Journal Pre-proof

The information content of short-term options

Ioannis Oikonomou, Andrei Stancu, Lazaros Symeonidis, Chardin Wese Simen

PII: S1386-4181(18)30305-7

DOI: https://doi.org/10.1016/j.finmar.2019.07.003

Reference: FINMAR 504

To appear in: *Journal of Financial Markets*

Received Date: 18 October 2018

Revised Date: 27 July 2019

Accepted Date: 29 July 2019

The information content of short-term options*

Ioannis Oikonomou†, Andrei Stancu‡, Lazaros Symeonidis§, and Chardin Wese Simen¶

Abstract

We exploit weekly options on the S&P 500 index to compute the weekly implied variance. We show that the weekly implied variance is a strong predictor of the weekly realized variance. In an encompassing regression test, it crowds out the information content of the monthly implied variance. Further tests reveal that the weekly implied variance outperforms not only the monthly implied variance but also well-established time series models of realized variance. This result holds both in- and out-of-sample and the forecast accuracy gains are significant.

JEL classification: G11, G12

Keywords: Implied variance, Predictability, Realized variance, Weekly options

*We are grateful to an anonymous reviewer, Torben Andersen, Fabian Hollstein, Ming-Tsung Lin (discussant), Marcel Prokopczuk, Rogier Quaedvlieg, Viktor Todorov, and participants at the International Association for Applied Econometrics Conference (2019), the 4th Annual Conference of Brunel Studies in Economics & Finance (2019), the American Economic Association Meeting (2018), the Financial Econometrics Conference at the University of Lancaster (2018), the British Accounting and Finance Association Conference (2018), the Western Economic Association International Conference (2018), the Econometrics and Financial Data Science Workshop (2017), the EFiC Conference in Banking and Finance (2017), the International Conference on Computational and Financial Econometrics (2017), Leibniz University Hannover, Norwich Business School, the University of Glasgow, the University of Kent, and the University of Stirling for helpful comments and discussions.

†ICMA Centre, Henley Business School, University of Reading, Reading, RG6 6BA, United Kingdom. Contact: i.oikonomou@icmacentre.ac.uk.

‡Norwich Business School, University of East Anglia, Norwich, NR4 7TJ, United Kingdom. Contact: a.stancu@uea.ac.uk

§Norwich Business School, University of East Anglia, Norwich, NR4 7TJ, United Kingdom. Contact: l.symeonidis@uea.ac.uk (corresponding author)

¶ICMA Centre, Henley Business School, University of Reading, Reading, RG6 6BA, United Kingdom. Contact: c.wese-simen@icmacentre.ac.uk.

1 Introduction

Several studies document that the variance implied by monthly options is a strong predictor of the monthly realized variance (e.g., Jiang and Tian, 2005; Carr and Wu, 2009; Busch et al., 2011).[1] While this fact is generally well-accepted, we know relatively little about the forecasting power of short-term implied variance for short-term realized variance. Although it is tempting to speculate that the same findings will hold for the shorter horizon, there are a couple of reasons to suggest otherwise. First, short-term options are routinely discarded in the literature on the grounds that they are illiquid and noisy. Consequently, their information content is assumed to be limited. Second, for the sample period of most studies there were very few short-term options available in the market, making it challenging to obtain a long enough time series of implied variance (e.g., Jiang and Tian, 2005; Busch et al., 2011). This limitation is important because the statistical tests may lack power in a short sample. As a work around this issue, some studies use the monthly implied variance to predict the short-term realized variance (e.g., Blair et al., 2001; Pong et al., 2004; Kourtis et al., 2016). Alas, this approach introduces a mismatch between the maturity of the options and the forecasting horizon. It is unclear how big an issue this disconnect may be.

In this paper, we exploit weekly option contracts (weeklies), which were launched by the Chicago Board of Options Exchange (CBOE) in 2005, to study the predictability of weekly realized variance. Essentially, these option contracts allow market participants to better manage their short-term risk (e.g., the weekly realized variance). Andersen et al. (2017) document that weeklies account for nearly 50 % of the total trading volume in the S&P 500 index options in 2015, indicating that these options are quite liquid. We use

[1]See also the survey by Poon and Granger (2003).

the Bakshi et al. (2003) estimator to compute the weekly option–implied variance and analyze its predictive power for the weekly realized variance estimated using 5-minute sampled data on the S&P 500 index. For our main analysis, we consider the time period spanning from March 5, 2008 to August 31, 2015.

In a regression of the daily time series of the weekly realized variance on a constant and the lagged weekly implied variance, we obtain a statistically significant slope estimate and a high predictive power (Adj R^2=64.9%). In an effort to understand the channel through which this predictability arises, we decompose the realized variance into its continuous and jump components (Barndorff-Nielsen, 2002). We find that the weekly implied variance predicts both components. Next, we evaluate the incremental information content of the monthly implied variance relative to that of the weekly implied variance. To this end, we formulate and estimate an encompassing model. The regression results suggest that the weekly implied variance crowds out the orthogonalized monthly implied variance. Furthermore, the forecasting performance of the weekly implied variance is significantly better than that of time series models of the heterogeneous autoregressive (HAR) realized variance family. These results hold both in- and out-of-sample.

We run a battery of tests to assess the robustness of our results. The findings are unaffected when using a sampling frequency of 1-minute to estimate the realized variance. Our conclusions are also robust to the method of interpolation used to compute the implied variance series. Furthermore, the key findings are not driven by the choice of the implied variance estimator. Indeed, we obtain qualitatively similar results when using the Britten-Jones and Neuberger (2000) implied variance estimator. Finally, our findings are qualitatively similar if we extend the sample period back to 1996 and adopt a non-overlapping monthly sampling scheme.

Our research relates to the stream of the literature in which the implied variance of a given maturity is used to predict the realized variance of a shorter horizon. Blair et al. (2001), Pong et al. (2004), and Kourtis et al. (2016) are relevant studies along those lines. We confirm their findings that the monthly implied variance predicts the weekly realized variance. However, we find that there are significant gains in forecasting accuracy once the maturity of the option contracts and the forecasting horizon are aligned. In fact, our encompassing regression estimates suggest that the weekly implied variance crowds out the monthly implied variance. To the best of our knowledge, we are the first to document this result at the short horizon.

Our study also adds to the broader literature on the predictability of realized variance. Corsi (2009) proposes the HAR model and documents its superior performance relative to the random walk model. Andersen et al. (2007) decompose the historical variance terms of the HAR model into continuous and jump components. Patton and Sheppard (2015) propose an extension that separately uses positive and negative semivariances. Bollerslev et al. (2016) extend the HAR model to account for heteroskedastic measurement errors in realized variance. We show that the weekly implied variance provides significantly more accurate forecasts of short-term risk than the HAR model and its aforementioned extensions.

The remainder of this paper proceeds as follows. In Section 2, we describe the methodology and the dataset. In Section 3, we discuss the performance of the weekly implied variance relative to that of the monthly implied variance. In Section 4, we extend the analysis to various time series models of the HAR family. The robustness checks are presented in Section 5. We conclude in Section 6.

2 Methodology and data

In this section, we introduce the methodology used to construct the main variables. We also discuss the dataset.

2.1 Methodology

2.1.1 Realized variance

We focus on the predictability of next week's realized variance. We start with the definition of the intraday return:

$$r_{j,k} = \log\left(\frac{S_{j,k}}{S_{j,k-1}}\right), \tag{1}$$

where $r_{j,k}$ denotes the intraday return at the end of the k^{th} intraday interval of day j. $S_{j,k}$ and $S_{j,k-1}$ are the asset prices at the end of the k^{th} and $(k-1)^{th}$ intraday interval of trading day j, respectively.

We compute the (annualized) weekly realized variance as follows:

$$RV^w_{t+7} = 52 \times \sum_{j=0}^{N^w_{t+7}-1} \sum_{k=1}^{m} r^2_{t+7-j,k}, \tag{2}$$

where RV^w_{t+7} is the (annualized) weekly realized variance for the week ending on day $t+7$. The number 52 indicates that the realized variance estimate is annualized. N^w_{t+7} denotes the number of trading days during the week ending on day $t+7$. There are m returns observed on each trading day. The case where $k=1$ corresponds to the overnight return.[2]

[2]It is standard in the literature to account for the overnight returns. See Bollerslev et al. (2009), Drechsler and Yaron (2011), and Bekaert and Hoerova (2014), among others.

2.1.2 Implied variance

In the literature, the Britten-Jones and Neuberger (2000) estimator is often used to proxy for the risk-neutral expectation of the total variation of returns. However, this estimator captures the risk-neutral expectation of the continuous variation that is equal to the total variation of returns only if the underlying return process does not jump. This result arises because the total variation is the sum of the continuous and jump variations. In the presence of jumps, the Britten-Jones and Neuberger (2000) estimator is a biased estimator of the risk-neutral expectation of the future total variation and the magnitude of this bias increases with the contribution of jumps to the total variation of returns (Du and Kapadia, 2013).

Andersen et al. (2007) and Lee and Mykland (2008), among others, use non-parametric statistical tests to show that the S&P 500 index jumps.[3] Du and Kapadia (2013) conduct an extensive simulation exercise to assess the impact of jumps on the implied variance and recommend the Bakshi et al. (2003) estimator as a jump-robust estimator of implied variance. Heeding on their recommendation, we use the Bakshi et al. (2003) formula to compute the implied variance:[4]

$$IV_t^{\tau} = \frac{360}{\tau}\left[e^{rf_t \frac{\tau}{360}} QUAD_t - \left(e^{rf_t \frac{\tau}{360}} - 1 - \frac{e^{rf_t \frac{\tau}{360}}}{2} QUAD_t\right)^2\right], \tag{3}$$

where

$$QUAD_t = \int_0^{S_t} \frac{2\left(1+\ln\left[\frac{S_t}{K}\right]\right)}{K^2} P_t(\tau, K)dK + \int_{S_t}^{+\infty} \frac{2\left(1-\ln\left[\frac{K}{S_t}\right]\right)}{K^2} C_t(\tau, K)dK, \tag{4}$$

[3] The documented jumps are not spuriously induced by the fact that the index is not directly tradable. Prokopczuk and Wese Simen (2016) show that the liquid S&P 500 E-Mini futures contract, which is tradable, also jumps.

[4] To make our analysis more comparable to prior studies, we also consider the Britten-Jones and Neuberger (2000) implied variance. The results are in Subsection 5.3. Note, however, that these findings should be interpreted cautiously, keeping in mind that this specific estimator is not robust to jumps in the underlying return process.

