Climate Risks and State-Level Stock-Market Realized Volatility

September 2022

Matteo Bonato*, Oğuzhan Çepni†, Rangan Gupta‡, Christian Pierdzioch§

Abstract

We analyze the predictive value of climate risks for state-level realized stock-market volatility, computed, along with other realized moments, based on high-frequency intra-day U.S. data (September, 2011 to October, 2021). A model-based bagging algorithm recovers that climate risks have predictive value for realized volatility at intermediate and long (one and two months) forecast horizons. This finding also holds for upside ("good") and downside ("bad") realized volatility. The benefits of using climate risks for predicting state-level realized stock-market volatility depend on the shape and (as-)symmetry of a forecaster's loss function.

JEL Classifications: C22; C53; G10; G17; Q54.

Keywords: Finance; State-level data; Realized stock-market volatility; Climate-related predictors; Forecasting

Conflicts of interest: The authors declare no conflict of interest.

*Department of Economics and Econometrics, University of Johannesburg, Auckland Park, South Africa; IPAG Business School, 184 Boulevard Saint-Germain, 75006 Paris, France. Email address: matteobonato@gmail.com.

†Copenhagen Business School, Department of Economics, Porcelænshaven 16A, Frederiksberg DK-2000, Denmark; Central Bank of the Republic of Turkey, Haci Bayram Mah. Istiklal Cad. No:10 06050, Ankara, Turkey. Email address: oce.eco@cbs.dk.

‡Department of Economics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa. Email address: rangan.gupta@up.ac.za.

§Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany; Email address: macroeconomics@hsu-hh.de.

1 Introduction

Wachter (2013) and, more recently, Tsai and Wachter (2015), develop theoretical models in which aggregate consumption follows a normal distribution with low volatility most of the time but with some probability of a far out-in-the-left-tail realization of consumption, creating risk associated with rare disaster events. The possibility of such a poor outcome substantially raises the equity premium, and the time-variation in the probability of such a disaster fosters high stock-market volatility. Extending the benchmark theoretical models of rare disaster risks (Rietz, 1988; Barro 2006) by making the probability of a disaster stochastically varying over time (and agents having recursive preferences rather than power utility functions) Wachter (2013) and Tsai and Wachter (2015) are able, for reasonable values of the structural parameters of their models, to generate stock-market volatility close to that observed in U.S. data. In other words, a theoretical link exists between rare disaster risk and stock-market volatility, in addition to the well-known link between rare disaster risk and the equity premium.

In light of these recent theoretical advances, our empirical research aims to forecast U.S. stock-market volatility in the presence of rare disaster risks, but at the level of individual states rather than at the aggregate U.S. level. The underlying reason for taking such a regional perspective is derived from the premise that core business activities of firms often occur close to their headquarters (Pirinsky and Wang, 2006; Chaney et al., 2012) and, hence, equity prices should contain a non-negligible regional component, so much so that investors overweight local firms in their portfolios (Coval and Moskowitz, 1999, 2001; Korniotis and Kumar, 2013). Obviously, then, the forecasting exercise that we undertake in this research should be of immense value to investors, given that accurate forecasts of stock-market volatility carries widespread implications for portfolio selection, derivative pricing, risk management, and also for policy-making (Poon and Granger 2003; Rapach et al., 2008). In the process, we add to the vast associated literature on U.S. stock-market volatility, wherein earlier researchers have used a wide array of (univariate and multivariate, linear and nonlinear) forecasting models and (behavioral, macroeconomic, and financial) predictors (see, for example, Ben Nasr et al., 2016; Liu et al., 2020; Gupta et al., 2021; Liu and Gupta, 2022; Salisu et al., 2022a, b; Segnon et al., forthcoming), by bringing to the forefront a regional perspective and by considering the role of rare disaster risk.

At this juncture, it is important to highlight the following integral features of our analyses: First, traditionally, disaster events are generally captured by cumulative declines in output and/or consumption of at least 10% over one or more years (see Ćorić (2021), and Ćorić and Šimić (2021) for detailed discussions in this regard, with these authors extending the original data sets of Barro and Ursúa (2008, 2012)). Given this, a major obstacle to full-fledged empirical verification of rare disaster models is that individual countries rarely face such major disasters, resulting in a small

sample problem inherent in the use of actual rare disasters, which, in turn, explains why earlier researchers studying the implications of rare disasters for asset pricing have relied theoretical models calibrated on rare-disaster-risk probabilities derived from historical cross-country evidence of major declines in output and/or consumption.[1] In contrast, we consider, in order to capture the role of rare disaster risks in predicting daily state-level volatility, daily proxies related to corresponding state-level climate risks, besides those of the aggregate U.S., as motivated by the growing literature on "Climate Finance" (Engle et al., 2020; Giglio et al., 2021; Stroebel and Wurgler, 2021; van Benthem et al., 2022).

The risks associated with climate change can be typically categorized into physical risks and transition risks. The former arises due to rising temperatures, higher sea levels, heavy storms and floods, and wildfires. The latter arise due to a gradual switch-over to a low-carbon economy and comprise risks due to changes in climate policy, the development of disruptive green technologies, and changing consumer preferences. Hence, every future scenario includes climate-related financial risks (though their level, as well as the source of uncertainty, may vary) due to the occurrence of rare disasters (Battiston et al., 2021; Flori et al., 2021). Unsurprisingly, researchers now routinely use climate risks, as captured by textual and narrative analyses of climate-change-related news (involving natural disasters, global warming, international summits, and climate policy) or via movements of temperature and precipitation, as high-frequency proxies of potential forthcoming rare disaster risks (Choi et al., 2020; Kapfhammer et al., 2020; Faccini et al., 2021; Bonato et al. 2022, forthcoming; Bua et al., 2022). In line with this research, we use the information content of seasonal, predictable, and abnormal patterns of temperature, precipitation, heating degree days, cooling degree days, and wind speed to capture climate-related risks. In addition, we rely on the daily Google Search Volume Index (SVI) on the topic "global warming" in a particular state and the news trends (NT) function of the Bloomberg terminal to compile the news counts involving the term "climate change" for every state. The intensity of such keywords-based searches can reflect the perceived transition-risks component of climate change, in addition to physical risks.

We forecast state-level daily stock-market realized volatility using an extended version of the heterogeneous autoregressive realized volatility (HAR-RV) model of Corsi (2009). Our extension incorporates the role of climate risks of the corresponding state and the overall U.S. over the period from September, 2011 to October, 2021. Realized volatility, as captured by the square root of the sum of squared intraday log-returns (of equities domiciled in a given state) over a day (following Andersen and Bollerslev (1998)), is considered as an accurate, observable, and unconditional met-

[1] In this regard, Berkman et al. (2011, 2017), propose a solution to the small sample problem that would make the empirical estimation of such models feasible. They focus on a large sample of potential disasters, i.e., international political crises, that are likely to cause significant changes in perceived rare disaster probabilities. Using a detailed database of international political crises, they document that various international crises, over the period from 1918 to 2006, affect equity returns and volatility of a large number of developed and emerging economies.

ric of volatility (McAleer and Medeiros, 2008), unlike the measures of the same derived from the popular generalized autoregressive conditional heteroscedastic (GARCH) and stochastic volatility (SV) models. At the same time, the benchmark HAR-RV model, though being simple in formulation, is elaborate enough to capture long-memory and multi-scaling properties often observed for stock-market volatility (Mei et al., 2017), as it uses volatilities from different time resolutions to model and to predict realized volatility. Therefore, the HAR-RV model can easily mimic the idea of the so-called heterogeneous-market hypothesis (Müller et al. 1997), which states that different types of market participants that populate the stock market are characterized by differences in their sensitivity to flows of information that is available at different time horizons.

Since we have as many as 14 state-level and nation-wide climate-related predictors, as well as intraday-data-based realized moments such as realized jumps, realized upside and downside tail risks, realized skewness, and realized kurtosis, for the 50 U.S. states, estimation of such a large number of prediction models is not feasible. Due to this, we use a model-based bagging (MOBA) algorithm to analyze whether climate risks help to improve predictive performance beyond that of benchmark HAR-RV models (without and with the moments-based controls). We must point out that the decision to look at both state-specific and aggregate climate risks involving temperature and precipitation emanates from the recent study by Gil-Alana et al. (forthcoming), which points towards a starkly contrasting degree of persistence of these climate shocks for the U.S. as a whole, and across the states. In essence, we need to account for the heterogeneity involving the underlying data-generating processes of aggregate and localized rare disaster risks that will originate from climate change. This is more so, given that recent regional analyses have revealed that climate risks impact state-level business cycles and uncertainties (Sheng et al., 2022a, b), which, in turn, are likely to feed indirectly into state-level stock-market realized volatility, given the importance of these two predictors identified by the studies cited above.

To the best of our knowledge, this is the first paper to predict state-level realized volatility based on rare disaster events emanating from climate risk. Besides being of relevance to investors, predicting daily realized state-level stock-market volatility traces out the future path of state-level financial uncertainty, which can then be incorporated into mixed data sampling (MIDAS) models to predict low-frequency (monthly or quarterly) real-activity variables and allow policymakers to come up with appropriate localized policy responses to tackle a potential recessionary impact of uncertainty in a time-efficient-manner. We organize the remainder of this research as follows. In Section 2, we describe our data. In Section 3, we describe the methods we use in our empirical research. In Section 4, we report our empirical results. In Section 5, we conclude.

2 Data

In order to study the predictive value of climate risks for state-level stock-market volatility, we study 5-min intraday data on the Bloomberg State level stock-market indices covering 50 US States. Bloomberg terminal composes these state-level stock-market indices as the capitalization-weighted index of equities domiciled in a state.[2]

We compute the sum of squared intraday returns of the indices and, thus, consider the classical estimator of realized variance (Andersen and Bollerslev, 1998). We obtain the state-level realized variance as follows (we do not introduce a subindex to denote a state for notational convenience).

