0.301	0.210
Window=100 / h=6	0.048	0.417	0.143	0.109	0.063	0.085
Window=100 / h=12	0.004	0.859	0.305	0.465	0.084	0.535
Window=120 / h=1	0.002	0.790	0.072	0.454	0.018	0.191
Window=120 / h=3	0.000	0.808	0.289	0.382	0.004	0.048
Window=120 / h=6	0.000	0.124	0.981	0.513	0.012	0.000
Window=120 / h=12	0.016	0.877	0.123	0.936	0.038	0.563
Window=140 / h=1	0.005	0.179	0.030	0.283	0.317	0.100
Window=140 / h=3	0.002	0.216	0.852	0.732	0.093	0.044
Window=140 / h=6	0.004	0.789	0.724	0.177	0.032	0.006
Window=140 / h=12	0.000	0.891	0.267	0.844	0.370	0.609

Panel B: Results for $\sqrt{RV}$

Window / horizon	Model 1 / Model 2	Model 2 / Model 3	Model 3 / Model 4	Model 4 / Model 5	Model 5 / Model 6	Model 2 / Model 6
Window=100 / h=1	0.083	0.088	0.020	0.432	0.062	0.072
Window=100 / h=3	0.077	0.403	0.299	0.225	0.530	0.454
Window=100 / h=6	0.052	0.092	0.144	0.146	0.051	0.075
Window=100 / h=12	0.005	0.830	0.605	0.174	0.109	0.490
Window=120 / h=1	0.004	0.235	0.033	0.593	0.064	0.109
Window=120 / h=3	0.003	0.129	0.937	0.281	0.001	0.060
Window=120 / h=6	0.000	0.655	0.488	0.452	0.019	0.001
Window=120 / h=12	0.016	0.764	0.845	0.893	0.015	0.481
Window=140 / h=1	0.010	0.059	0.898	0.071	0.377	0.165
Window=140 / h=3	0.002	0.356	0.878	0.831	0.096	0.084
Window=140 / h=6	0.004	0.170	0.182	0.488	0.023	0.005
Window=140 / h=12	0.001	0.091	0.803	0.750	0.605	0.596

Note: This table reports MAFE ratios, computed for out-of-sample forecasts. The first model is the baseline and the second model is the alternative model. A ratio larger than unity indicates that the alternative model outperforms the corresponding baseline model. Estimation is by a rolling window, where "window" in the first column gives the length of the rolling-estimation window. The parameter *h* denotes the forecast horizon (in months). The random forests are built using 2,000 trees.