IV_t^τ is the (annualized) implied variance of time to maturity τ (expressed in days) observed on day t. Throughout this paper, we use the expressions "weekly" and "monthly" to denote the case where $\tau = 7$ and $\tau = 30$ calendar days, respectively. rf_t is the τ-day (annualized) discount rate on day t. S_t is the underlying price on day t. $P_t(\tau, K)$ and $C_t(\tau, K)$ denote the price on day t of the European put and call options of time to maturity τ and strike price K, respectively. Note that the formula in equation (4) involves only out-of-the-money (OTM) options. For each option maturity available on that day, we compute the Black and Scholes (1973) implied volatility for all OTM options. We then average the OTM implied volatility estimates of the same maturity. Equipped with this average implied volatility, denoted σ, we define the variables $K_{t,L}$ and $K_{t,U}$ as follows:

$$K_{t,L} = S_t e^{-8\sigma_t} \tag{5}$$

$$K_{t,U} = S_t e^{8\sigma_t}, \tag{6}$$

where σ_t is the average implied volatility at time t of all OTM options of the same maturity.

Similar to Carr and Wu (2009), we linearly interpolate the implied volatilities for 2,000 equally–spaced strike prices between $K_{t,L}$ and $K_{t,U}$ defined in equations (5) and (6), respectively. In practice, the strike prices traded in the market do not completely span the interval starting at $K_{t,L}$ and ending at $K_{t,U}$, raising the question of extrapolation. We follow Jiang and Tian (2005) and Carr and Wu (2009), among others, and perform the nearest neighbourhood extrapolation. To be precise, for strike prices greater (lower) than $K_{t,L}$ ($K_{t,U}$) but lower (higher) than the lowest (highest) strike available in the market, we use the implied volatility associated with the lowest (highest) strike available in the market. Next, we map the grid of 2,000 implied volatilities into Black and Scholes (1973) OTM option prices. Finally, we use the trapezoidal rule to numerically evaluate the integrals in equation (4) and compute the implied variance as in equation (3).[5]

We repeat the steps above for all maturities observed on that day to obtain the term structure of implied variances. From this term structure, we linearly interpolate the implied variance of weekly (IV^w) and monthly (IV^m) horizons. We emphasize that we only interpolate between maturities and never extrapolate since this could introduce spurious spikes in the constant maturity implied variance series.[6]

2.2 Data

We obtain high-frequency data on the S&P 500 index from Thomson Reuters Tick History (TRTH) to build the realized variance series. Our interest in high-frequency data, as opposed to daily data, is motivated by the studies of Andersen and Bollerslev (1998), Barndorff-Nielsen and Shephard (2002), and Andersen et al. (2003), who recommend the use of intraday data to efficiently measure realized variance. The dataset spans the period from January 1996 to August 2015.[7] It contains bid and ask quotes pertaining to regular business hours, i.e., from 8:30 AM to 3:00 PM (Chicago Time). Similar to Bollerslev et al. (2009) and Bollerslev and Todorov (2011), among others, we use a 5-minute sampling frequency.[8] At the end of each 5-minute interval, we use the most recent mid-quote price

[5]Note that by using options with strike prices ranging from $K_{t,L}$ to $K_{t,U}$, we essentially truncate the integrals in equation (4). This choice is consistent with earlier work (e.g., Carr and Wu, 2009). In a simulation setting, Jiang and Tian (2005) show that the truncation error is negligible if the truncation points are more than two standard deviations from the current underlying price.

[6]As a robustness check, we do not linearly interpolate the constant maturity contracts but instead simply use the option contracts with maturity closest to the target maturity. On average, the actual maturities of the options are 6 and 27 calendar days for the weekly and monthly horizons, respectively. Our main findings are unchanged, leading us to conclude that the method of interpolation plays a minimal role in our results. See Subsection 5.2 for further details.

[7]The sample period is dictated by the dataset available from TRTH at the time we started this project.

[8]As a robustness check, we consider a higher sampling frequency of 1-minute and obtain qualitatively similar findings. See Subsection 5.1 for further details.

to proxy for the closing price of that interval.

To estimate our implied variance series, we obtain end-of-day S&P 500 index options data for the same period from IvyDB OptionMetrics. These options trade on the CBOE and are of the European type. For each trading day and option contract, the database contains information about the bid and ask prices, the open interest, the strike price, and the expiration date. The dataset includes weekly and standard option contracts, among others. Generally, weekly options expire on Friday of each week, except the third Friday of each month when the standard options expire.[9] Figure 3 of Andersen et al. (2017) reveals a rapid growth in the trading volume of the weekly contracts from less than 5 % of the total S&P 500 index option volume during the first few years of trading to 50 % towards the end of our sample period (i.e. in 2015).[10]

Although the OptionMetrics dataset spans the same time period as the TRTH dataset, we face two limitations that require us to start our sample in March 2008. The first limitation is due to the fact that weekly options on the S&P 500 index were initiated in October 2005. As a result, we can only analyze the period beginning from that point onwards. The second limitation is driven by the way OptionMetrics reports the closing option and underlying prices. Prior to March 5, 2008, OptionMetrics records derivatives prices at 3:15 PM (the market closing), whereas the underlying spot price is recorded at 3:00 PM, introducing a bias in studies that rely on synchronous observations of the

[9] At the time we started the project, the term structure of weekly options included up to 12 maturities. For further information about weekly options, we refer the interested reader to the following webpage: *http://www.cboe.com/micro/weeklys/introduction.aspx*. For an up-to-date list of weekly option contracts on offer, we direct the reader to the following link: *http://www.cboe.com/micro/weeklys/availableweeklys.aspx*.

[10] Absent official data on the identity/profile of market participants who trade weekly options, it is difficult to definitely ascertain their trading motives. Andersen et al. (2017) do not find a significant change in the trading activity of these contracts around important macroeconomic news announcements. This finding suggests that speculation does not seem to be the main driver of their trading activity. This leaves open the possibility that the increased trading activity in weeklies is primarily driven by a desire to improve short-term risk management.

derivatives and underlying prices. To mitigate this issue, we focus on the March 5, 2008 to August 31, 2015 sample period.[11,12]

We process the options data as follows. We discard observations with missing or zero prices. We implement this filter on bid and ask prices separately. In doing so, we address the concern that our dataset includes contracts that are not actively quoted. As is standard in the literature (e.g., Carr and Wu, 2009), we compute the mid-quote price of the option, which we refer to as the option price. Next, we remove all option observations that are in-the-money. We take this step because the computation of the implied variance only involves OTM option prices (see equation (4)). Furthermore, we download the discount rates from OptionMetrics. These discount rates are based upon the London Interbank Offered Rate (LIBOR) and the Eurodollar futures. For each trading day and option contract, we linearly interpolate the discount rate of the same time to maturity as the option contract. We then match the discount rates with the panel of options data. We also match the time series of the daily S&P 500 index prices and that of the dividend yield, both obtained from OptionMetrics, with the panel of options data.

Our analysis involves daily observations of all key variables. Table 1 presents the summary statistics of the main series. The weekly and monthly implied volatility have average (annualized) values of 21.345% and 21.837%, respectively.

[11] One may argue that we should start our sample period at a later date (e.g., in 2011 as in Andersen et al., 2017) to allow the trading activity in weekly options to pick up. We also considered this alternative starting date and reached qualitatively similar results. If one holds the view that our sample period includes illiquid weeklies, then this low trading activity should work against the predictive power of the weekly options. Viewed in this way, the gains in forecasting performance we document represent a "worst case" scenario.

[12] Once a month, we observe a standard option contract with 7 days left to expiration. Thus, if one changes the sampling frequency to the monthly level, it is possible to extend the sample period back to 1996 and repeat the main analyses. We explore this possibility in Subsection 5.4 and show that the main results are robust to this change. We thank the reviewer for suggesting this analysis.

3 Weekly vs. monthly implied variance

In this section, we examine the in-sample predictive power of the weekly implied variance for next week's realized variance and compare it to that of the monthly implied variance. We also explore the channels through which this predictability arises.

3.1 The information content of the weekly implied variance

3.1.1 Univariate evidence

We begin by evaluating the information content of implied variance for the week ahead realized variance. To this end, we estimate the following Mincer and Zarnowitz (1969) regression:

$$RV_{t+7}^{w} = \alpha + \beta IV_{t}^{x} + \epsilon_{t+7}, \tag{7}$$

where α is the intercept. β denotes the slope parameter. IV_t^x is the implied variance on day t of time to maturity x, where x can be the weekly (w) or monthly (m) maturity. ϵ_{t+7} is the residual of the regression at $t+7$. If the implied variance is informative about the future weekly realized variance, then the slope parameter will be significantly different from 0.

Starting with the monthly implied variance, Table 2 reports a positive and statistically significant (t–stat=6.676) slope estimate of 0.785. The explanatory power associated with this regression (Adj R^2=61.6%) confirms that the monthly implied variance predicts the weekly realized variance.

Turning to the weekly implied variance, we can see that it predicts the future realized variance with a slope estimate of 0.713 (t–stat=7.240). The corresponding predictive power is equal to 64.9%. Several points are worth highlighting. First, the explanatory power of this regression model is higher than that of the model that relies on the monthly implied variance. This result indicates that the weekly implied variance has superior predictive ability for short-term risk than the often-used monthly implied variance. Second, the slope estimates obtained from these univariate regressions are similar and significantly different from 1. In the earlier literature (e.g., Canina and Figlewski, 1993; Lamoureux and Lastrapes, 1993), a slope parameter that is significantly different from 1 was interpreted as evidence against the expectations hypothesis. However, Chernov (2007) and Prokopczuk and Wese Simen (2014) point out that this result can arise in a setting where the variance risk premium is time-varying.[13] It is thus useful to analyze the average variance risk premium, defined as the difference between the implied variance and the contemporaneously estimated realized variance (Bollerslev et al., 2009), of each maturity. Table 1 reports an average of 2.280% with a volatility of 4.659% for the (annualized) weekly variance risk premium and an average of 2.300% with a volatility of 3.107% for the (annualized) monthly variance risk premium. Thus, there is evidence of a non-zero and time-varying variance risk premium in both the weekly and monthly implied variances.

3.1.2 Multivariate evidence

The preceding analysis shows that, when used alone, the weekly and monthly implied variance predict the future weekly realized variance. However, it does not directly shed light on the incremental information content of these two predictors.

[13]Note that the high persistence of the implied variance series, along with the short sample, could also easily generate a spurious mean-reversion leading to a slope coefficient lower than 1. We thank the reviewer for providing this insight.

