Information Gains from using Short-Dated Options for Measuring and Forecasting Volatility∗

Viktor Todorov† and Yang Zhang‡

First draft: November 19, 2019, This draft: September 23, 2020

Abstract

We study the gains from using short-dated options for volatility measurement and forecasting. Using option portfolios, we estimate nonparametrically spot volatility under weak assumptions for the underlying asset. This volatility estimator complements existing ones constructed from high-frequency returns. We show empirically, using the market index and Dow 30 stocks, that combining optimally return and option data can lead to nontrivial gains for volatility forecasting. These gains are due to “diversification” of the measurement error in the two volatility proxies. The information content of short-dated options, not spanned by the current spot volatility, is of limited relevance for volatility forecasting.

Keywords: high-frequency data, nonparametric volatility estimation, options, return predictability, volatility forecasting.

JEL classification: C51, C52, G12.

∗Research partially supported by NSF grant SES-1530748.
†Kellogg School of Management, Northwestern University; e-mail: v-todorov@northwestern.edu.
‡Evanston, IL 60201; e-mail: yang.zhang1@kellogg.northwestern.edu. Yang Zhang completed the majority of this work as a postdoc in the Finance Department, Kellogg School of Management, Northwestern University, and the rest on her personal time. The views expressed in the article are the author’s own and not those of her current employer.
1 Introduction

The availability of reliable high-frequency financial data allows for efficient nonparametric measurement of volatility. In the last two decades, a number of alternative jump-robust high-frequency volatility estimators have been developed. These include the multipower variations of Barndorff-Nielsen and Shephard (2004, 2006), the truncated variation of Mancini (2001, 2009) as well as estimators based on the empirical characteristic function developed by Jacod and Todorov (2014, 2018). The improved precision in measuring volatility, offered by the high-frequency data, leads to nontrivial gains in the estimation of stochastic volatility models, see e.g., Barndorff-Nielsen and Shephard (2002), Bollerslev and Zhou (2002), Corradi and Distaso (2006), Takahashi et al. (2009), Todorov (2009), Dobrev and Szerszen (2010) and Koopman and Schart (2013), the nonparametric analysis of the volatility distribution, see e.g., Todorov and Tauchen (2012) and Christensen et al. (2019), as well as in volatility forecasting and management, see e.g., Andersen et al. (2003), Corsi (2009), Shephard and Sheppard (2010), Hansen et al. (2012), Bollerslev et al. (2016) and Buccheri and Corsi (2020), among many others.

Parallel to the increased availability of reliable high-frequency return data, there has been also a recent sharp increase in the trading of options, particularly those written on market indices and with short time-to-maturity, see e.g., Andersen et al. (2017). The goal of the current paper is to quantify empirically the incremental gains from adding option observations to high-frequency return data when measuring and forecasting volatility. On one hand, the differences in the information content of options and stock returns make unclear whether options and option portfolios can help volatility forecasts based solely on returns. On the other hand, if (true) volatility recovery is possible from options, this should lead unambiguously to better measurement and forecasting of volatility because of “diversification” in the measurement error of return- and option-based volatility estimators. In this paper, we show empirically that option data offers nontrivial gains for the purposes of volatility forecasting mostly because it allows for more precise measurement of spot volatility. Additional forecasting gains due to the different information content of short-dated options are small.

Given the nonlinear payoff of options, their prices naturally depend on the volatility of the underlying asset. Indeed, in the classical Black-Scholes volatility model (Black and Scholes (1973)), there is a one-to-one map between the option price normalized by the current stock price and the (constant) volatility of the underlying asset. As a result, the volatility parameter of the Black-Scholes model can be backed out from the price of a single option written on the underlying

\footnote{The original estimators have been developed under the assumption of no microstructure noise in the observed price. Many of them have been extended to settings with microstructure noise, see e.g. Jacod and Protter (2011).}
asset, with the resulting option-based volatility estimator being referred to as Black-Scholes implied volatility (abbreviated henceforth as BSIV). The assumptions behind the Black-Scholes model are known to be violated in practice due to the time-variation in volatility, the presence of jumps in asset prices and the risk premia demanded by investors for bearing these risks. This leads to a well-established empirical pattern where BSIV is significantly higher than “historical” volatility measured from past returns, and this pattern is particularly pronounced for BSIV of out-of-the-money puts written on the market index, see e.g., Bates (2000) and references therein.

An alternative to the BSIV, often used in volatility forecasting, is a model-free measure of risk-neutral expected integrated quadratic variation known as the VIX volatility index. It is constructed from a portfolio of out-of-the-money options with different strikes and the same time-to-maturity, see e.g., Britten-Jones and Neuberger (2000) and Carr and Wu (2008). While the construction of the VIX index, unlike BSIV, does not rely on a model for the underlying asset dynamics, the VIX nevertheless differs substantially from return-based volatility measures. The reason for this is that the VIX index is a conditional expectation of return variance over a period in the future, and further the expectation is under the so-called risk-neutral probability measure. Therefore, the mean-reversion in volatility and the presence of nontrivial volatility and jump risk premia create a wedge between true (spot) volatility and the VIX index. Given that the risk premium component of the VIX index is rather nontrivial, see e.g., Carr and Wu (2008) and Bollerslev et al. (2009), the latter (just like BSIV) is on average significantly above a “historical” measure of volatility formed from past returns and this wedge is time-varying.