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2, \tag{1}$$

where $r_{t,i}$ denotes the intraday $M \times 1$ return vector, and $i = 1, ..., M$ denotes the number of intraday returns. Equipped with Equation (1), we compute the realized volatility as the standard deviation of the realized variance: $\tilde{RV}_t = \sqrt{RV_t}$. Given that the realized variance typically exhibit occasional large peaks, we study, in our empirical analysis, the realized volatility defined as the square root of the realized variance given in Equation (1).[3]

– Figures 1 to 3 about here. –

We plot the mean of the realized volatility for every state in Figure 1. The mean of realized volatility varies substantially across states, with returns of companies domiciled in New Mexico, Vermont, and Wyoming displaying a particularly large average realized volatility. Hence, realized volatility exhibits a substantial heterogeneity in the cross-sectional dimension. This cross-sectional heterogeneity can also be seen in Figure 2, which we use to plot the time series of the cross-sectional mean of realized volatility (black line) along with the time series of the cross-sectional mean plus one standard deviation (gray line). The cross-sectional mean exhibits the usual temporal peaks that give rise to the volatility clustering characteristic of financial market data. When we add one standard deviation to the cross-sectional mean, we observe again the type of substantial cross-sectional heterogeneity that we already observed in Figure 3. However, a look at the time-series dimension of the data further reveals that this cross-sectional heterogeneity is scattered throughout the entire sample period and not concentrated in a few short sub-sample periods. Finally, we

[2] Upon further detailed inspection of the raw data, we spotted periods (a small fraction of the data for most states) during which the data exhibit more or less a straight trend, which probably indicates that Bloomberg stopped producing the indices during those periods and then interpolated the resulting gap in the data. We keep those periods in the sample for computing the descriptive statistics we shall present in this section, but we remove those periods when we compute the estimation and simulation results we shall present in Section 4. We refer the interested reader to the the Supplementary Material, where we present further details regarding the "gap periods'. In the Supplementary Material, we also report some further descriptive statistics of the data.

[3] In order to avoid notational clutter, we do not introduce a separate abbreviation for the realized volatility, but emphasize that it is the realized volatility that we use to estimate our prediction models and to evaluate predictions (see Sections 3.1 –3.3).

plot in Figure 3 the auto-correlation function of realized volatility, where we average the coefficients of auto-correlation across states. The auto-correlation function exhibits the usual slowly decaying pattern and, thereby, illustrates that the HAR-RV model is a good starting point for our empirical analysis.

We collect daily weather data from the Bloomberg terminal, as compiled by the National Climatic Data Center (NCDC), for the US and our sample of 50 states. The weather data captures meteorological phenomena along several dimensions, including temperature, precipitation, number of heating degree days (HDD), number of cooling degree days (CDD), and wind speed as described below:

- **Temperature**: The average temperature (usually of the high and low) that was observed between 7am and 7pm local time, expressed in Fahrenheit.
- **HDD**: The number of degrees below 65 degrees Fahrenheit of the mean temperature used to estimate the energy needed to heat a building.
- **CDD**: The number of degrees above 65 degrees Fahrenheit of the mean temperature used to estimate the energy needed to cool a building.
- **Precipitation**: The amount of rain, snow, sleet or hail that falls in a specific location.
- **Wind speed**: Average of sustained winds which does not include wind gust, expressed in knots.

Like Choi et al., (2020), we decompose the weather-related variables into three components that account for seasonal, predictable, and abnormal patterns. In particular, for each country, i, and day, t, we compute the daily $WeatherMeasure_{it}$ using the following formula:

$$WeatherMeasure_{it} = Aver\_WeatherMeasure_{it} + Mon\_WeatherMeasure_{it} + Ab\_WeatherMeasure_{it}, \quad (2)$$

where $WeatherMeasure_{it} = \{$temperature, HDD, CDD, precipitation, wind speed$\}$, and the term $Aver\_WeatherMeasure_{it}$ denotes the mean of $WeatherMeasure_{it}$ in state i over the 120 months prior to t. Moreover, the variable $Mon\_WeatherMeasure_{it}$ denotes the difference of the mean of the deviation of the $WeatherMeasure_{it}$ from the monthly average temperature in state i in the same calendar month over the last ten years and $Aver\_WeatherMeasure_{it}$. Finally, the variable $Ab\_WeatherMeasure_{it}$ is the remainder (that is, the abnormal deviation of weather conditions) and, hence, captures unusual deviations from local weather conditions. For this reason, we focus on this variable in our analysis. We standardize the abnormal deviations, commonly known as the standardized anomaly (Kim et al., 2021).

It is also important to measure how people react to abnormal local weather conditions. To this end, we use the daily Google Search Volume Index (SVI) of the topic "global warming" in a state (and a nation-wide measure for the US). This index captures retail investor attention to climate risks. Moreover, the attention to abnormal weather conditions also can be fostered via communication channels and the media. Hence, we use the news trends (NT) function of the Bloomberg terminal to compile data on the news counts, including the term "climate change" for every state (and a nation-wide measure for the US) The NT function, being based on a vast archive of news stories and social media posts from over 150,000 sources, renders it possible to search specific keywords and obtain the historical volume of relevant news. We use this NT-based measure to capture institutional investors' attention. Institutional investors (that is, market participants who work in asset management, banking, and institutional financial services) are known to use Bloomberg terminals as a major source of information (Ben-Rephael et al., 2017). Hence, we assume that institutional investors, who have more resources and incentives to pay attention to news quickly, will follow the news that appears on Bloomberg terminal.

3 Methods

3.1 Defining the Core Model

We use the core HAR-RV model of Corsi (200) as our benchmark model. The following equation represents this model:

$$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{w,t} + \beta_3 RV_{m,t} + u_{t+h}, \tag{3}$$

where estimation is done by the ordinary-least-squares techniques (OLS), $\beta_j, j = 0, .., 3$ are the coefficients to be estimated, u_{t+h} denotes the usual disturbance term, and RV_{t+h} is the average realized volatility over the prediction horizon, h. We set $h = 1, 5, 22, 44$ to cover short and long prediction horizons. The model features three predictors: the daily realized volatility, RV_t, the weekly realized volatility, $RV_{t,w}$, and the monthly, $RV_{t,m}$ realized volatility. We compute the weekly realized volatility as the average realized volatility from period $t-5$ to period $t-1$, and the monthly realized volatility as the average realized volatility from period $t-22$ to period $t-1$.

3.2 Constructing the Extended Model

We are mainly interested in the question of whether extending the core HAR-RV model by the climate-related predictors improves the preditive performance of the model. In total, we have 14 (state-level and nation-wide) climate-related predictors and, therefore, a total of 2^{14} combinations of these predictors that we can use to extend the core HAR-RV model. Given that our plan is to

estimate the predictions models for the 50 states in our sample, and given that we study a large dataset that consists of daily data, estimation of such a large number of prediction models is not feasible. We instead use the following model-based bagging (MOBA) algorithm to analyze whether the climate-related predictors help to improve predictive performance beyond that of the core HAR-RV model.

□

MOBA Algorithm

1. Fix the number of simulation runs, S_{max}. Set $s = 1$. At the start of every simulation run, use sampling without replacement to split the data into an estimation, a validation, and a test sample.

   (a) Fix a maximum number, n_{max}, of iterations and some small number, v. Set $n = 1$.

      i. For $n = 1$: Estimate the core HAR-RV model on the training data by OLS. Use the model along with the validation data to compute a vector of initial validation predictions, $RV_{t+h,1}$.

      ii. Initialize the vector of average validation predictions for the extended model: $\bar{RV}_{t+h,n=1} = RV_{t+h,1}$.

      iii. For $2 \leq n \leq n_{max}$: Sample a new prediction model.

         A. Sample the number, $N \in \{1, 2, ..., 14\}$, of climate-related predictors to be used for an extended prediction model.

         B. Use N to sample without replacement from the climate-related predictors.

         C. Use the sample climate-related predictors to form an extended prediction model of the format

         $$RV_{t+h} = \beta_0 + \beta_1 RV_t + \beta_2 RV_{w,t} + \beta_3 RV_{m,t} + \theta X_t + u_{t+h}, \tag{4}$$

         where X_t is the $N \times 1$ vector of sampled predictors and θ is the $1 \times N$ corresponding coefficient vector to be estimated.

         D. Estimate the extended model on the training data by OLS. Use the estimated extended model and the validation data to compute a new validation-prediction vector, $RV_{t+h,2}$.

      iv. Update the vector of average validation predictions, $\bar{RV}_{t+h,n} = [\bar{RV}_{t+h,n-1}(n-1) + RV_{t+h,2}]/n$.

      v. Update: $RV_{t+h,2} = \bar{RV}_{t+h,n}$.

   (b) Track the maximum absolute percentage change, $\%\bar{RV}_{t+h,n}$, in the vector of average validation predictions.

i. If $\%\bar{RV}_{t+h+1} \leq v$:

A. Apply the extended models estimated in Step (a), iii. to the test data, and average the predictions from the models to compute a vector of out-of-bag predictions.

B. Return to Step 1 and start a new simulation run.

ii. If $\%\bar{RV}_{t+h+1} > v$: increase n by one increment, $n+1$.

A. If $n+1 \leq n_{max}$: Return to Step (a), iii. and sample a new extended model.

B. If $n+1 > n_{max}$: Proceed as in Step (b), i., A. and. Then return to Step 1 and start a new simulation run.

2. Update $s = ++1$.

■

Our MOBA algorithm has only two hyper-parameters: n_{max} and v, plus the maximum number of simulation runs. In our empirical analysis, we set $n_{max} = 100$, $v = 0.01$, and $S_{max} = 500$, but the algorithm typically terminates much earlier before n_{max} is reached and that is cheap in terms of computational time. For example, when the algorithm terminates after about 30 iterations, a standard boosting algorithm with 14 weak learners would have to terminate after only about two iterations to reach a similar average computational speed. The computational speed of the MOBA algorithm also is high as compared to random forests, an algorithmic statistical-learning technique tailored to analyze datasets with many predictors. Forecasts computed by means of random forests typically require the estimation of many regression trees to build a random forest (for a textbook exposition of various statistical learning algorithms, see Hastie et al. 2009). Our MOBA algorithm further implies that the forecasts implied by the core HAR-RV model are, by construction, nested versions of the forecasts computed using the algorithmically constructed extended model, making it easy to compare the forecasts implied by the core and extended models through standard statistical tests and, thereby, to trace out whether climate risks contribute to improving the accuracy of forecasts of state-level realized stock-market volatility.