To address this question, we estimate the following model:

$$RV^w_{t+7} = \alpha + \beta IV^w_t + \gamma \hat{\nu}^{m/w}_t + \epsilon_{t+7}, \tag{8}$$

where α is the intercept. β and γ are the slope parameters. IV^w_t is the weekly implied variance on day t. $\hat{\nu}^{m/w}_t$ is the estimated residual at time t of the regression of the monthly implied variance on a constant and the weekly implied variance (see equation (9)):[14]

$$IV^m_t = \phi_0 + \phi_1 IV^w_t + \nu^{m/w}_t. \tag{9}$$

The last row of Table 2 reports that the slope estimate associated with the weekly component is significant (t–stat=7.025), whereas that of the orthogonalized monthly implied variance is not (t–stat=0.631). Moreover, the explanatory power of the encompassing model (Adj R^2=65.0%) is very similar to that of the univariate model, which uses the weekly implied variance as a forecasting variable (Adj R^2=64.9%). Taken as a whole, the results suggest that the orthogonalized monthly implied variance does not add to the predictive power of the weekly implied variance. These findings are important because empirical studies routinely discard short-term options data on the grounds that they are noisy and thus uninformative. Our results caution that, by following this approach, one throws away valuable information about short-term risk. These findings are also interesting given the growing practice of using the monthly implied variance to predict the short-term realized variance. Our evidence reveals that this methodology may not be the best way of modeling the short-term realized variance.

[14] In an earlier version of the paper, we included the monthly implied variance rather than its orthogonalized component with respect to the weekly implied variance. While this analysis led to qualitatively similar conclusions, it was vulnerable to the concern that the results are difficult to interpret given the high correlation between the two implied variance series. By using the orthogonalized component of the monthly implied variance, we assuage this concern. We thank the reviewer for suggesting this analysis.

3.2 Dissecting the predictability

Having established the in-sample predictive power of the weekly implied variance for weekly realized variance, we next explore the channel through which this predictability arises.

3.2.1 Framework

Our starting point is the theory of quadratic variation (Barndorff-Nielsen and Shephard, 2002), which posits that the realized variance of asset returns can be decomposed into components linked to (i) the significant continuous variation and (ii) the significant jump variation of the asset returns. More formally, we have:

$$RV^w_{t+7} = C^w_{t+7} + J^w_{t+7}, \tag{10}$$

where C^w_{t+7} and J^w_{t+7} are the significant weekly continuous and jump variations of the asset returns computed over the week ending on day $t+7$, respectively.

This insight suggests that there are two channels through which short-term implied variance may be informative about next week's realized variance. The first possibility is that the weekly implied variance contains information about the significant continuous variation of returns. The second possibility is that the weekly implied variance is informative about the significant jump variation.

Barndorff-Nielsen and Shephard (2002) propose the bipower variation as an estimator of the continuous variation of asset returns. Andersen et al. (2012) subsequently establish that the $MedRV$ estimator has better properties than the bipower variation. Thus, we use the $MedRV$ to estimate the continuous variation of returns:

$$CV^w_{t+7} = \underbrace{\frac{52 \times m\pi}{(6 - 4\sqrt{3} + \pi)(m-2)} \sum_{j=0}^{N^w_{t+7}-1} \sum_{k=3}^{m} \text{median}(|r_{t+7-j,k}|, |r_{t+7-j,k-1}|, |r_{t+7-j,k-2}|)^2}_{MedRV\ \ Estimator}, \tag{11}$$

where median$(\cdot)$ is the median operator and all other variables are as previously defined. Next, we modify the test statistic presented in Huang and Tauchen (2005), which relies on the bipower variation and the realized quarticity to take advantage of the more robust estimators of the continuous variation ($MedRV$) and realized quarticity ($MedRQ$). We are thus able to compute the significant continuous and jump variations:

$$z^w_{t+7} = m^{1/2} \left(\frac{\frac{RV^w_{t+7} - CV^w_{t+7}}{RV^w_{t+7}}}{\sqrt{0.96 \max(1, \frac{MedRQ^w_{t+7}}{(CV^w_{t+7})^2})}} \right) \tag{12}$$

$$MedRQ^w_{t+7} = A_w \times \sum_{j=0}^{N^w_{t+7}-1} \sum_{k=3}^{m} \text{median}(|r_{t+7-j,k}|, |r_{t+7-j,k-1}|, |r_{t+7-j,k-2}|)^4 \tag{13}$$

$$A_w = \frac{52^2 \times 3m^2\pi}{(72 - 52\sqrt{3} + 9\pi)(m-2)} \tag{14}$$

$$C^w_{t+7} = I_{z^w_{t+7} > \phi_{1-\alpha}} CV^w_{t+7} + I_{z^w_{t+7} \leq \phi_{1-\alpha}} RV^w_{t+7} \tag{15}$$

$$J^w_{t+7} = I_{z^w_{t+7} > \phi_{1-\alpha}} (RV^w_{t+7} - CV^w_{t+7}), \tag{16}$$

where $\phi_{1-\alpha}$ is the critical value from the cumulative standard normal distribution at confidence level $1 - \alpha$. I is the indicator function. Similar to Andersen et al. (2007), we employ $\alpha = 99.9\,\%$ throughout this paper.

Figures 1 to 3 show the daily time series of the realized variance, the significant continuous variation, and the significant jump variation, respectively. Consistent with the existing evidence (e.g., Andersen et al., 2007), the realized variance and continuous variation series are quite persistent and increase during turmoil periods, such as the 2007–2009 financial crisis. Moreover, from Figure 3 we see that the jump intensity varies over time and significant jumps tend to cluster during more volatile periods.

3.2.2 Significant continuous variation

We regress the time series of the significant continuous variation on a constant and the lagged implied variance series:

$$C^w_{t+7} = \alpha + \beta IV^x_t + \epsilon_{t+7}, \tag{17}$$

where all variables are as previously defined.

Panel A of Table 3 shows that, in univariate regressions, each maturity of the implied variance predicts the significant continuous component of the realized variance. This conclusion is borne out by the significant slope estimates in the univariate regressions. Similar to our analysis of the realized variance, we find that the weekly implied variance has the higher predictive power (Adj R^2=63.6%) of the two variables. The encompassing model (last two rows of Panel A) yields an Adj R^2 of 63.7%, which is very close to that of the univariate model which relies on the weekly implied variance alone.

3.2.3 Significant jump variation

We now estimate the following predictive regression:

$$J^w_{t+7} = \alpha + \beta IV^x_t + \epsilon_{t+7}, \tag{18}$$

where all variables are as previously defined.

Panel B of Table 3 documents that it is harder to accurately model the significant jump variation than the significant continuous variation. This conclusion is evidenced by the lower explanatory power for the significant jump variation compared to that of the significant continuous variation (see Panel A).[15] We can see that each maturity of the implied variance individually predicts the significant weekly jump variation with very similar explanatory power. However, the encompassing regression reveals that the weekly implied variance crowds out the orthogonalized monthly implied variance.

4 Implied variance vs. time series models

The previous section shows that the weekly implied variance is superior to the monthly implied variance in-sample. However, it is not clear how it compares to other sophisticated time series models proposed in the literature. In this section, we present the competing models. Next, we assess their in- and out-of-sample predictive performance.

4.1 Introducing the competing models

We use the heterogeneous autoregressive (HAR) realized variance model (Corsi, 2009) as our benchmark:

$$RV^w_{t+7} = \alpha + \beta RV^d_t + \gamma RV^w_t + \delta RV^m_t + \epsilon_{t+7}, \tag{19}$$

where α is the intercept. β, γ, and δ are the slope parameters. RV^d_t and RV^m_t are the (annualized) daily and monthly realized variance at time t, respectively. These series are

[15] Using various time series models, Busch et al. (2011) document a similar result at the monthly horizon.

computed using the following estimators:

$$
\begin{aligned}
RV_t^d &= 252 \times \sum_{k=1}^{m} r_{t,k}^2 && (20)\\
RV_t^m &= 12 \times \sum_{j=0}^{N_t^m-1} \sum_{k=1}^{m} r_{t-j,k}^2, && (21)
\end{aligned}
$$

where the numbers 252 and 12 serve to annualize the daily and monthly realized variance estimates, respectively. N_t^m is the number of trading days in the calendar month ending on day t.

Building on the work of Andersen et al. (2007), we also consider the continuous heterogeneous autoregressive (CHAR) model, in which each historical variance in the HAR model is replaced with the continuous variation of the corresponding horizon:

$$
RV_{t+7}^w = \alpha + \beta CV_t^d + \gamma CV_t^w + \delta CV_t^m + \epsilon_{t+7}, \tag{22}
$$

where CV_t^d and CV_t^m are the (annualized) daily and monthly continuous variations at time t, respectively:

$$
\begin{aligned}
CV_t^d &= \frac{252 \times m\pi}{(6-4\sqrt{3}+\pi)(m-2)} \sum_{k=3}^{m} \text{median}(|r_{t,k}|, |r_{t,k-1}|, |r_{t,k-2}|)^2 && (23)\\
CV_t^m &= \frac{12 \times m\pi}{(6-4\sqrt{3}+\pi)(m-2)} \sum_{j=0}^{N_t^m-1} \sum_{k=3}^{m} \text{median}(|r_{t-j,k}|, |r_{t-j,k-1}|, |r_{t-j,k-2}|)^2. && (24)
\end{aligned}
$$

We also analyze the performance of the HAR–J model (Andersen et al., 2007), which augments the HAR model with the lagged significant daily jump variation:

$$RV_{t+7}^{w} = \alpha + \beta RV_t^d + \gamma RV_t^w + \delta RV_t^m + \eta J_t^d + \epsilon_{t+7}, \tag{25}$$

where α, β, γ, δ, and η are parameters to estimate. J_t^d is the (annualized) statistically significant daily jump variation at time t computed as follows:

$$\begin{aligned}
J_t^d &= I_{z_t^d > \phi_{1-\alpha}}(RV_t^d - CV_t^d) && (26)\\
z_t^d &= m^{1/2}\left(\frac{\frac{RV_t^d - CV_t^d}{RV_t^d}}{\sqrt{0.96\max(1, \frac{MedRQ_t^d}{(CV_t^d)^2})}}\right) && (27)\\
MedRQ_t^d &= A_d \times \sum_{k=3}^{m} \text{median}(|r_{t,k}|, |r_{t,k-1}|, |r_{t,k-2}|)^4 && (28)\\
A_d &= \frac{252^2 \times 3m^2\pi}{(72 - 52\sqrt{3} + 9\pi)(m-2)}. && (29)
\end{aligned}$$