The increased availability and trading of options with very short tenor, however, allow for developing nonparametric volatility measures from options that are robust to the time variation in volatility and to the presence of jumps (and their pricing) in asset prices. Todorov (2019) develops such an estimator (henceforth referred to as OV) which is nearly rate-efficient and is based on an option portfolio of short-dated options that “spans”, i.e., estimates nonparametrically, the risk-neutral conditional characteristic function of the price increment over the time span of the options. The option portfolio behind this nonparametric volatility estimator is formed from long and short positions in options whose strikes are in the vicinity of the current stock price, with the magnitude of the weights gradually declining as the strikes of the options move further away from the current stock price. This portfolio puts significantly less weight to deep out-of-the-money options than the one used in the construction of the VIX index. The option-based OV estimator, when combined with one formed from high-frequency returns, can provide efficiency gains in the measurement of volatility. Indeed, the Central Limit Theorem (CLT) for the return-based volatility estimators is
governed by the martingale component of the price over the interval over which the estimator is computed while the CLT for the option-based volatility estimator is driven by the measurement error in observed option prices at the point in time they are recorded (typically at market close). Therefore, asymptotically, the return- and option-based volatility estimators are uncorrelated and by combining them, we can gain precision in measuring volatility. This, in turn, should lead to gains in volatility measurement and forecasting.

We quantify empirically the information gains from adding an option-based volatility measure to one formed from high-frequency returns using data for the S&P 500 market index and 28 stocks in the Dow Jones Industrial Average index (Dow 30) with good option coverage. The sample period for the S&P 500 index is 2008-2018 and that for the individual stocks is 2010-2019. The option-based volatility estimator for the market index and a truncated variation estimator formed from high-frequency price record of S&P 500 index futures are on average very close. Nevertheless, the option-based estimator is far less noisy than the return-based one. This increased precision in measuring spot diffusive volatility from options leads to a nonparametric estimate of its marginal distribution which is significantly more concentrated around its mean compared to one recovered from returns via the nonparametric deconvolution method of Todorov and Tauchen (2012). For the individual stocks, option measures provide less reduction in the volatility measurement error. This is not surprising given the lower number of available individual stock options.

We next assess the gains offered from the short-dated options for the purposes of volatility forecasting. Earlier work that compares option and return-based volatility measures, see e.g., Fleming (1998), Christensen and Prabhala (1998), Blair et al. (2001), Jiang and Tian (2005), Koopman et al. (2005), Neely (2009), Martin et al. (2009), Busch et al. (2011), Bekaert and Hoerova (2014), Kambouroudis et al. (2016) and Oikonomou et al. (2019) as well as the many references in the review articles of Poon and Granger (2003) and Christoffersen et al. (2013), uses either BSIV or the VIX index. These measures, however, are biased estimates of volatility, with the bias due to mean-reversion in volatility and the presence of volatility and jump risk premia. The bias in the BSIV and VIX option measures can lead to their information inefficiencies as pointed by Chernov (2007). Therefore, the comparison between the forecasting performance of return-based volatility estimators on one hand and the BSIV and the VIX index on the other hand will depend not only on the measurement error in these estimates but also on the forecasting ability of the wedge between the BSIV or the VIX index and the true volatility. Perhaps in part due to this, the existing evidence for the benefits from using option data for volatility forecasting is mixed. Some of the above-cited studies have found gains from using BSIV or VIX for volatility forecasting.
(such as Blair et al. (2001), Busch et al. (2011), Bekaert and Hoerova (2014), Kambouroudis et al. (2016) and Oikonomou et al. (2019)) while others have found no such gains (such as Koopman et al. (2005), Neely (2009) and Martin et al. (2009)).

As already discussed above, the OV estimator, unlike BSIV and VIX, is an option-based estimator of the true spot volatility. Therefore, at least theoretically, it should always provide gains in volatility forecasting when used together with return-based volatility proxies. This is because the measurement errors in the two volatility proxies are uncorrelated. As a result, regardless of the forecasting model, replacing the return-based volatility proxy with a linear combination of the return- and option-based volatility proxies should always provide forecasting gains. We illustrate this using the heteroskedastic autoregressive (HAR) model of Corsi (2009), which is a constrained autoregressive model of high order aimed at mimicking in a parsimonious way long-memory features of volatility. Our benchmark model is a HAR model in which we use solely truncated variance to predict future volatility. We compare this model with one in which the truncated volatility is replaced with the option-based volatility estimator and mixture models in which both volatility proxies are used as predictors or are combined in a single one based on optimal weighting according to estimates of their asymptotic variances. Our empirical evidence confirms the theoretical prediction of the superiority of the mixture models. They provide around 30% reduction in the time series median of the daily forecasting loss for the market and around 5% reduction in that for the individual stocks. Combining option and return data continues to offer advantages when the forecasting horizon is one week or one month, with the size of the gains being statistically significant and larger than the daily ones in the case of individual stocks. Moreover, the above-reported gains become much bigger when using a less noisy proxy of future volatility.