3.3 Evaluating Forecasts

As baseline statistics to evaluate the out-of-sample predictions, we use the out-of-bag $R^2_{oob,s}$ statistic defined for simulation run, s, as

$$R^2_{oob,s} = 1 - PE_{m,s}/PE_{c,s}, \tag{5}$$

where $PE_{j,s}, j = m, c$ denotes the sum of the squared out-of-bag prediction error for simulation run, s, for the MOBA, m, and the HAR-RV core, c, models. We then form, for every state, the average, R^2_{oob} of the resulting vector of $R^2_{oob,s}$ statistics across all simulation runs. A positive R^2_{oob} statistic indicates that the extended model produces on average a lower out-of-bag prediction error than

the core HAR-RV model. In a final step, we use the sampling distribution of the $R^2_{oob,s}$ statistics to compute a p-value, pval $= 1 - \sum_s \mathbf{1}(R^2_{oob,s} > 0)/S$, where $\mathbf{1}(\cdot)$, denotes the indicator function.

The out-of-bag R^2_{oob} statistic represents the case of a quadratic symmetric loss function, that is, the basic assumption underlying the statistic is that the loss a forecaster suffers is increasing in the quadratic prediction error, irrespective of whether a model under- or over-predicts realized volatility. In practical settings, however, behavioral biases or trading strategies involving derivative securities (e.g., trading strategies involving options) may imply that a forecaster has a loss function that is asymmetric in the prediction error. We account for this possibility by studying the following loss function (Elliott et al. 2005):

$$L(p, a) = [a + (1 - 2a)I_{[PE_{t+h}<0]}]|PE_{t+h}|^p. \tag{6}$$

For $p = 1$, we obtain a lin-lin loss function (the absolute prediction error matters), and for $p = 2$ we obtain a quad-quad loss function (the quadratic prediction error matters). The shape parameter, $a \in (0, 1)$, determines the asymmetry of the loss function, where $a = 0.5$ gives a symmetric loss function. Hence, for $a = 0.5$ and $p = 1$, the loss is a function of the absolute prediction error, sand for $a = 0.5$ and $p = 2$ we get a loss function that is symmetric in the squared prediction error. For $a > 0.5$, a forecaster suffers a larger loss from an under-prediction of realized volatility (as compared to an over-prediction of the same absolute size), and for $a < 0.5$ an over-prediction gives rise to a larger loss than a comparable under-prediction.

Upon using Equations (5) and (6), we can then compute the out-of-bag R^2_{oob} statistic for alternative shapes of the loss function for both the core HAR-RV model and the MOBA models and examine whether specific types of forecasters (as represented by the different shapes of the loss function) benefit from considering climate risks for predicting state-level realized stock-market volatility.

3.4 Computing Control Variables

While our primary focus is whether the climate risks improve predictions of RV from the core HAR-RV model, we also consider an extended HAR-RV model as an alternative benchmark model. The extended HAR-RV model features as control variables several other intraday-data-based realized moments that have been extensively studied in earlier research. Specifically, we consider the following control variables: realized jumps, $JUMPS$, realized upside and downside tail risks, TR_u and TR_d, and realized skewness, RSK, as well as realized kurtosis, RKU.

We study RSK to trace the asymmetry of the returns distribution, while we use RKU to capture

its extremes (see, e.g., Amaya et al., 2015). We calculate RSK and RKU as follows:

$$RSK_t = \frac{\sqrt{M}\sum_{i=1}^{M} r_{(i,t)}^3}{RV_t^{3/2}}, \tag{7}$$

$$RKU_t = \frac{M\sum_{i=1}^{M} r_{(i,t)}^4}{RV_t^2}. \tag{8}$$

The scaling terms, $(M)^{1/2}$ and M, turn RSK and RKU into their daily values.

We use the formula derived by Barndorff-Nielsen and Shephard (2004) to trace out the realized jumps, and use the fact that the realized variance converges into a discontinuous (jump) and a permanent component. We have

$$\lim_{M\to\infty} RV_t^2 = \int_{t-1}^{t} \sigma^2(s)ds + \sum_{j=1}^{N_t} k_{t,j}^2, \tag{9}$$

where N_t denotes the number of jumps within day t and $k_{t,j}$ denotes the jump size. It follows from Equation (9) that RV_t is a consistent estimator of the jump contribution plus the integrated variance $\int_{t-1}^{t} \sigma^2(s)ds$. Building on asymptotics, Barndorff-Nielsen and Shephard (2004, 2006) show that

$$\lim_{M\to\infty} BV_t^2 = \int_{t-1}^{t} \sigma^2(s)ds, \tag{10}$$

where BV_t denotes the daily realized bipolar variation, which is defined as

$$BV_t = \mu_1^{-2}\left(\frac{M}{M-1}\right)\sum_{i=2}^{M} |r_{t,i-1}||r_{i,t}| = \frac{\pi}{2}\sum_{i=2}^{M} |r_{t,i-1}||r_{i,t}|, \tag{11}$$

where we define

$$\mu_a = E(|Z|^a), Z \sim N(0,1), a > 0. \tag{12}$$

Upon using the continuous component of realized variance, we define the consistent estimator of the pure daily jump contribution:

$$J_t = RV_t - BV_t. \tag{13}$$

The formal test estimator proposed by Brandorff-Nielsen and Shephard (2006) can be used to inspect the significance of the jumps. Specifically, we make use of the following test statistic:

$$JT_t = \frac{RV_t - BV_t}{(v_{bb} - v_{qq})\frac{1}{N}QP_t}, \tag{14}$$

where $v_{bb} = \left(\frac{\pi}{2}\right) + \pi - 3$ and $v_{qq} = 2$, and QP_t is defined as the daily Tri-Power Quarticity:

$$TP_t = M \frac{M}{M-2} \left(\frac{\Gamma(0.5)}{2^{2/3}\Gamma(7/6)} \right) \sum_{i=3}^{M} |r_{t,i}|^{4/3} |r_{t,i-1}|^{4/3} |r_{t,i-2}|^{4/3}, \tag{15}$$

which converges to

$$TP_t \to \int_{t-1}^{t} \sigma^4(s) ds, \tag{16}$$

even in the presence of jumps. For each t, $JT_t \sim N(0,1)$ as $M \to \infty$.

Equation (13) makes clear that the jump contribution to RV_t is non-negative. Hence, in order to rule out negative empirical contributions, we redefine the jump measure as (see Zhou and Zhu, 2012):

$$RJ_t = \max(RV_t - BV_t; 0). \tag{17}$$

Last, we compute two measures of tail risk using the Hill estimator (Hill, 1975). We construct $X_{t,i}$, the set of reordered intraday returns $r_{t,i}$, in such a way that

$$X_{t,i} \geq X_{t,j} \text{ for } i < j. \tag{18}$$

We compute the Hill positive tail risk estimator (our predictor TR_u) as

$$H_t^{up} = \frac{1}{k} \sum_{j=1}^{k} \ln(X_{t,i}) - \ln(X_{t,k}) \tag{19}$$

and the negative tail risk estimator (our predictor TR_d) as

$$H_t^{down} = \frac{1}{k} \sum_{j=n-k}^{n} \ln(X_{t,i}) - \ln(X_{t,n-k}) \tag{20}$$

where k is the observation denoting the chosen α tail interval.

4 Empirical Results

We present in Panel A of Figure 4 violin plots to summarize our baseline findings for the 50 states.[4] The violin plots depict the cross-sectional distribution of the average R_{oob} statistic, where a positive statistic shows that the MOBA model outperforms the core HAR-RV model on average, indicating that the climate risks help to improve predictive accuracy. The white dot represents the median,

[4]We compute all our empirical results using the R language and environment for statistical computing (R Core Team 2021). All results that we report in this section are based on (random) partitions of the as follows: We use a train fraction of 50%, a validation fraction of 20%, and a test fraction of 30%. We report results for an alternative partition scheme at the end of the paper (Supplementary Material).

the thick black bar represents the interquartile range, and the thin black bar denotes ± 1.5 times the interquartile range. The shaded gray area represents the area under kernel density estimates of the cross-section of statistics and, thus, informs in detail about the distribution across states. The main message conveyed by the violin plots is that the cross-sectional distribution of the statistic shifts upward when we increase the length of the prediction horizon. The center of the distributions can be found above unity for the two long prediction horizons, while it settles close to but below unity for the two short prediction horizons.

– Figure 4 about here. –

We summarize the p-values for the R_{oob} statistic in Panel B of Figure 4. In line with the results we report in Panel A, we observe that the center of the cross-sectional distribution of the p-values clearly shifts downward when we increase the prediction horizon. For the long prediction horizons, the mass of the kernel densities settles below the two dashed horizontal lines, representing the 10% and 5% significance levels.

– Figure 5 about here. –

Figure 5, we go beyond our baseline scenario and compare an extended MOBA model with an extended HAR-RV model. In addition to the core HAR-RV predictors, both models feature the realized skewness, realized kurtosis, and realized jumps, and our measures of realized upside and downside tail risk as control variables. As in our baseline scenario, we find that the importance of climate risks for predictive accuracy strengthens in the cross-section of states when we increase the length of the predictive horizon.

– Figure 6 about here. –

We summarize the results (p-values) for good and bad realized volatility in Figure 6. This decomposition of realized volatility is motivated by the observation made by Giot et al. (2010) that market participants, in addition to the level of volatility, also care about the nature of volatility, with market participants typically differentiating between upside and downside volatilities. The results for good and bad realized state-level stock-market volatility are similar across good and bad realized volatility and lend further support to our main finding that the predictive value of the climate risks for state-level realized stock-market volatility strengthens in the cross-section of states when we increase the prediction horizon.