We also evaluate the forecasting performance of the HAR–C–J model (Andersen et al., 2007). Essentially, this model decomposes each historical variance in the HAR model into the corresponding significant continuous and jump variations:

$$RV_{t+7}^{w} = \alpha + \beta C_t^d + \gamma J_t^d + \delta C_t^w + \eta J_t^w + \theta C_t^m + \kappa J_t^m + \epsilon_{t+7}, \tag{30}$$

where α, β, γ, δ, η, θ, and κ are parameters to estimate. C_t^d and C_t^m are the corresponding daily and monthly significant continuous variation at time t, respectively. J_t^m denotes the (annualized) monthly significant jump variation at time t. We obtain C_t^d, C_t^m, and J_t^m as follows:

$$C_t^d = I_{z_t^d > \phi_{1-\alpha}} CV_t^d + I_{z_t^d \le \phi_{1-\alpha}} RV_t^d \tag{31}$$

$$C_t^m = I_{z_t^m > \phi_{1-\alpha}} CV_t^m + I_{z_t^m \leq \phi_{1-\alpha}} RV_t^m \tag{32}$$

$$J_t^m = I_{z_t^m > \phi_{1-\alpha}} (RV_t^m - CV_t^m) \tag{33}$$

$$z_t^m = m^{1/2} \left(\frac{\frac{RV_t^m - CV_t^m}{RV_t^m}}{\sqrt{0.96 \max(1, \frac{MedRQ_t^m}{(CV_t^m)^2})}} \right) \tag{34}$$

$$MedRQ_t^m = A_m \times \sum_{j=0}^{N_t^m - 1} \sum_{k=3}^{m} \text{median}(|r_{t-j,k}|, |r_{t-j,k-1}|, |r_{t-j,k-2}|)^4 \tag{35}$$

$$A_m = \frac{12^2 \times 3m^2\pi}{(72 - 52\sqrt{3} + 9\pi)(m-2)}. \tag{36}$$

Patton and Sheppard (2015) document the good empirical performance of the semi-variance heterogeneous autoregressive (SHAR) model. Essentially, this model modifies the HAR specification by decomposing each historical variance term into positive and negative semivariance components:[16]

$$RV_{t+7}^w = \alpha + \beta SV_t^{d+} + \gamma SV_t^{d-} + \delta SV_t^{w+} + \eta SV_t^{w-} + \theta SV_t^{m+} + \kappa SV_t^{m-} + \epsilon_{t+7}, \tag{37}$$

where α, β, γ, δ, η, θ, and κ are parameters to estimate. SV_t^{d+}, SV_t^{w+}, and SV_t^{m+} are the positive (annualized) daily, weekly, and monthly semivariances at time t, respectively. SV_t^{d-}, SV_t^{w-}, and SV_t^{m-} are the negative (annualized) daily, weekly, and monthly semivariances at time t, respectively. We compute these variables as follows:

$$SV_t^{d+} = 252 \times \sum_{k=1}^{m} r_{t,k}^2 I_{r_{t,k}>0} \tag{38}$$

$$SV_t^{d-} = 252 \times \sum_{k=1}^{m} r_{t,k}^2 I_{r_{t,k}<0} \tag{39}$$

$$SV_t^{w+} = 52 \times \sum_{j=0}^{N_t^w - 1} \sum_{k=1}^{m} r_{t-j,k}^2 I_{r_{t-j,k}>0} \tag{40}$$

[16] We have also considered more parsimonious specifications of the SHAR model where we only decompose the historical variance of a specific horizon into positive and negative semivariance components. We found very little to distinguish between these alternative specifications.

$$SV_t^{w-} = 52 \times \sum_{j=0}^{N_t^w - 1} \sum_{k=1}^{m} r_{t-j,k}^2 I_{r_{t-j,k}<0} \tag{41}$$

$$SV_t^{m+} = 12 \times \sum_{j=0}^{N_t^m - 1} \sum_{k=1}^{m} r_{t-j,k}^2 I_{r_{t-j,k}>0} \tag{42}$$

$$SV_t^{m-} = 12 \times \sum_{j=0}^{N_t^m - 1} \sum_{k=1}^{m} r_{t-j,k}^2 I_{r_{t-j,k}<0}. \tag{43}$$

4.2 In-sample evidence

Table 4 shows the in-sample forecasting performance of each model. Starting with the HAR model, we can see that it yields an explanatory power of 59.8 %. The CHAR model yields a comparable explanatory power of 59.6 %. The fit of the HAR–J model to the data (Adj R^2=59.8 %) is similar to that of the HAR model. This result arises because the exposure to the significant daily jump variation is not statistically significant. Turning to the HAR–C–J and SHAR models, we can see that they slightly improve on the benchmark HAR model as evidenced by their Adj R^2s of 60.8 % and 60.7 %, respectively. This conclusion is consistent with the in-sample finding of Patton and Sheppard (2015).

Comparing the Adj R^2s in Tables 2 and 4, we can see that the weekly implied variance achieves the highest explanatory power (Adj R^2=64.9 %). This result leads us to conclude that the weekly implied variance performs better than the HAR model and its extensions in-sample.

4.3 Out-of-sample evidence

We next investigate whether the in-sample predictability results also extend out-of-sample. To this end, we define the models in equation (7) based on the IV^w and IV^m as the IVW and IVM models, respectively. We use a rolling window of four years of daily data to estimate the forecasting models in equations (7), (19), (22), (25), (30), and (37).[17] We then use all relevant information available in real-time to generate the conditional expectation of next week's realized variance. Similar to Bollerslev et al. (2016), we subject these forecasts to the "insanity filter" to guard against implausible variance forecasts. If the forecast is higher (lower) than the highest (lowest) weekly realized variance observed in the estimation window, we set the forecast to the average weekly realized variance in the estimation window. This filter also enables us to avoid the situation where the variance forecast could be negative.[18]

Repeating the steps above for each rolling window, we obtain the time series of the out-of-sample variance forecasts, which we then compare to the realized variance. We compute the mean squared error (MSE) and quasi-likelihood ($QLIKE$) loss functions as follows:

$$MSE = \frac{1}{T}\sum_{t=1}^{T}\left(RV_{t+7}^{w} - E_t(RV_{t+7}^{w})\right)^2 \tag{44}$$

$$QLIKE = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{RV_{t+7}^{w}}{E_t(RV_{t+7}^{w})} - \log\frac{RV_{t+7}^{w}}{E_t(RV_{t+7}^{w})} - 1\right), \tag{45}$$

where T is the total number of out-of-sample forecasts. $E_t(RV_{t+7}^{w})$ is the forecast at time t of the variance to be realized at $t+7$. All other variables are as previously defined.

Patton (2011) shows that these two loss functions are robust to the presence of noise in the proxy used for the unobservable variance, making them well-suited for our analysis.[19] Table 5 reports the ratio of the loss function (name in row) associated with the model (name in column) over that of the benchmark HAR model. An entry equal to 1 indicates

[17] A window of four years is consistent with the research of Bollerslev et al. (2016) and Patton and Sheppard (2015).

[18] As a robustness check, we remove the filter and obtain similar results.

[19] In Subsection 5.1, we further discuss the implications of measurement errors.

that the model does as well as the benchmark HAR model. Entries lower than 1 suggest that the model achieves lower average forecast errors than the HAR model. Conversely, entries greater than 1 indicate that the average forecast errors of the model are higher than those of the HAR model.

Focusing on the entries reported for the IVW model, we can see that the MSE and $QLIKE$ ratios are equal to 0.797 and 0.870, respectively. This set of numbers reveals that a forecasting model based on the weekly implied variance reduces the average forecast errors of the benchmark HAR model by 20.3 % (MSE) and 13.0 % ($QLIKE$). Looking across the competing models, we observe that their performance is generally inferior to that of the weekly implied variance.

We formally test the null hypothesis that the average forecast loss associated with the weekly implied variance is equal to or greater than that of the best forecasting model among its competitors (IVM, HAR, CHAR, HAR–J, HAR-C–J, and SHAR). The alternate hypothesis is that the weekly implied variance delivers lower average forecast errors than its competitors. To implement this test, we modify the Reality Check of White (2000) as in Bollerslev et al. (2016). In our empirical implementation, we use the stationary bootstrap of Politis and Romano (1994) with 9,999 re-samplings and an average block length of 10.[20] We find that the null hypothesis is rejected at the 5% significance level with p-values of 1.9% and 1.6% for the MSE and $QLIKE$ loss functions, respectively. We thus conclude that the weekly implied variance achieves significantly lower forecasting errors relative to its competitors.

[20] We experiment with different block lengths and obtain very similar results.

5 What about ...

This section presents several robustness checks. First, we explore whether our results are affected by measurement errors in the realized variance estimates. Second, we examine the impact of potential errors induced by the method of interpolation we use to obtain the implied variance series. Third, we assess the sensitivity of our findings to the estimator of implied variance. Fourth, we consider a longer sample period that dates back to 1996. Fifth, we analyze the predictability of quarterly realized variance.

5.1 The noise in the realized variance?

Bollerslev et al. (2016) propose a forecasting model that extends the HAR model by taking into account the measurement errors in the historical variance. These errors arise from the fact that the historical variance is not directly observable. Thus, one needs to estimate the historical variance before using it for forecasting, leading to the errors-in-variables problem. The authors introduce a modeling framework, termed HAR–RQ, which aims to capture the heteroskedasticity of the measurement errors and improve the realized variance forecasts:[21]

$$\begin{aligned} RV_{t+7}^{w} &= \alpha + (\beta + \gamma\sqrt{MedRQ_t^d})RV_t^d + (\delta + \eta\sqrt{MedRQ_t^w})RV_t^w \\ &\quad + (\theta + \kappa\sqrt{MedRQ_t^m})RV_t^m + \epsilon_{t+7}, \end{aligned} \tag{46}$$

[21] Bollerslev et al. (2016) use a square root specification for the measurement error correction on the basis that it has an imbued robustness. We also experiment with the logarithmic specification and find it delivers inferior forecasting performance compared to the square root specification. This finding is consistent with the authors' argument and their own empirical results.

where α is the intercept. β, γ, δ, η, θ, and κ are slope parameters.[22] All other variables are as previously defined.

Using the same methodology as before, we analyze the out-of-sample predictive performance of the HAR–RQ model. Consistent with Bollerslev et al. (2016), Table 5 shows that this model yields more accurate forecasts than the CHAR, HAR–J, and HAR–C–J models. However, its performance is inferior to that of the weekly implied variance, indicating that the short-term implied variance provides more accurate forecasts than the HAR–RQ model. Our untabulated analysis also reveals that the difference in the performance of the two models is statistically significant.

In addition, we consider more parsimonious specifications of the HAR–RQ model in equation (46), namely: the HAR–RQ–D (equation(47)), HAR–RQ–W (equation(48)), and HAR–RQ–M (equation(49)):

$$RV^w_{t+7} = \alpha + (\beta + \gamma\sqrt{MedRQ^d_t})RV^d_t + \delta RV^w_t + \theta RV^m_t + \epsilon_{t+7} \quad (47)$$

$$RV^w_{t+7} = \alpha + \beta RV^d_t + (\delta + \eta\sqrt{MedRQ^w_t})RV^w_t + \theta RV^m_t + \epsilon_{t+7} \quad (48)$$

$$RV^w_{t+7} = \alpha + \beta RV^d_t + \delta RV^w_t + (\theta + \kappa\sqrt{MedRQ^m_t})RV^m_t + \epsilon_{t+7}. \quad (49)$$

Our untabulated analysis reveals that these parameterizations do not outperform the weekly implied variance. For instance, the HAR–RQ–D and HAR–RQ–W models yield MSE ($QLIKE$) loss ratios of 0.864 (0.917) and 0.861 (0.902), respectively. While the HAR–RQ–W performs better than the more general HAR–RQ specification in equation (46), a finding consistent with the work of Bollerslev et al. (2016), it does not outperform the weekly implied variance.

[22] Bollerslev et al. (2016) use the realized quarticity as defined in Barndorff-Nielsen (2002) rather than the more robust $MedRQ$ estimator of Andersen et al. (2012). We also repeat the analysis using the same estimator as Bollerslev et al. (2016). The results are qualitatively similar.

From a theoretical standpoint, the noise in the realized variance series should be larger at lower sampling frequencies (Barndorff-Nielsen, 2002). Thus, by sampling the data at finer frequencies, one should be able to dampen the effect of the noise. This insight motivates us to increase the sampling frequency from 5-minute to 1-minute and repeat our out-of-sample analysis. Table 6 points to the same conclusion as for Table 5: the weekly implied variance yields the highest improvement in forecast accuracy.

5.2 The method of interpolation?

We obtain our implied variance series of constant maturity by applying a linear interpolation across maturities. It is thus interesting to assess the sensitivity of our results to the method of interpolation. To this end, we consider the nearest neighborhood interpolation method as an alternative approach. Essentially, we use the variance implied by option contracts of maturity closest to 1 week (month) in order to estimate the implied variance of the weekly (monthly) maturity. With this new time series, we repeat our out-of-sample analysis. The results in Table 7 confirm our main findings presented in Table 5. The weekly implied variance is a strong predictor of the weekly realized variance. It reduces the forecasting error of the benchmark HAR model by 33.7 % and 18.9 % when considering the MSE and $QLIKE$ loss functions, respectively. Moreover, the weekly implied variance outperforms the monthly implied variance as well as the other variance forecasting models of the HAR family.

5.3 The implied variance estimator?

Our interest in the Bakshi et al. (2003) estimator is motivated by its robustness to jumps (Du and Kapadia, 2013). However, most studies use the Britten-Jones and

Neuberger (2000) estimator, making our results difficult to directly compare to those in the extant literature (e.g., Jiang and Tian, 2005; Taylor et al., 2010). As such, we use the numerical scheme presented in Subsection 2.1.2 to implement the Britten-Jones and Neuberger (2000) estimator of implied variance:[23]

$$IV_t^{\tau} = \frac{2e^{rf_t \frac{\tau}{360}}}{\frac{\tau}{360}} \left(\int_0^{S_t} \frac{P_t(\tau, K)}{K^2} dK + \int_{S_t}^{\infty} \frac{C_t(\tau, K)}{K^2} dK \right), \tag{50}$$

where all variables are as previously defined.

We use the resulting time series to repeat our out-of-sample analysis. The results in Table 8 show that the loss ratios associated with the implied variance series are generally higher than those based on the Bakshi et al. (2003) estimator (see Table 5). This is not surprising since the Britten-Jones and Neuberger (2000) estimator is biased in the presence of jumps (Du and Kapadia, 2013), resulting in larger forecast errors. However, most important for our purpose, the weekly implied variance series outperforms its monthly counterpart and all the competing models of the HAR family. This is true for both the MSE and the $QLIKE$ loss functions. Overall, these results are consistent with our main findings presented in Table 5.

[23]Generally in the literature, the squared value of the volatility index (VIX) is employed instead of the monthly implied variance series computed using the Britten-Jones and Neuberger (2000) estimator and the numerical method presented in Subsection 2.1.2. It is interesting to compare the performance of the implied variance estimates obtained from these two approaches. By doing so, one can provide insight on the impact of the numerical method on the results. Untabulated results show that the time series of the square of the VIX and the Britten-Jones and Neuberger (2000) series of monthly maturity we computed are highly correlated and very similar. Empirically, the squared VIX series yields a $QLIKE$ loss ratio of 1.251, which is comparable to the 1.217 figure based on the monthly implied variance using the Britten-Jones and Neuberger (2000) estimator (see Table 8). We thus conclude that the numerical method does not have a major impact on the results.

5.4 A longer sample period?

Our main analysis focuses on the period following the introduction of weekly options. However, one could extend the sample period back to January 1996 since, each month, we have one observation of a standard option with seven days left to maturity (i.e., a standard option eventually becomes a weekly option once a month). We can thus check whether the main results hold in the extended sample period from January 4, 1996 to August 31, 2015. Note that, by extending the sample period to 1996, we are forced to specify the regressions to the monthly frequency.

Accordingly, we sample the implied variance series one week before the expiration of the closest to expiration standard option. We use this new dataset to repeat our main analyses. Tables A.1–A.4 in the Online Appendix confirm our main findings. In particular, the weekly implied variance outperforms the monthly implied variance and the recently proposed time series models of the HAR family. This finding holds both in- and out-of-sample.

5.5 A longer forecasting horizon?

Up to this point, our results show that, if one wants to predict the realized variance over the next week, it is better to use options with the same maturity as the forecasting horizon as opposed to simply using monthly options, the usual practice in the literature. Naturally, one may wonder: Is this a more general phenomenon? For instance, does the quarterly implied variance outperform the monthly implied variance when it comes to predicting the quarterly realized variance? Although this is not the main goal of our study, it is interesting to explore this possibility.[24]

[24]We thank the reviewer for suggesting this analysis.

We compute the (annualized) quarterly realized variance as follows:

$$RV^q_{t+90} = 4 \times \sum_{j=0}^{N^q_{t+90}-1} \sum_{k=1}^{m} r^2_{t+90-j,k}, \tag{51}$$

where RV^q_{t+90} is the (annualized) quarterly realized variance for the quarter ending on day $t + 90$. The number 4 indicates that the realized variance estimate is annualized. N^q_{t+90} indicates the number of trading days during the quarter ending on day $t + 90$.

Next, we estimate the following variance forecasting regression:

$$RV^q_{t+90} = \alpha + \beta IV^x_t + \epsilon_{t+90}, \tag{52}$$

where α is the intercept. β denotes the slope parameter. IV^x_t is the implied variance on day t of time to maturity x, where x can be the quarterly (q) or monthly (m) maturity. ϵ_{t+90} is the residual of the regression at $t + 90$.

Given the large amount of overlap between consecutive daily observations of the realized variance, we sample all data at the quarterly frequency to obtain non-overlapping data samples. By taking this step, we address concerns related to the overlapping observation biases that typically plague long-horizon regressions. Unfortunately, this choice means that we have a very limited sample of independent observations. We attempt to mitigate this issue by extending the sample period to 1996, as in Subsection 5.4. Nonetheless, the results of the in-sample analysis should be interpreted with caution.[25] Table A.5 in the Online Appendix shows that, in the univariate regressions, the quarterly implied variance yields an Adj R^2 of 17.8 %, which is slightly higher than that of the monthly implied variance.

[25]Another implication of the limited sample size is that we are not able to carry out a reliable out-of-sample analysis.

Our working hypothesis is that the quarterly implied variance contains all relevant information to predict the quarterly realized variance. Consequently, the orthogonal component of the monthly implied variance with respect to the quarterly implied variance should not contain information about the future quarterly realized variance. This insight motivates the following encompassing model:

$$RV_{t+90}^{q} = \alpha + \beta IV_t^q + \gamma \hat{\nu}_t^{m/q} + \epsilon_{t+90}, \tag{53}$$

where α is the intercept. β and γ are the slope parameters. IV_t^q is the implied variance on day t of quarterly time to maturity. $\hat{\nu}_t^{m/q}$ is the estimated residual at time t of the regression of the monthly implied variance on a constant and the quarterly implied variance (see equation (54)):

$$IV_t^m = \phi_0 + \phi_1 IV_t^q + \nu_t^{m/q}. \tag{54}$$

The results presented in the last row in Table A.5 of the Online Appendix are consistent with our main hypothesis. Namely, the orthogonalized monthly implied variance does not add to the information content of the quarterly implied variance.

6 Conclusion

We exploit weekly options on the S&P 500 index to construct the weekly implied variance series and explore its information content for predicting short-term realized variance. Our results reveal that the weekly implied variance is a powerful predictor of future weekly realized variance. Out-of-sample, the weekly implied variance outperforms not only the monthly implied variance but also the HAR model and its various extensions.

Our evidence carries implications for both academics and practitioners. For practitioners, it would be interesting for the Chicago Board of Options Exchange (CBOE) to compute and disseminate the time series of the weekly implied volatility index. This series, which would sit alongside the popular 30-day volatility index, would be useful for market participants to better gauge and manage short-term risk. For academics, our results suggest that one would benefit from not discarding short-term options on the grounds that they are illiquid and thus uninformative. Our analysis clearly shows that these options are more informative about future short-term variance than the monthly options that have been analyzed.

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.

Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *Review of Economics and Statistics*, 89(4):701–720.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modelling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Andersen, T. G., Fusari, N., and Todorov, V. (2017). Short-term market risks implied by weekly options. *Journal of Finance*, 72(3):1335–1386.

Bakshi, G., Kapadia, N., and Madan, D. (2003). Stock return characteristics, skew laws, and the differential pricing of individual equity options. *Review of Financial Studies*, 16(1):101–143.

Barndorff-Nielsen, O. E. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(5):457–477.

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

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Blair, B. J., Poon, S.-H., and Taylor, S. J. (2001). Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics*, 105(1):5–26.

Bollerslev, T., Patton, A. J., and Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192(1):1–18.

Bollerslev, T., Tauchen, G., and Zhou, H. (2009). Expected stock returns and variance risk premia. *Review of Financial Studies*, 22(11):4463–4492.

Bollerslev, T. and Todorov, V. (2011). Tails, fears, and risk premia. *Journal of Finance*, 66(6):2165–2211.

Britten-Jones, M. and Neuberger, A. (2000). Option prices, implied price processes and stochastic volatility. *Journal of Finance*, 55(2):839–866.

Busch, T., Christensen, B. J., and Nielsen, M. Ø. (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics*, 160(1):48–57.

Canina, L. and Figlewski, S. (1993). The informational content of implied volatility. *Review of Financial Studies*, 6(3):659–681.

Carr, P. and Wu, L. (2009). Variance risk premiums. *Review of Financial Studies*, 22(3):1311–1341.

Chernov, M. (2007). On the role of risk premia in volatility forecasting. *Journal of Business & Economic Statistics*, 25(4):411–426.

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

Drechsler, I. and Yaron, A. (2011). What's vol got to do with it. *Review of Financial Studies*, 24(1):1–45.