– Figure 7 about here. –

We plot the results for an asymmetric loss function in Figure 7, where we report results for a loss function of the quad-quad type in Panel A, and the results for a loss function of the lin-lin

type in Panel B. We present the results for different values of the shape parameter, a. The results show that the center of the cross-sectional distribution of the R^2_{oob} statistic shifts upward as we increase the prediction horizon for a forecaster who suffers a higher loss from an under-prediction of realized volatility than from an over-prediction of the same absolute size. In contrast, the center of the cross-sectional distribution shifts downward as the length of the forecast horizon increases for a forecaster who incurs a higher loss from an over-prediction than from a comparable under-prediction. Hence, the magnitude of the benefits of using climate risks for predicting state-level realized stock-market volatility clearly depends on the shape of a forecaster's loss function, where the results for the quad-quad loss function are stronger than those for the lin-lin loss function.

In Figure 8, we present results for a scenario in which the MOBA model features either only nation-wide or only state-level climate risks as additional predictors relative to the core HAR-RV model. In this scenario, the MOBA model features only seven additional climate predictors. Hence, it would be straightforward to estimate all of the possible $2^7 = 128$ forecasting models every period for all states and then rely on some model-selection or model-averaging criterion to form a forecast (for such a combinatorial approach, see, for example, Pesaran and Timmermann, 2000). To ensure that the results for this scenario are directly comparable to the results in Figure 4, however, we continue to use our MOBA algorithm for this simplified scenario.

– Figure 8 about here. –

The test results show that, in the cross section of states, both nation-wide and state-level climate risks contain predictive value for state-level realized stock-market volatility at the two long prediction horizons, where, in the cross section, the test results for the nation-wide climate risks are even somewhat stronger in terms of statistical significance than those for the state-level climate risks. The state-level climate risks, in turn, appear to work slightly better than the nation-wide climate risks for several states in case of the intermediate prediction horizon. In any case, the test results show that a forecaster should consider both nation-wide and state-level climate risks to predict state-level realized stock-market volatility at the two long forecast horizons. Given the long list of candidate preiction models with so many predictor variables, the MOBA algorithm is an efficient technique to extract predictive information from the nation-wide and state-level climate-risk predictors.

5 Concluding Remarks

Based on high-frequency US data for the period beginning in September, 2011 and ending in October, 2021, we have analyzed the predictive value of climate risks for state-level realized stock-market volatility at short and long prediction horizons. To this end, we have computed out-of-bag predic-

tions using a simple and efficient MOBA algorithm to extract the predictive information from the various nation-wide, and state-level climate risks predictors. Our main finding is that, in the cross section of states, climate risks have predictive value for state-level realized stock-market volatility at the two (one month and two months) forecast horizons that we have studied in our empirical research. This finding also obtains for good, and bad state-level realized stock-market volatility and when we use various realized moments of state-level stock-market returns as control variables. Moreover, we have shown that, in the cross section of states, a forecaster who suffers a larger loss from an under-prediction of state-level realized stock-market volatility tends to benefit to a larger extent from tracing climate risks than a forecaster who incurs a larger loss from an over-prediction of the same absolute magnitude. Our results, concerning the importance of local climate-related rare disasters for accurate prediction of state-level stock-market volatility, convey valuable information for investors, portfolio managers, and derivate traders, especially when one accounts for the fact that stock market players tend to overweigh local firms in their portfolios (Pham et al., 2021).

In terms of future research, it is interesting to combine our results with mixed-frequency models of the US economy to improve the accuracy of low-frequency predictions of state-level real-activity variables, which is of obvious importance for policymakers. In technical terms, it is interesting to compare the prediction performance of the MOBA algorithm that we have used in this research with the performance of other widely-studied techniques like, for example, the Lasso or random forests that have been used in earlier research in setting with many predictors.

References

Amaya, D., Christoffersen, P., Jacobs, K., and Vasquez, A. (2015). Does realized skewness predict the cross-section of equity returns? Journal of Financial Economics, 118, 135–167.

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.

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

Barndorff-Nielsen, O.E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics, 4, 1–30.

Barro, R.J. (2006). Rare disasters and asset markets in the twentieth century. Quarterly Journal of Economics, 121, 823–866.

Barro, R.J., and Ursúa, J.F. (2008). Macroeconomic Crises since 1870. Brooking Papers on Economic Activity, 39(1), 255–350.

Barro, R.J., and Ursúa, J.F. (2012). Rare Macroeconomic Disasters, Annual Review of Economics, 4(1), 83–109.

Battiston, S., Dafermos, Y., and Monasterolo, I. (2021). Climate risks and financial stability. Journal of Financial Stability, 54, 100867.

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.

Ben-Rephael, A., Da, Z., and Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. Review of Financial Studies, 30(9), 3009–3047.

van Benthem, A.A., Crooks, E., Giglio, S., Schwob, E., and Stroebel, J. (2022). The effect of climate risks on the interactions between financial markets and energy companies. Nature Energy, 7, 690–697.

Berkman, H., Jacobsen, B., and Lee, J.B. (2011). Time-varying rare disaster risk and stock returns, Journal of Financial Economics, 101, 313–332.

Berkman, H., Jacobsen, B., and Lee, J.B. (2017). Rare disaster risk and the expected equity risk premium, Accounting and Finance, 57(2), 351–372.

Bonato, M., Cepni, O., Gupta, R., and Pierdzioch, C. (2022). Climate risks and realized volatility of major commodity currency exchange rates. Journal of Financial Markets. DOI: https://doi.org/10.1016/j.finmar.2022.100760.

Bonato, M., Cepni, O., Gupta, R., and Pierdzioch, C. (Forthcoming). El Niño, La Niña, and forecastability of the realized variance of agricultural commodity prices: Evidence from a machine learning approach. Journal of Forecasting.

Bua, G., Kapp, D., Ramella, F., and Rognone, L. (2022). Transition versus physical climate risk pricing in European financial markets: A text-based approach. European Central Bank Working Paper No. 2022/2677.

Chaney, T., Sraer, D., and Thesmar, D. (2012). The collateral channel: how real estate shocks affect corporate investment. American Economic Review, 102(6), 2381−2409.

Choi, D., Gao, Z., and Jiang, W. (2020). Attention to global warming. Review of Financial Studies, 33 (3), 1112−1145.

Ćorić, B. (2021). Economic Disasters: A New Data Set. Finance Research Letters, 39(C), 101612.

Ćorić, B., and Šimić, V. (2021). Economic disasters and aggregate investment. Empirical Economics, 61(6) 3087−3124.

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

Coval, J.D., and Moskowitz, T.J. (1999). Home bias at home: local equity preference in domestic portfolios. Journal of Finance, 54, 2045−2073.

Coval, J.D., and Moskowitz, T.J. (2001). The geography of investment: Informed trading and asset prices. Journal of Political Economy, 199(4), 811−841.

Elliott, G., Komunjer, I., and Timmermann, A. (2005). Estimation and testing of forecasting rationality under flexible loss. Review of Economic Studies, 72: 1107−1125.

Engle, R.F., Giglio, S., Kelly, B., Lee, H., and Stroebel, J. (2020). Hedging climate change news. Review of Financial Studies, 33(3), 1184−1216.

Faccini, R., Matin, R., and Skiadopoulos, G. (2021). Dissecting climate risks: Are they reflected in stock prices? Available at SSRN: https://ssrn.com/abstract=3795964.

Flori, A., Pammolli, F., and Spelta, A. (2021). Commodity prices co-movements and financial stability: a multidimensional visibility nexus with climate conditions. Journal of Financial Stability, 54, 100876.

Giglio, S., Kelly, B., and Stroebel, J. (2021). Climate finance. Annual Review of Financial Economics, 13, 15–36.

Gil-Alana, L.A., Gupta, R., Sauci, L., and Carmona-Gonzalez, N. (Forthcoming). Temperature and precipitation in the US states: Long memory, persistence and time trend. Theoretical and Applied Climatology.

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

Gupta, R., Nel, J., and Pierdzioch, C. (2021). Investor confidence and forecastability of US stock market realized volatility: Evidence from machine learning. Journal of Behevaioral Finance. DOI: https://doi.org/10.1080/15427560.2021.1949719.

Hastie, T., Tibshirani, R., Friedman, J. (2009) The elements of statistical learning: Data mining, inference, and prediction, 2nd ed.: Springer: New York, NY, USA.

Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3, 1163-1173.

Kapfhammer, F., Larsen, V.H., and Thorsrud, L.A. (2020). Climate risk and commodity currencies. Centre for Applied Macro and Petroleum economics (CAMP), BI Norwegian Business School, Working Papers No. 10/2020.

Kim, H.S., Matthes, C., and Phan, T. (2021). Extreme weather and the macroeconomy. Federal reserve Bank of Richmond, Working Paper No. 21–14.

Korniotis, G.M., and Kumar, A. (2013). State-level business cycles and local return predictability. Jorunal of Finance, 68(3), 1037–1096.

Liu, R., Demirer, R., Gupta, R., and Wohar, M.E. (2020). Do bivariate multifractal models improve volatility forecasting in financial time series? An application to foreign exchange and stock markets. Journal of Forecasting, 39(2), 155–167.

Liu, R., and Gupta, R. (2022). Investors' uncertainty and forecasting stock market volatility. Journal of Behavioral Finance, 23(3), 327–337.

McAleer, M., and Medeiros, M.C. (2008). Realized volatility: A review. Econometric Reviews, 27, 10–45.

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.

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.

Pesaran, M.H., and Timmermann, A. (2000). A recursive modelling approach to predicting UK stock returns. Economic Journal, 110, 159−191.

Pham, A.V., Adrian, C., Garg, M., Phang, S-Y., and Truong, C. (2021). State-level COVID-19 outbreak and stock returns. Finance Research Letters, 43, 102002.

Pirinsky, C., and Wang, Q. (2006). Does corporate headquarters location matter for stock returns? Journal of Finance, 61, 1991−2015.

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 (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

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, 381−416.

Rietz, T. (1988). The equity risk premium: A solution. Journal of Monetary Economics, 22, 117−131.

Salisu, A.A., Demirer, R., and Gupta, R. (2022a). Financial turbulence, systemic risk and the predictability of stock market volatility. Global Finance Journal, 52(1), 100699.

Salisu, A.A., Gupta, R., and Ogbonna, A.E. (2022b). 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, 27(1), 384−400.

item Segnon, M., Gupta, R., and Wilfling, B. (Forthcoming). Forecasting stock market volatility with regime-switching GARCH-MIDAS: The role of geopolitical risks. International Journal of Forecasting.

Sheng, X., Gupta, R., and Çepni, O. (2022a). The effects of climate risks on economic activity in a panel of US states: The role of uncertainty. Economics Letters, 213, 110374.

Sheng, X., Gupta, R., and Çepni, O. (2022b). Persistence of state-level uncertainty of the United States: The role of climate risks. Economics Letters, 215, 110500.

Stroebel, J., and Wurgler, J. (2021). What do you think about climate finance? Journal of Financial Economics, 142, 487−498.

Tsai, J., and Wachter, J.A. (2015). Disaster risk and its implications for asset pricing. Annual Review of Financial Economics, 7, 219−252.

Wachter, J.A. (2013). Can time-varying risk of rare disasters explain aggregate stock market volatility?. Journal of Finance, 68(3), 987−1035.

Zhou, H., and Zhu, J.Q. (2012). An empirical examination of jump risk in asset pricing and volatility forecasting in China's equity and bond markets. Pacific-Basin Finance Journal, 20(5), 857−880.

Figure 1: State-level mean of realized volatility

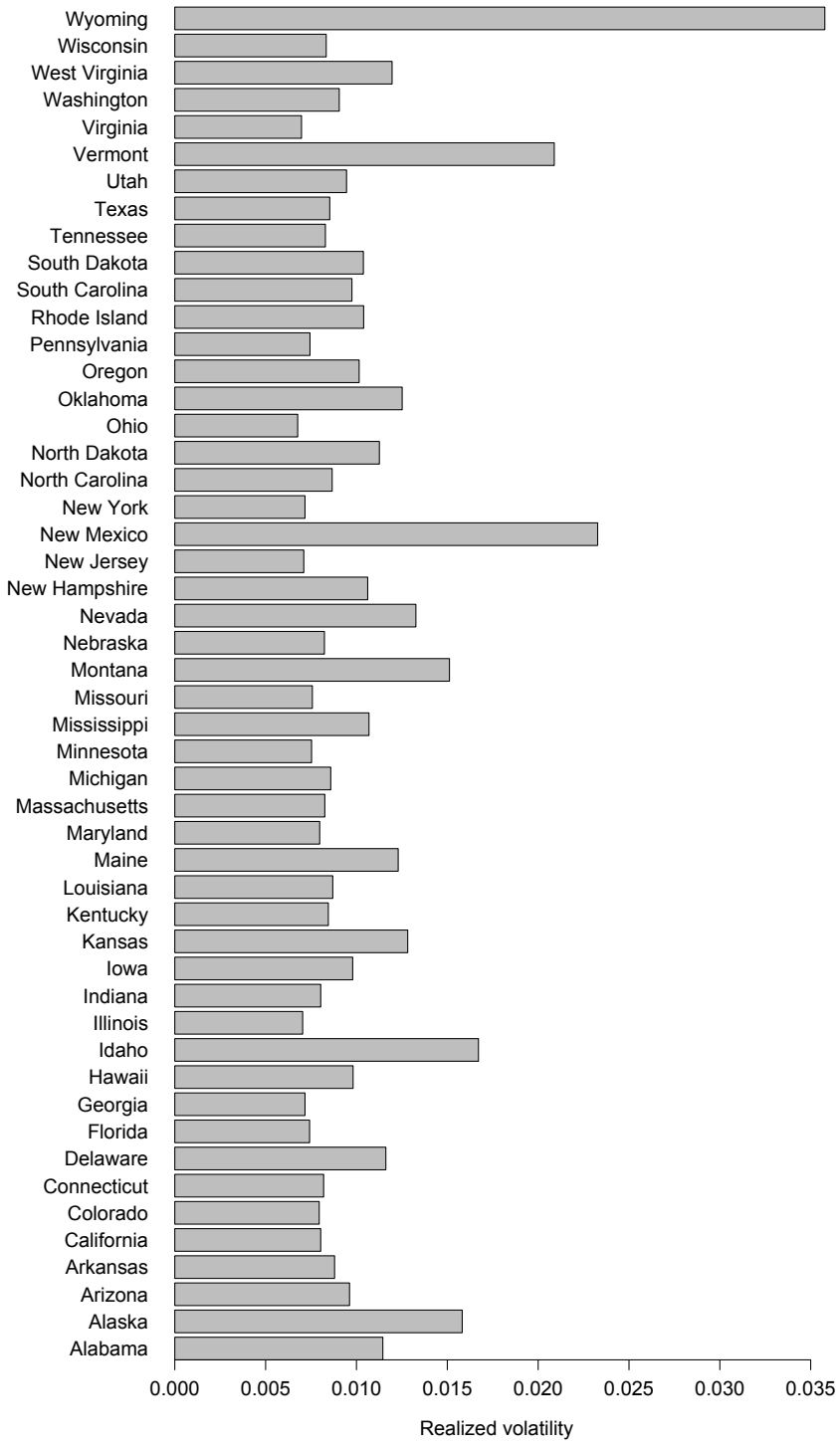
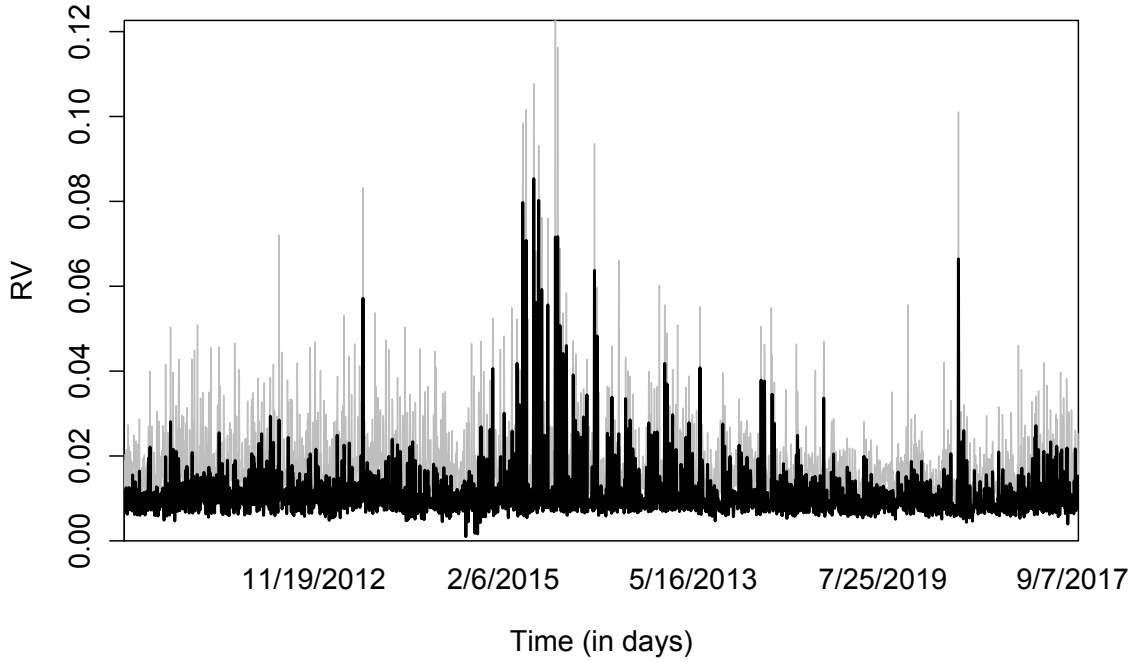
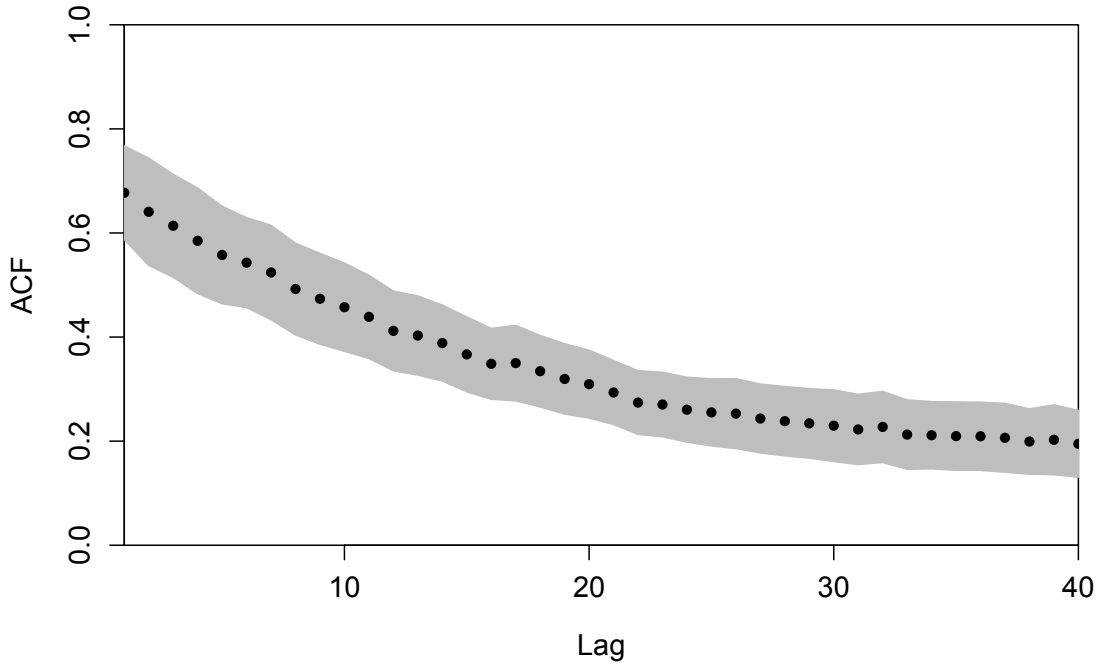