Du, J. and Kapadia, N. (2013). The tail in the volatility index. *EFA Meetings 2013 Paper*.

Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4):456–499.

Jiang, G. J. and Tian, Y. S. (2005). The model-free implied volatility and its information content. *Review of Financial Studies*, 18(4):1305–1342.

Kourtis, A., Markellos, R. N., and Symeonidis, L. (2016). An international comparison of implied, realized, and GARCH volatility forecasts. *Journal of Futures Markets*, 36(12):1164–1193.

Lamoureux, C. G. and Lastrapes, W. D. (1993). Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *Review of Financial Studies*, 6(2):293–326.

Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies*, 21(6):2535–2563.

Mincer, J. A. and Zarnowitz, V. (1969). The evaluation of economic forecasts. In Mincer, J., editor, *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pages 1–46. Cambridge, MA: Elsevier.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256.

Patton, A. J. and Sheppard, K. (2015). Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 97(3):683–697.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

Pong, S., Shackleton, M. B., Taylor, S. J., and Xu, X. (2004). Forecasting currency volatility: A comparison of implied volatilities and AR(FI)MA models. *Journal of Banking & Finance*, 28(10):2541–2563.

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

Prokopczuk, M. and Wese Simen, C. (2014). The importance of the volatility risk premium for volatility forecasting. *Journal of Banking & Finance*, 40:303–320.

Prokopczuk, M. and Wese Simen, C. (2016). What makes the market jump? *EFA Meetings 2016 Paper*.

Taylor, S. J., Yadav, P. K., and Zhang, Y. (2010). The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks. *Journal of Banking & Finance*, 34(4):871–881.

White, H. (2000). A reality check for data snooping. *Econometrica*, 68(5):1097–1126.

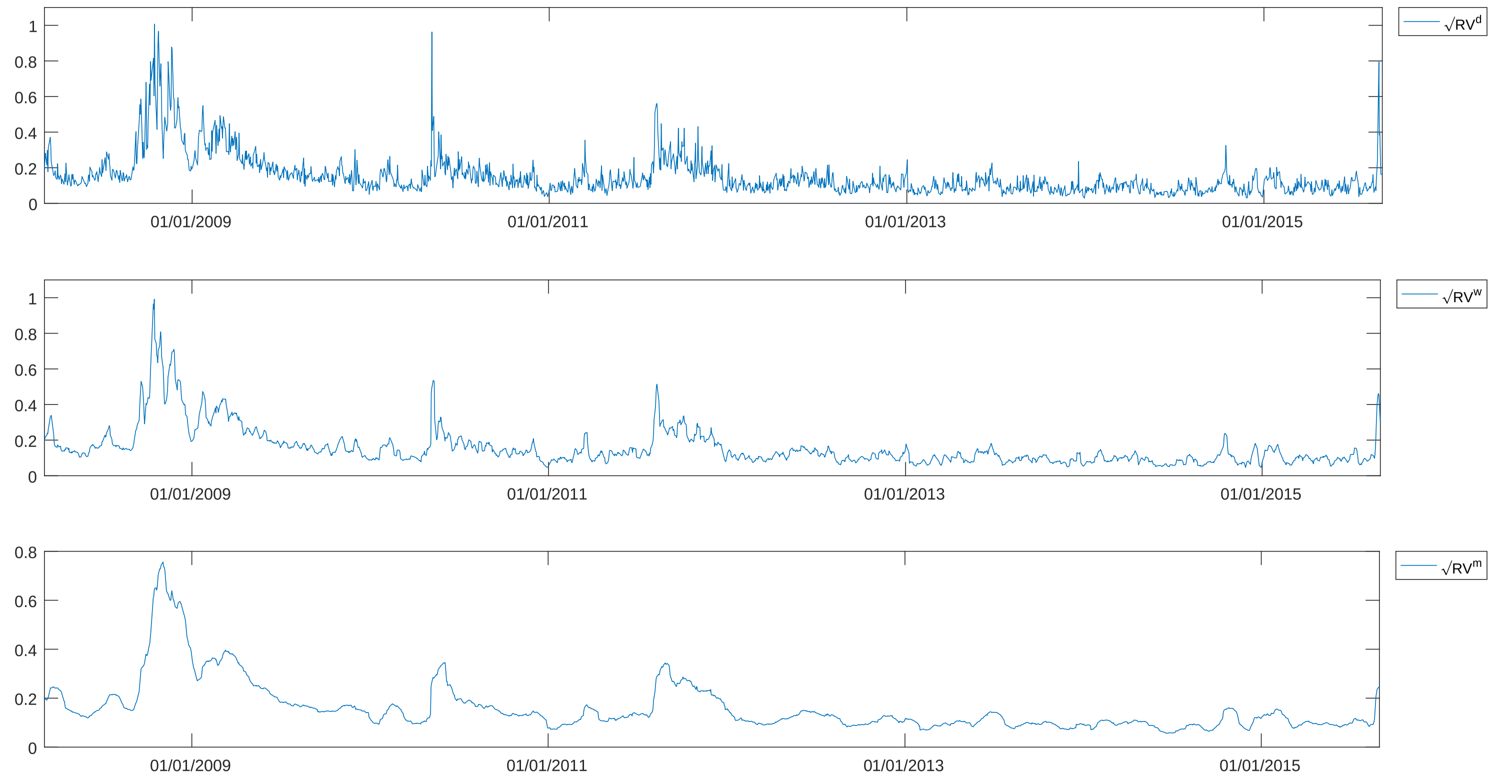

Figure 1: The dynamics of $\sqrt{RV}$

This figure shows the time series dynamics of the square root of the realized variance of daily (top panel), weekly (middle panel) and monthly (bottom panel) horizons, respectively. The horizontal axis shows the observation date, while the vertical axis indicates the (annualized) values of the series.

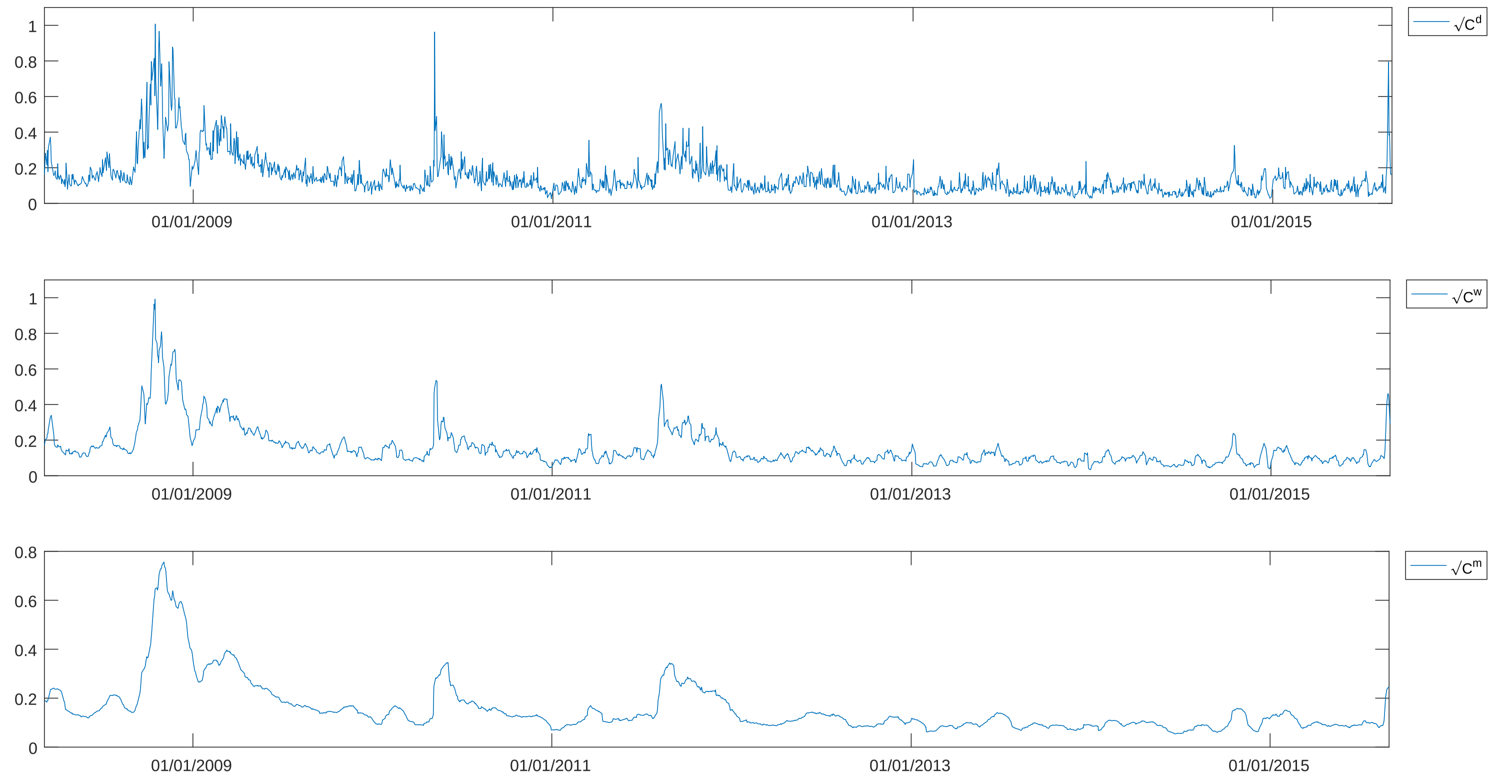

Figure 2: The dynamics of $\sqrt{C}$

This figure shows the time series dynamics of the square root of the significant continuous variation of daily (top panel), weekly (middle panel) and monthly (bottom panel) horizons, respectively. The horizontal axis shows the observation date, while the vertical axis indicates the (annualized) values of the series.