Figure 2: Cross-sectional mean and standard deviation of realized volatility



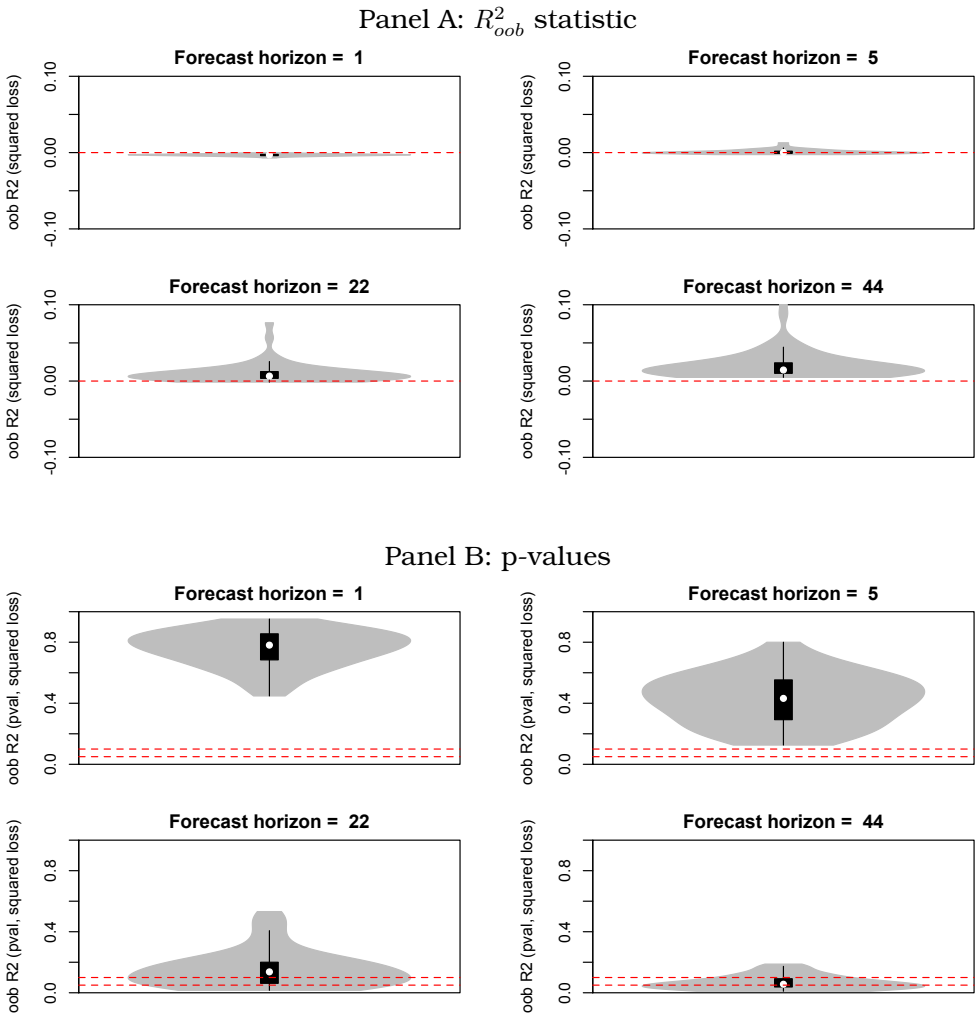
The black line denotes the cross-sectional mean of realized volatility, where we merged the data along the date column. The gray line is denotes the cross-sectional mean plus one cross-sectional standard deviation of realized volatility.

Figure 3: Cross-sectional mean of autocorrelation of realized volatility



The dots denote the cross-sectional mean of coefficient of autocorrelation of realized volatility, where we merged the data along the date column. The shaded gray area denotes the cross-sectional mean coefficient of autocorrelation plus/minus one standard deviation of the coefficient of autocorrelation autocorrelation of realized volatility.

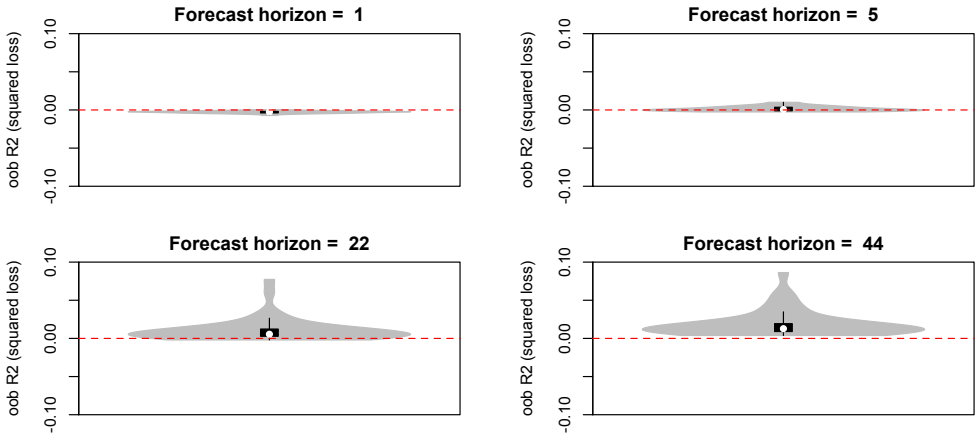
Figure 4: Baseline results



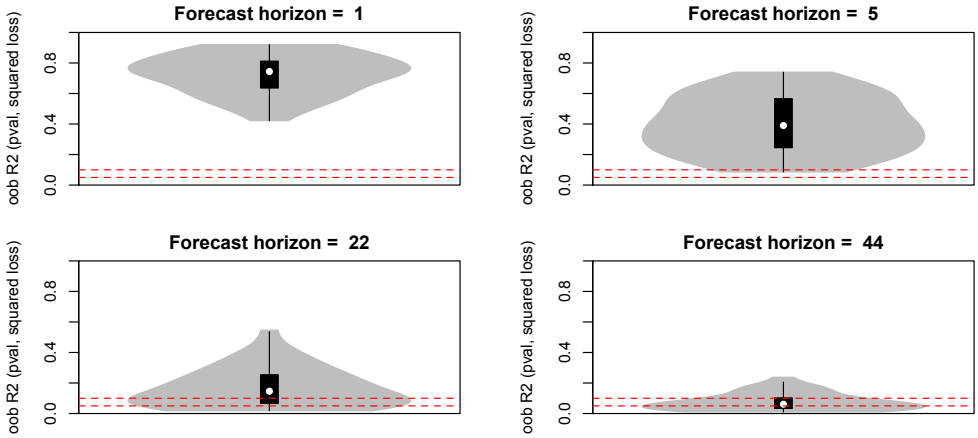
A positive R^2oob statistic indicates that the MOBA model outperforms the core HAR-RV model on predicting the test data (on average across all simulation runs). The white dot represents the median, the thick black bar denotes the interquartile range, and the thin black bar denotes ± 1.5 times the interquartile range. The boundaries of the shaded area represent kernel density estimates estimated on the R^2oob statistics for the states. The dashed horizontal lines (Panel B) denote the 10% and 5% significance levels.