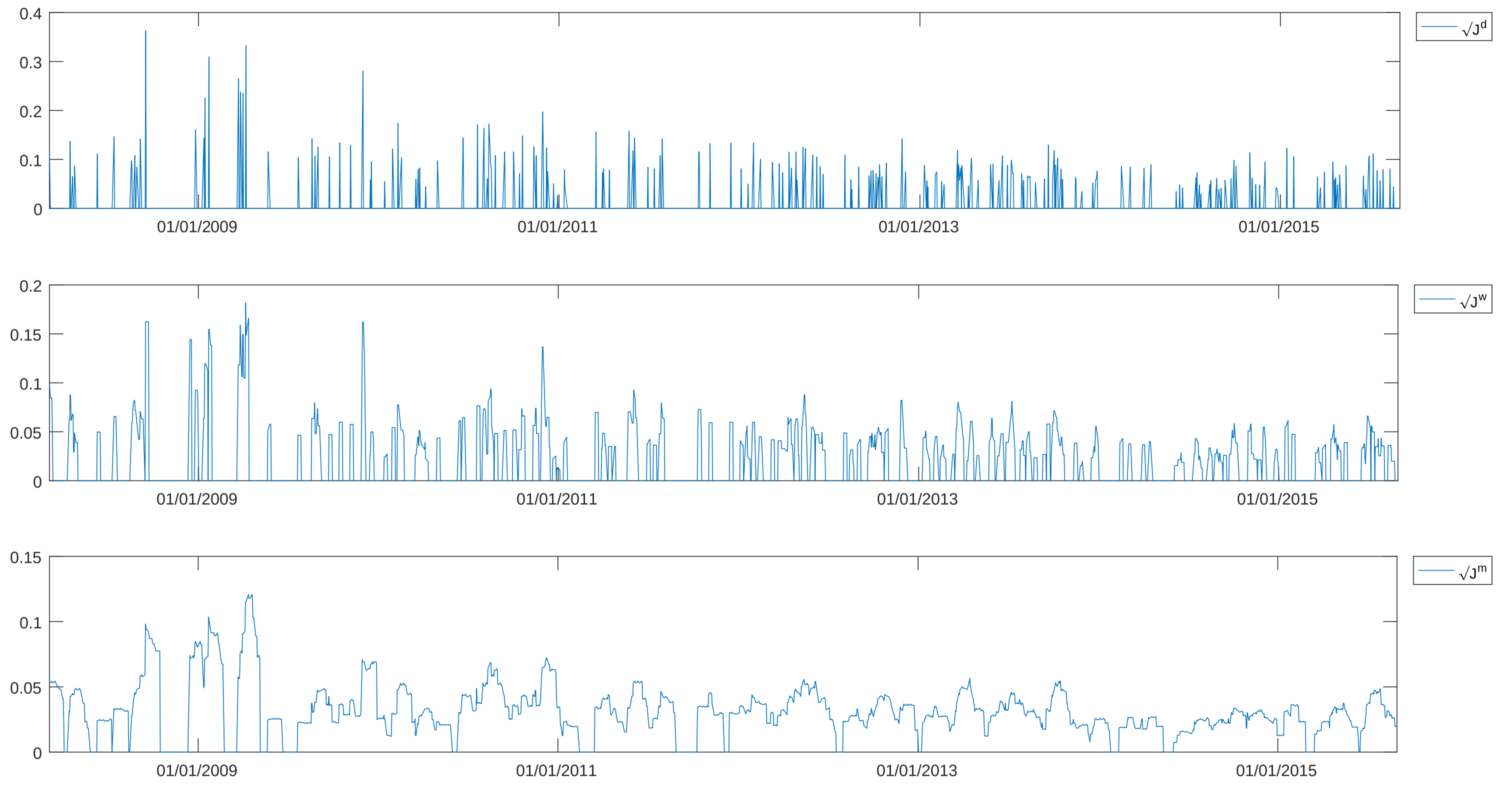

Figure 3: The dynamics of $\sqrt{J}$

This figure shows the time series dynamics of the square root of the significant jump variation of daily (top panel), weekly (middle panel) and monthly (bottom panel) horizons, respectively. The horizontal axis shows the observation date, while the vertical axis indicates the (annualized) values of the series.

Table 1: Descriptive statistics

This table presents key summary statistics. IV^w and IV^m denote the (annualized) implied variance of weekly, and monthly horizons, respectively. RV^d, RV^w, and RV^m denote the (annualized) realized variance of daily, weekly, and monthly horizons, respectively. VRP^w and VRP^m are the weekly and monthly variance risk premia, respectively. Similar to Bollerslev et al. (2009), the variance risk premium of a given maturity is defined as the difference between the implied variance of that maturity and the contemporaneously computed realized variance of the same maturity. Returns data are sampled at the 5-minute frequency. Mean is the average value of the daily time series of the variable. Std, Skew, and Kurt denote the standard deviation, skewness, and kurtosis of the variable, respectively.

	Mean	Std	Skew	Kurt
$\sqrt{IV^w}$	21.345%	11.850%	2.404	10.218
$\sqrt{IV^m}$	21.837%	10.996%	2.257	9.088
$\sqrt{RV^d}$	15.177%	11.429%	2.967	15.388
$\sqrt{RV^w}$	15.633%	11.121%	2.903	14.288
$\sqrt{RV^m}$	15.957%	10.635%	2.740	12.146
VRP^w	2.280%	4.659%	0.506	40.845
VRP^m	2.300%	3.107%	1.109	29.283

Table 2: In-Sample results: Implied variance

This table contains the results of regressions of the daily time series of the (annualized) weekly realized variance on a constant and the lagged forecasting variable(s). α denotes the intercept parameter. IV^w and IV^m are the Bakshi et al. (2003) weekly and monthly implied variances, respectively. The last two rows show the results from a multiple regression of the weekly realized variance on a constant, the lagged weekly implied variance and the lagged orthogonal component of the monthly implied variance (i.e., the estimated residual series from a regression of the monthly implied variance on a constant and the weekly implied variance). We present in parentheses the Newey and West (1987) t-statistics with 10 lags. Adj R^2 is the adjusted R-squared of the regression model. Returns are sampled at the 5-minute frequency.

α	IV^w	IV^m	Adj R^2
-0.005	0.713		0.649
(-1.376)	(7.240)		
-0.010		0.785	0.616
(-2.140)		(6.676)	
-0.005	0.713	0.137	0.650
(-1.338)	(7.025)	(0.631)	

Table 3: Continuous vs. jump variation

Panel A (Panel B) presents the results of regressions of the daily time series of the annualized significant weekly continuous (jump) variation on a constant and the lagged variable(s). α denotes the intercept parameter. IV^w and IV^m are the Bakshi et al. (2003) weekly and monthly implied variances, respectively. The last two rows of Panel A (Panel B) present the results from a multiple regression of the significant weekly continuous (jump) variation on a constant, the lagged weekly implied variance, and the lagged orthogonal component of the monthly implied variance. The latter corresponds to the residual series from a regression of the monthly implied variance on a constant and the weekly implied variance. We present in parentheses the Newey–West t-statistics with 10 lags. Adj R^2 is the adjusted R-squared of the regression. Returns are sampled at the 5-minute frequency.

α	IV^w	IV^m	Adj R^2
Panel A: Significant continuous variation			
-0.007	0.695		0.636
(-1.728)	(6.924)		
-0.011		0.764	0.602
(-2.424)		(6.455)	
-0.007	0.695	0.119	0.637
(-1.689)	(6.755)	(0.535)	
Panel B: Significant jump variation			
0.002	0.018		0.103
(6.280)	(3.121)		
0.001		0.021	0.109
(4.264)		(2.840)	
0.002	0.018	0.018	0.108
(6.151)	(3.031)	(1.026)	

Table 4: In-sample results: Time series models

This table contains the results of regressions of the daily time series of the (annualized) weekly realized variance on a constant and the lagged forecasting variable(s). α denotes the intercept parameter. RV^d, RV^w, and RV^m denote the (annualized) daily, weekly, and monthly realized variance series, respectively. C and J indicate the (annualized) significant continuous and (annualized) significant jump variations, respectively. The associated superscripts indicate that we compute these quantities for the daily (d), weekly (w), and monthly (m) horizons, respectively. SV^{d+} and SV^{d-} are the (annualized) positive and negative daily semivariances, respectively. SV^{w+} and SV^{w-} denote the (annualized) positive and negative weekly semivariances, respectively. SV^{m+} and SV^{m-} are the (annualized) positive and negative monthly semivariances, respectively. We present in parentheses the Newey and West (1987) t-statistics with 10 lags. Adj R^2 is the adjusted R-squared of the regression. Returns are sampled at the 5-minute frequency.

	HAR	CHAR	HAR-J	HAR-C-J	SHAR
α	0.005	0.007	0.005	0.002	0.006
	(2.592)	(3.271)	(2.816)	(0.464)	(2.620)
RV^d	0.342		0.343		
	(2.650)		(2.607)		
RV^w	0.216		0.215		
	(3.254)		(3.214)		
RV^m	0.311		0.310		
	(3.679)		(3.669)		
C^d		0.340		0.338	
		(2.603)		(2.705)	
C^w		0.229		0.219	
		(3.387)		(2.938)	
C^m		0.299		0.298	
		(3.514)		(3.483)	
J^d			-0.084	0.443	
			(-0.397)	(3.070)	
J^w				-1.911	
				(-1.255)	
J^m				5.292	
				(1.297)	
SV^{d+}					0.075
					(0.704)
SV^{d-}					0.482
					(2.851)
SV^{w+}					-0.081
					(-0.142)
SV^{w-}					0.577
					(1.085)
SV^{m+}					0.084
					(0.076)
SV^{m-}					0.546
					(0.543)
Adj R^2	0.598	0.596	0.598	0.608	0.607

Table 5: Out-of-sample results

This table presents the ratio of the average loss of each model over that of the HAR model. We use a rolling window of four years of daily data to estimate the parameters of the forecasting models. HAR denotes the forecasting model that uses the daily, weekly, and monthly lagged realized variance series to predict the weekly realized variance. IVW and IVM are the forecasting models based on the Bakshi et al. (2003) weekly and monthly implied variances, respectively. CHAR is the continuous heterogeneous autoregressive model. HAR-J extends the HAR model by including the significant lagged daily jump variation. The HAR-C-J model decomposes the realized variances in the HAR model into their significant continuous and jump components. The SHAR model extends the HAR by splitting the lagged historical variance terms into the corresponding positive and negative semivariance components. HAR-RQ takes into account the heteroskedasticity of the measurement error in all three maturities of lagged realized variance. Similar to Bollerslev et al. (2016), we proxy the heteroskedasticity of the measurement error in the realized variance with the square root of the realized quarticity of corresponding maturity. Returns are sampled at the 5-minute frequency.

	IVW	IVM	CHAR	HAR-J	HAR-C-J	SHAR	HAR-RQ
MSE	0.797	0.955	1.012	1.001	0.998	0.887	0.836
$QLIKE$	0.870	1.049	1.022	1.000	1.016	0.951	0.939

Table 6: Out-of-sample results (1-minute sampling frequency)

This table presents the ratio of the average loss of each model over that of the HAR model. We use a rolling window of four years of daily data to estimate the parameters of the forecasting models. HAR denotes the forecasting model that uses the daily, weekly, and monthly lagged realized variance series to predict the future weekly realized variance. IVW and IVM are the forecasting models based on the Bakshi et al. (2003) weekly and monthly implied variances, respectively. CHAR is the continuous heterogeneous autoregressive model. HAR-J extends the HAR model by including the significant lagged daily jump variation. The HAR-C-J model decomposes the realized variances in the HAR model into their significant continuous and jump components. The SHAR model extends the HAR by splitting the lagged historical variance terms into the corresponding positive and negative semivariance components. HAR-RQ takes into account the heteroskedasticity of the measurement error in all three maturities of lagged realized variance. Similar to Bollerslev et al. (2016), we proxy the heteroskedasticity of the measurement error in the realized variance with the square root of the realized quarticity of corresponding maturity. Returns are sampled at the 1-minute frequency.