Figure 5: The HAR-RV model includes realized moments

Panel A: Core HAR-RV model is the benchmark model

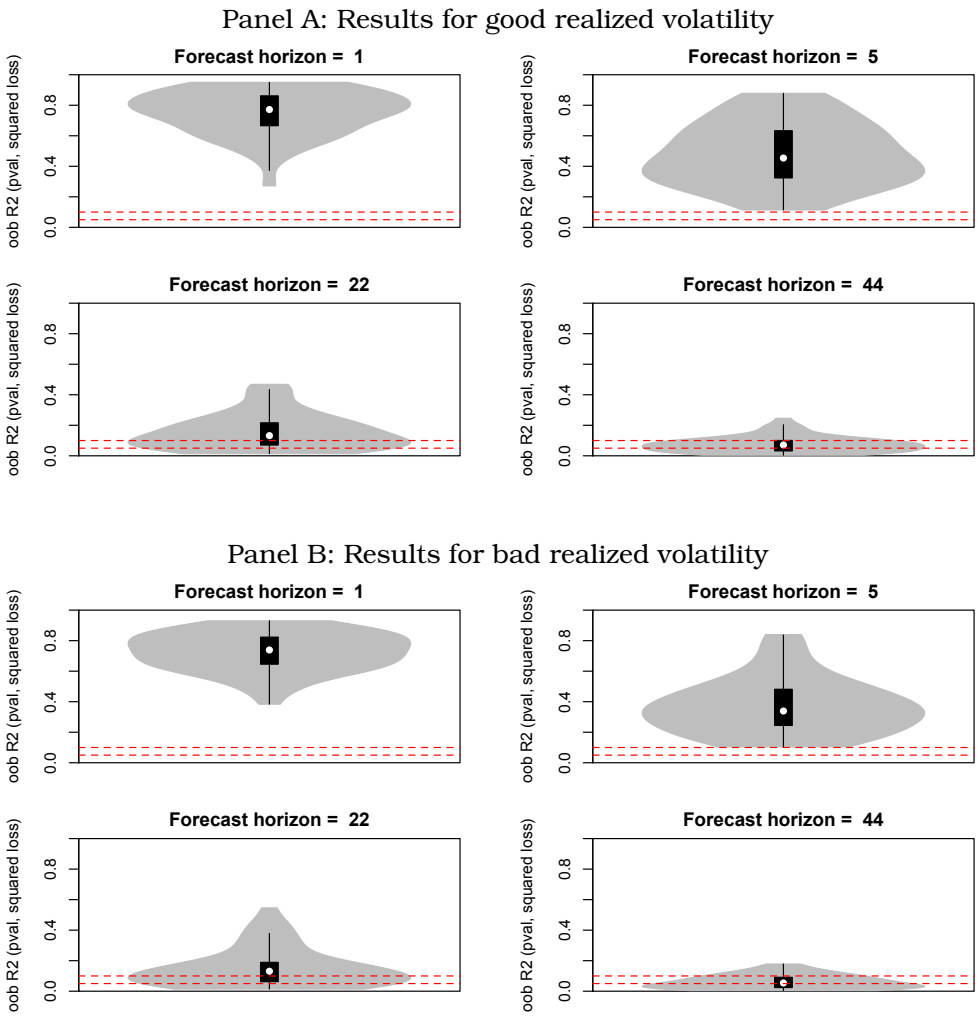


Panel B: HAR-RV model featuring realized moments is the benchmark model



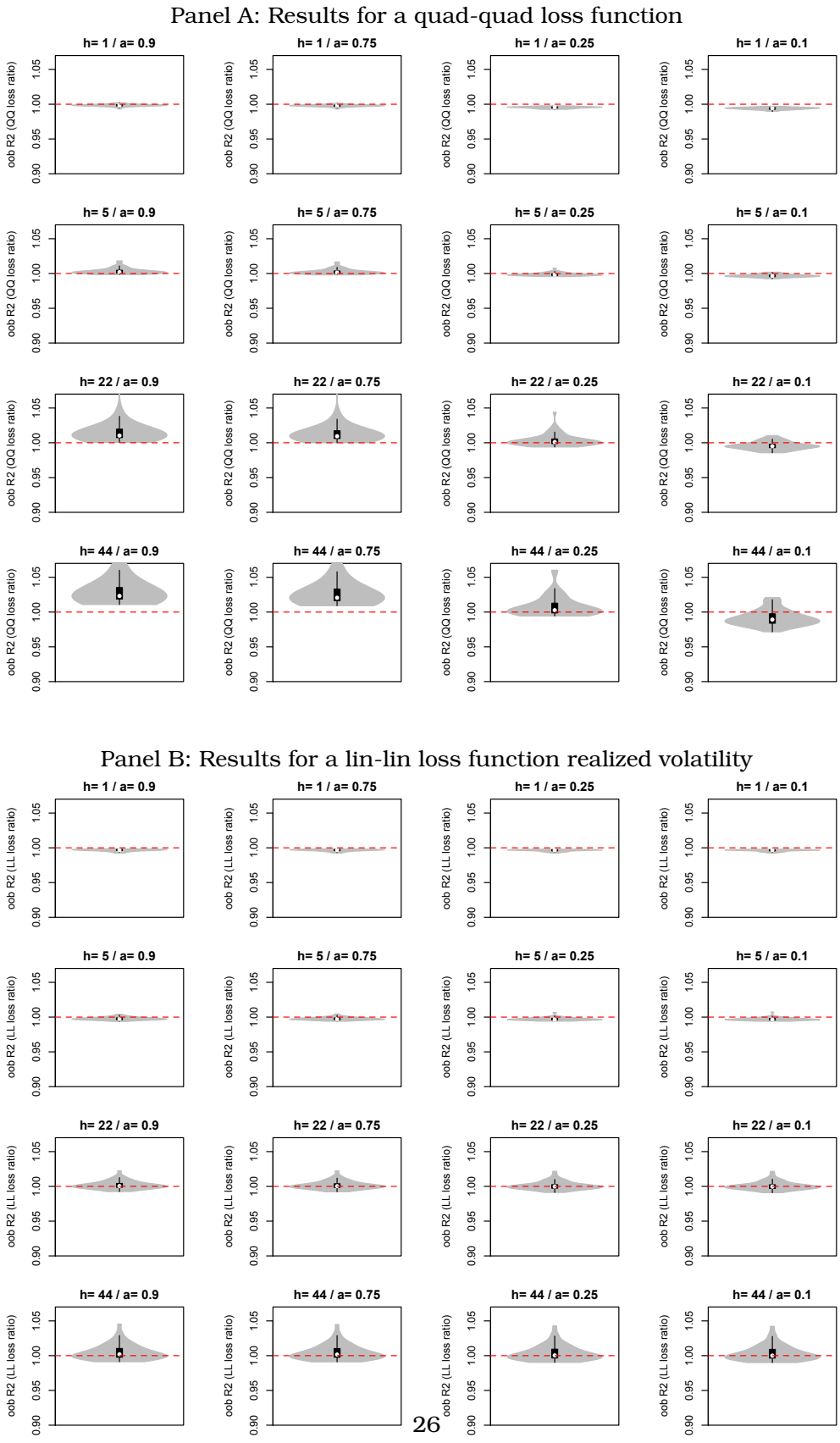
A positive R^2oob statistic indicates that the MOBA model outperforms the core HAR-RV moments-model on predicting the test data (on average across all simulation runs). The realized moments are: realized skewness, realized kurtosis, realized jumps, realized upside tail risk, realized downside tail risk. The white dot represents the median, the thick black bar denotes the interquartile range, and the thin black bar denotes ± 1.5 times the interquartile range. The boundaries of the shaded area represent kernel density estimates estimated on the R^2oob statistics for the states. The dashed horizontal lines (Panel B) denote the 10% and 5% significance levels.

Figure 6: Results for realized good and bad volatility



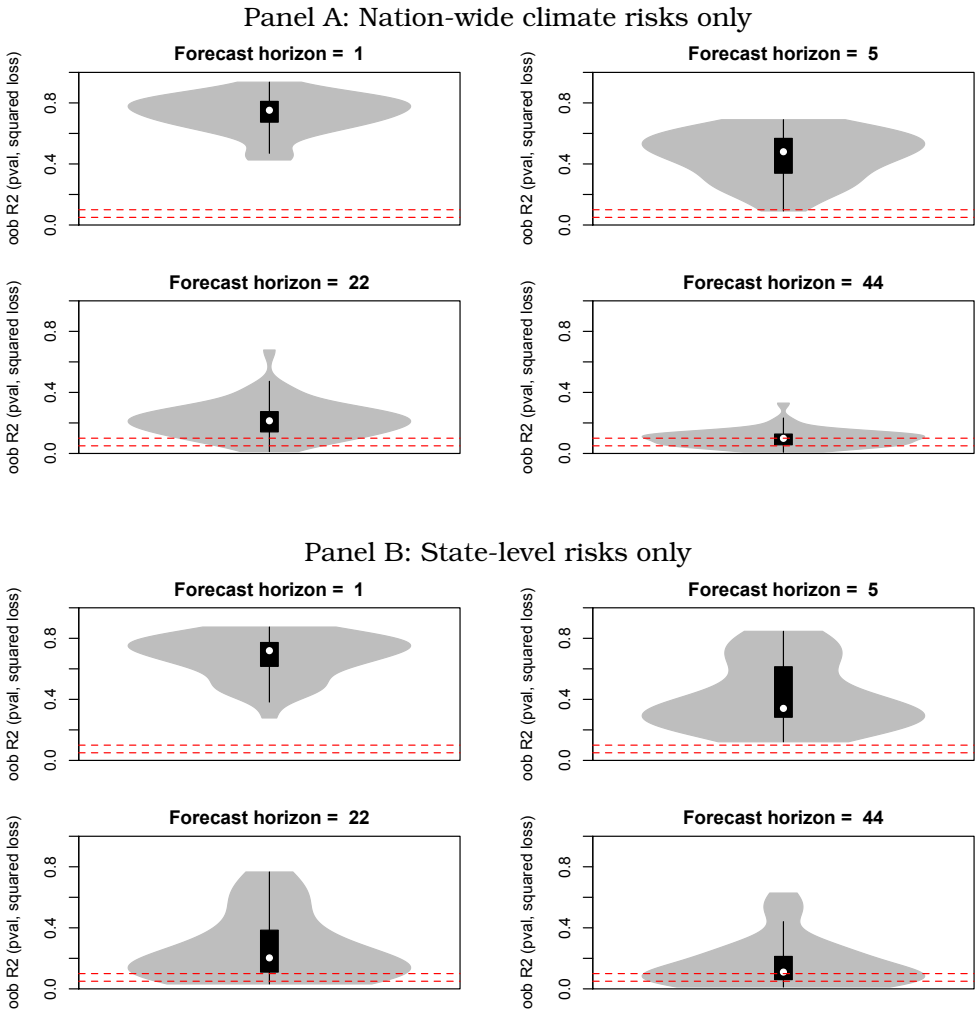
p-values for the R^2oob statistic. A positive R^2oob statistic indicates that the MOBA model outperforms the core HAR-RV model on predicting the test data (on average across all simulation runs). The white dot represents the median, the thick black bar denotes the interquartile range, and the thin black bar denotes ± 1.5 times the interquartile range. The boundaries of the shaded area represent kernel density estimates estimated on the R^2oob statistics for the states. The dashed horizontal lines denote the 10% and 5% significance levels.

Figure 7: Results for an asymmetric loss function



Results for an asymmetric loss function. For $a > 0.5$, a forecaster suffers a larger loss from an under-prediction of realized volatility (as compared to an over-prediction of the same absolute size), and for $a < 0.5$ an over-prediction gives rise a larger loss than a comparable under-prediction. A positive r^2_{oob} statistic implies that the MOBA model performs better than the core HAR-RV model model. The white dot represents the median, the thick black bar denotes the interquartile range, and the thin black bar denotes ± 1.5 times the interquartile range. The boundaries of the shaded area represent kernel density estimates estimated on the test results.