	IVW	IVM	CHAR	HAR-J	HAR-C-J	SHAR	HAR-RQ
MSE	0.791	1.013	0.936	1.014	1.002	1.021	0.883
$QLIKE$	0.897	1.145	1.159	1.024	0.970	0.948	0.962

Table 7: Out-of-sample results (nearest)

This table presents the ratio of the average loss of each model over that of the HAR model. We use a rolling window of four years of daily data to estimate the parameters of the forecasting models. HAR denotes the forecasting model that uses the daily, weekly, and monthly lagged realized variance series to predict the future weekly realized variance. IVW and IVM are the forecasting models based on the Bakshi et al. (2003) weekly and monthly implied variances computed using options of maturities nearest to the weekly and monthly horizons, respectively. CHAR is the continuous heterogeneous autoregressive model. HAR-J extends the HAR model by including the significant lagged daily jump variation. The HAR-C-J model decomposes the realized variances in the HAR into their significant continuous and jump components. The SHAR model extends the HAR by splitting the lagged historical variance terms into the corresponding positive and negative semivariance components. HAR-RQ takes into account the heteroskedasticity of the measurement error in all three maturities of lagged realized variance. Similar to Bollerslev et al. (2016), we proxy the heteroskedasticity of the measurement error in the realized variance with the square root of the realized quarticity of corresponding maturity. Returns are sampled at the 5-minute frequency.

	IVW	IVM	CHAR	HAR-J	HAR-C-J	SHAR	HAR-RQ
MSE	0.763	0.861	1.012	1.001	0.998	0.887	0.836
$QLIKE$	0.811	0.915	1.022	1.000	1.016	0.951	0.939

Table 8: Out-of-sample results (alternative IV estimator)

This table presents the ratio of the average loss of each model over that of the HAR model. We use a rolling window of four years of daily data to estimate the parameters of the forecasting models. HAR denotes the forecasting model that uses the daily, weekly, and monthly lagged realized variance series to predict the future weekly realized variance. IVW and IVM are the forecasting models based on the Britten-Jones and Neuberger (2000) weekly and monthly implied variances, respectively. CHAR is the continuous heterogeneous autoregressive model. HAR-J extends the HAR model by including the significant lagged daily jump variation. The HAR-C-J model decomposes the realized variances in the HAR into their significant continuous and jump components. The SHAR model extends the HAR by splitting the lagged historical variance terms into the corresponding positive and negative semivariance components. HAR-RQ takes into account the heteroskedasticity of the measurement error in all three maturities of lagged realized variance. Similar to Bollerslev et al. (2016), we proxy the heteroskedasticity of the measurement error in the realized variance with the square root of the realized quarticity of corresponding maturity. Returns are sampled at the 5-minute frequency.

	IVW	IVM	CHAR	HAR-J	HAR-C-J	SHAR	HAR-RQ
MSE	0.805	1.121	1.020	1.010	0.999	0.887	0.836
$QLIKE$	0.891	1.217	1.004	0.993	1.015	0.951	0.939

Appendix to

“The information content of short-term options”

Not Intended for Publication

Will be Provided as Online Appendix

Table A.1: Descriptive statistics (extended sample)

This table presents key summary statistics based on the extended sample covering the period starting from January 1996 to August 2015. IV^w and IV^m denote the (annualized) implied variance of weekly and monthly horizons, respectively. RV^d, RV^w, and RV^m denote the (annualized) realized variance of daily, weekly, and monthly horizons, respectively. VRP^w and VRP^m are the weekly and monthly variance risk premia, respectively. Similar to Bollerslev et al. (2009), the variance risk premium of a given maturity is defined as the difference between the implied variance of that maturity and the contemporaneously computed realized variance of the same maturity. Returns data are sampled at the 5-minute frequency. Mean is the average value of the daily time series of the variable. Std, Skew, and Kurt denote the standard deviation, skewness, and kurtosis of the variable, respectively.

	Mean	Std	Skew	Kurt
$\sqrt{IV^w}$	19.871%	9.950%	2.616	14.336
$\sqrt{IV^m}$	20.951%	9.434%	2.606	14.322
$\sqrt{RV^d}$	13.728%	10.663%	4.145	27.600
$\sqrt{RV^w}$	14.776%	9.033%	2.941	15.365
$\sqrt{RV^m}$	14.982%	8.287%	3.001	16.716
VRP^w	1.938%	3.301%	2.664	23.254
VRP^m	2.348%	2.627%	2.661	13.881

Table A.2: In-Sample results: Implied variance (extended sample)

This table contains the results of regressions of the monthly time series of the (annualized) weekly realized variance on a constant and the lagged forecasting variable(s). α denotes the intercept parameter. IV^w and IV^m denote the (annualized) implied variance of weekly and monthly horizons, respectively. The last two rows show the results from a multiple regression of the weekly realized variance on a constant, the lagged weekly implied variance and the lagged orthogonalized monthly implied variance (i.e., the estimated residual series from a regression of the monthly implied variance on a constant and the weekly implied variance). We present in parentheses the Newey and West (1987) t-statistics with 10 lags. Adj R^2 is the adjusted R-squared of the regression model. Returns are sampled at the 5-minute frequency. The extended sample covers the period from January 1996 to August 2015.

α	IV^w	IV^m	Adj R^2
-0.001	0.563		0.815
(-0.396)	(21.739)		
-0.003		0.576	0.769
(-2.099)		(20.731)	
-0.001	0.563	-0.047	0.814
(-0.391)	(21.340)	(-0.226)	

Table A.3: In-sample results: Time series models (extended sample)

This table shows the results of regressions of the monthly time series of the (annualized) weekly realized variance on a constant and the lagged forecasting variable(s). α denotes the intercept parameter. RV^d, RV^w, and RV^m denote the (annualized) daily, weekly, and monthly realized variance series, respectively. C and J indicate the (annualized) significant continuous and (annualized) significant jump variations, respectively. The associated superscripts indicate that we compute these quantities for the daily (d), weekly (w), and monthly (m) horizons, respectively. SV^{d+} and SV^{d-} are the (annualized) positive and negative daily semivariances, respectively. SV^{w+} and SV^{w-} denote the (annualized) positive and negative weekly semivariances, respectively. SV^{m+} and SV^{m-} are the (annualized) positive and negative monthly semivariances, respectively. We present in parentheses the Newey and West (1987) t-statistics with 10 lags. Adj R^2 is the adjusted R-squared of the regression. Returns are sampled at the 5-minute frequency. The extended sample covers the period from January 1996 to August 2015.

	HAR	CHAR	HAR-J	HAR-C-J	SHAR
α	0.010	0.010	0.010	0.009	0.009
	(3.914)	(4.000)	(3.869)	(3.503)	(4.221)
RV^d	0.291		0.291		
	(3.938)		(3.935)		
RV^w	0.013		0.012		
	(0.101)		(0.098)		
RV^m	0.288		0.289		
	(2.516)		(2.529)		
C^d		0.251		0.278	
		(2.785)		(3.131)	
C^w		0.124		0.110	
		(0.623)		(0.568)	
C^m		0.236		0.198	
		(1.915)		(1.615)	
J^d			-0.283	0.316	
			(-0.620)	(0.689)	
J^w				-0.620	
				(-2.130)	
J^m				1.701	
				(1.905)	
SV^{d+}					-0.050
					(-0.470)
SV^{d-}					0.408
					(4.049)
SV^{w+}					0.072
					(0.263)
SV^{w-}					0.358
					(1.187)
SV^{m+}					0.287
					(0.427)
SV^{m-}					0.156
					(0.211)
Adj R^2	0.757	0.757	0.756	0.764	0.778

Table A.4: Out-of-sample results (extended sample)

This table shows the ratio of the average loss of each model over that of the HAR model. We use a rolling window of four years of monthly data to estimate the parameters of the forecasting models. HAR denotes the forecasting model that uses the daily, weekly, and monthly lagged realized variance series to predict the future weekly realized variance. IVW and IVM are the forecasting models based on the Bakshi et al. (2003) weekly and monthly implied variances, respectively. CHAR is the continuous heterogeneous autoregressive model. HAR-J extends the HAR model by including the significant lagged daily jump variation. The HAR-C-J model decomposes the realized variances in the HAR into their significant continuous and jump components. The SHAR model extends the HAR by splitting the lagged historical variance terms into the corresponding positive and negative semivariance components. HAR-RQ takes into account the heteroskedasticity of the measurement error in all three maturities of lagged realized variance. Similar to Bollerslev et al. (2016), we proxy the heteroskedasticity of the measurement error in the realized variance with the square root of the realized quarticity of corresponding maturity. Returns are sampled at the 5-minute frequency. The extended sample covers the period from January 1996 to August 2015.

	IVW	IVM	CHAR	HAR-J	HAR-C-J	SHAR	HAR-RQ
MSE	0.885	0.941	1.012	1.013	1.085	1.059	1.158
$QLIKE$	0.841	0.940	1.023	1.162	1.531	1.042	1.226

Table A.5: In-sample results: Implied variance (extended sample and quarterly horizon)

This table shows the results of regressions of the non-overlapping time series of the (annualized) quarterly realized variance on a constant and the lagged forecasting variable(s). α denotes the intercept parameter. IV^q and IV^m are the Bakshi et al. (2003) quarterly and monthly implied variances, respectively. We present in parentheses the Newey and West (1987) t-statistics with 3 lags. Adj R^2 is the adjusted R-squared of the regression model. Returns are sampled at the 5-minute frequency. The first two sets of regressions are univariate. In the third model we regress the quarterly realized variance on a constant, the lagged quarterly implied variance and the lagged orthogonal component of the monthly implied variance with respect to the lagged quarterly implied variance. To be more specific, the orthogonal monthly implied variance is the estimated residual from the regression of the monthly implied variance on a constant and the contemporaneous quarterly implied variance. The extended sample covers the period from January 1996 to August 2015.

α	IV^q	IV^m	Adj R^2
0.013	0.289		0.178
(2.553)	(12.327)		
0.016		0.245	0.172
(2.944)		(9.818)	
0.013	0.289	-0.026	0.167
(2.572)	(12.504)	(-0.070)	

Highlights

- We employ weekly options on the S&P 500 index to compute the weekly implied variance
- Weekly implied variance is a strong predictor of weekly realized variance
- Weekly implied variance crowds out the information content of the monthly implied variance
- Weekly implied variance outperforms time series models of realized variance
- The results hold both in- and out-of-sample and the forecast gains are significant