Figure 8: Results for nation-wide and state-level climate risks



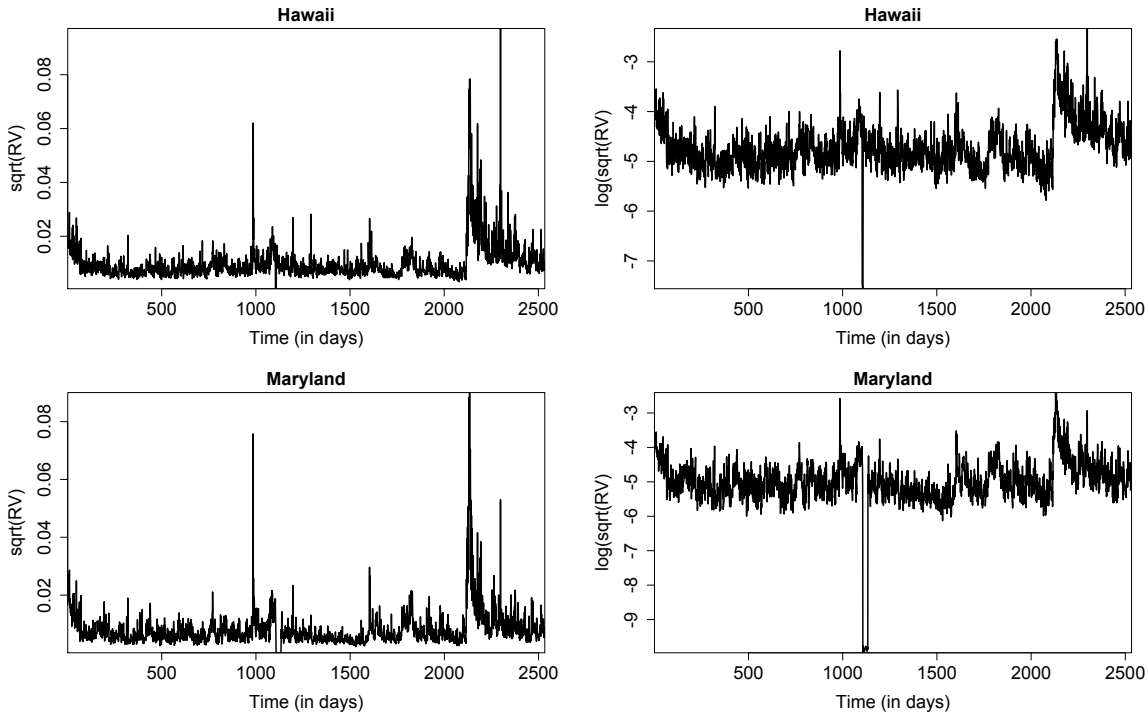
p-values for the R^2oob statistic. A positive R^2oob statistic indicates that the MOBA model outperforms the core HAR-RV model on predicting the test data (on average across all simulation runs). The white dot represents the median, the thick black bar denotes the interquartile range, and the thin black bar denotes ± 1.5 times the interquartile range. The boundaries of the shaded area represent kernel density estimates estimated on the R^2oob statistics for the states. The dashed horizontal lines denote the 10% and 5% significance levels.

Supplementary Material

As briefly discussed in Footnote 2 in the main body of the text, the raw data obtained from Bloomberg contain periods, usually a very small fraction of the data, during which the state-level stock-market indices follow a more or less a straight trend. Such peculiar price dynamics probably indicate that Bloomberg stopped producing the indices during those "gap periods" and interpolated the resulting gap in the data in retrospect. Given this pattern in the data, it is not surprising that realized volatility is unusually low during such a gap period.

In order to provide an example of how the gap periods look like, Figure S1 plots the realized volatilities and the natural logarithm of realized volatilities for two states, Hawaii and Maryland. The gap periods are hardly discernible in Figure S1 in the plots for realized volatilities. In fact, the gap periods become only visible in the plots for the log realized volatilities, where the unusually small realized volatilities during the gap periods translate into large negative realizations of the logarithm of realized volatility. It should be noted that, while the analysis of the logarithm of realized volatility, for example, it brings the data closer to normality, Figure S1 clearly illustrates why it is advantageous, in the case of our state-level data to study realized volatility, rather than its logarithm.

Figure S1: Examples of the data



We report some state-level descriptive statistics of the raw (high-frequency) data in Table S1. Specifically, we report the beginning and end dates for every state along with the total number of observations of high-frequency data that we used to compute the time series of state-level realized volatilities. Table S2, in turn, depicts the fraction of the gap periods as percent of the total daily observations per state. The gap periods are identified by visually inspection of the data as periods during which the log of realized volatility takes on a value smaller than -6 (-6.5 for Arizona). As we mentioned in Footnote 2, we remove the gap periods from the data for estimation of the prediction models.

Table S1: Descriptive statistics of the raw data

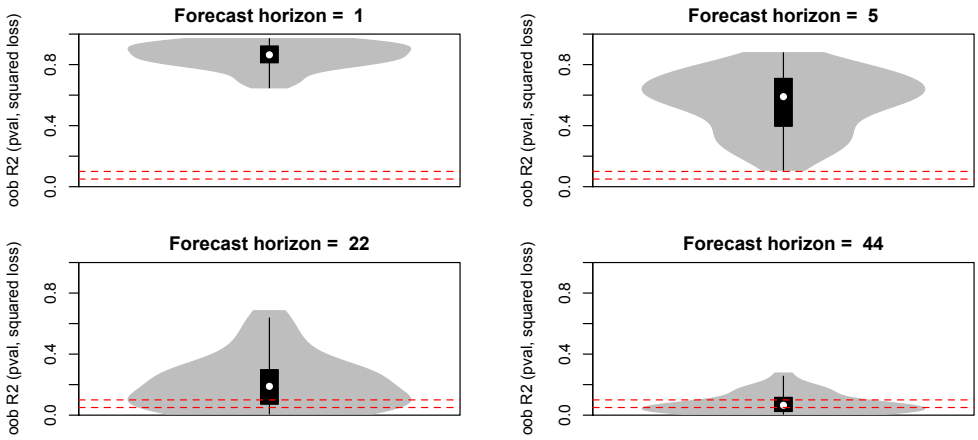
State Name	Start Date	End Date	Observations
Alabama	14/9/2011	22/10/2021	216432
Alaska	09/9/2011	22/10/2021	216676
Arizona	13/9/2011	22/10/2021	216434
Arkansas	14/9/2011	22/10/2021	216433
California	13/9/2011	22/10/2021	216519
Colorado	13/9/2011	22/10/2021	216120
Connecticut	15/9/2011	22/10/2021	215984
Delaware	15/9/2011	22/10/2021	215984
Florida	14/9/2011	22/10/2021	216280
Georgia	14/9/2011	22/10/2021	216215
Hawaii	15/9/2011	22/10/2021	216108
Idaho	13/9/2011	22/10/2021	216298
Illinois	14/9/2011	22/10/2021	216131
Indiana	15/9/2011	22/10/2021	216040
Iowa	14/9/2011	22/10/2021	216299
Kansas	13/9/2011	22/10/2021	216415
Kentucky	15/9/2011	22/10/2021	216210
Louisiana	14/9/2011	22/10/2021	216295
Maine	15/9/2011	22/10/2021	216125
Maryland	15/9/2011	22/10/2021	216033
Massachusetts	15/9/2011	22/10/2021	216106
Michigan	14/9/2011	22/10/2021	216130
Minnesota	13/9/2011	22/10/2021	216466
Mississippi	14/9/2011	22/10/2021	216040
Missouri	14/9/2011	22/10/2021	216201
Montana	13/9/2011	22/10/2021	216405
Nebraska	13/9/2011	22/10/2021	216347
Nevada	13/9/2011	22/10/2021	216337
New Hampshire	15/9/2011	22/10/2021	216033
New Jersey	21/9/2011	22/10/2021	215706
New Mexico	13/9/2011	22/10/2021	216233
New York	19/8/2011	22/10/2021	217318
North Carolina	14/9/2011	22/10/2021	216092
North Dakota	13/9/2011	22/10/2021	216139
Ohio	15/9/2011	22/10/2021	216264
Oklahoma	14/9/2011	22/10/2021	216350
Oregon	13/9/2011	22/10/2021	216179
Pennsylvania	14/9/2011	22/10/2021	216118
Rhode Island	15/9/2011	22/10/2021	216323
South Carolina	15/9/2011	22/10/2021	216340
South Dakota	13/9/2011	22/10/2021	216480
Tennessee	14/9/2011	22/10/2021	216344
Texas	09/9/2011	22/10/2021	216604
Utah	13/9/2011	22/10/2021	216516
Vermont	15/9/2011	22/10/2021	216041
Virginia	15/9/2011	22/10/2021	216348
Washington	14/9/2011	22/10/2021	216085
West Virginia	15/9/2011	22/10/2021	216254
Wisconsin	14/9/2011	22/10/2021	216349
Wyoming	13/9/2011	22/10/2021	216308

Table S2: Fraction of gap periods (in days) as percent of total data

State	Fraction (in %)
Alabama	0.79
Alaska	0.35
Arizona	0.71
Arkansas	0.87
California	1.10
Colorado	0.87
Connecticut	0.95
Delaware	1.15
Florida	1.10
Georgia	0.24
Hawaii	0.20
Idaho	0.16
Illinois	0.47
Indiana	1.11
Iowa	1.07
Kansas	0.99
Kentucky	1.10
Louisiana	1.10
Maine	1.11
Maryland	1.07
Massachusetts	0.99
Michigan	1.22
Minnesota	1.07
Mississippi	0.00
Missouri	1.11
Montana	0.91
Nebraska	1.10
Nevada	0.67
New Hampshire	0.51
New Jersey	0.39
New Mexico	0.39
New York	0.04
North Carolina	0.95
North Dakota	1.93
Ohio	1.50
Oklahoma	2.21
Oregon	1.11
Pennsylvania	2.29
Rhode Island	2.25
South Carolina	1.10
South Dakota	0.99
Tennessee	2.21
Texas	2.17
Utah	1.14
Vermont	2.29
Virginia	1.85
Washington	2.29
West Virginia	7.81
Wisconsin	1.70
Wyoming	0.91

Figure S2 depicts the results for an alternative partition of the data into training, validation, and test samples than that we use in the main body of the text. Specifically, we use a train fraction of 30%, a validation fraction of 20%, and a test fraction of 50%.

Figure S2: Results for an alternative partitioning scheme



p-values for the R^2oob statistic. A positive R^2oob statistic indicates that the MOBA model outperforms the core HAR-RV model on predicting the test data (on average across all simulation runs). The white dot represents the median, the thick black bar denotes the interquartile range, and the thin black bar denotes ± 1.5 times the interquartile range. The boundaries of the shaded area represent kernel density estimates estimated on the R^2oob statistics for the states. The dashed horizontal lines denote the 10% and 5% significance levels.