We finally study whether the additional information in the short-dated options that is not contained in the spot (diffusive) volatility can provide any gains in volatility forecasting. We answer this question by developing a test for independence of the future realized volatility and alternative option measures, conditional on the values of the option-based volatility. The alternative option measures are estimates of the jump variation, which together with the spot volatility are the two state variables that determine uniquely the short-dated options. Our results show that the risk-neutral jump variation does not contain statistically significant additional information, relative to

---

2Recent work by Buccheri and Corsi (2020) shows that the simple linear HAR model contains misspecification due to nonlinear volatility dependencies. While we use the HAR model (like most of the existing volatility forecasting literature) to illustrate the gains from combining option and return data, these gains are obviously not specific to our HAR illustration. Indeed, we performed the HAR volatility forecasting regressions in logs and we found similar gains to the ones reported in the paper from combining return- and option-based volatility proxies. These results are not reported in the paper in order to save space but are available upon request.
the one contained in spot volatility, that is relevant for forecasting future volatility. Any potential gains from including such measures for volatility forecasting are mostly due to reduction in the measurement error of the option-based quantities. Accordingly, upon augmenting the original forecasting models with measures of risk-neutral jump variation we find relatively small additional improvements in forecasting performance.

The rest of the paper is organized as follows. In Section 2, we introduce our return- and option-based volatility estimates. In Section 3 we illustrate theoretically the gains from combining alternative volatility estimates for the purposes of volatility forecasting. Section 4 compares empirically the performance of various market volatility forecasting models that use return and option volatility proxies and this analysis is repeated on an individual stock level in Section 5. In Section 6 we investigate whether short-dated options contain more information relevant for volatility forecasting in addition to that contained in the spot volatility and study the impact of improved volatility measurement for studying return predictability. Section 7 concludes. The Appendix contains additional details on the data and the construction of the volatility measures and tests.

2 Nonparametric Measurement of Volatility

We begin with introducing our nonparametric measures of volatility. The generic log-asset price is denoted with \( x_t \). It is defined on a filtered probability space, \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\), and is assumed to obey the following Itô semimartingale dynamics

\[
x_t = x_0 + \int_0^t \alpha_s ds + \int_0^t \sqrt{V_s} dW_s + \sum_{s \leq t} \Delta x_s,
\]

(1)

where \( \alpha_t \) is the drift term, \( V_t \) is the stochastic variance and \( W_t \) denotes a Brownian motion. We can measure nonparametrically \( V_t \) from high-frequency return observations or from short-dated options written on the asset. We introduce these estimators in the next two subsections.

2.1 Return-Based Measures

We start with the return-based volatility estimators. We assume that we sample the asset price at equidistant times during the trading hours of a business day. Our unit of time for defining the high-frequency measures is a business day. In particular, each unit interval \([t - 1, t]\) consists of an overnight period \([t - 1, t - 1 + \kappa]\) and a trading period \([t - 1 + \kappa, t]\), for \( t \in \mathbb{N}_+ \) and \( \kappa \in (0, 1) \). We have \( n + 1 \) equidistant observations during the trading hours. We define the high-frequency returns
as \( \Delta^{n}_{t,i} x = x_{t-1+(1-\kappa)i/n} - x_{t-1+(1-\kappa)(i-1)/n}, \) for \( t \in \mathbb{N}_+ \) and \( i = 1, \ldots, n. \) Our first high-frequency volatility estimator is the intraday realized volatility defined as

\[
RV_t = \sum_{i=1}^{n} (\Delta^{n}_{t,i} x)^2, \quad t \in \mathbb{N}_+.
\] (2)

\( RV_t \) is a consistent estimate of the integrated quadratic variation over the trading period, \( \int_{t-1+\kappa}^{t} V_s ds + \sum_{s \in [t-1+\kappa,t]} (\Delta x_s)^2. \) Next, we introduce the truncated volatility of Mancini (2009) which separates the volatility from jumps and is computed by

\[
TV_t = \sum_{i=1}^{n} (\Delta^{n}_{t,i} x)^2 1\left( |\Delta^{n}_{t,i} x| \leq 3 \sqrt{BV_t \wedge RV_t \times n^{-1/2}} \right), \quad t \in \mathbb{N}_+,
\] (3)

with \( BV_t = \pi \frac{2}{n} \sum_{i=2}^{n} \sqrt{|\Delta^{n}_{t,i-1} x||\Delta^{n}_{t,i} x|} \) being the bi-power variation of Barndorff-Nielsen and Shephard (2004). \( BV_t \) is a consistent and tuning-free estimator of the integrated volatility \( \int_{t-1+\kappa}^{t} V_s ds. \) We use it to determine the truncation level in \( TV_t, \) which is an efficient estimator of the integrated volatility. If we consider estimation in which the length of the time interval is asymptotically shrinking, as we sample more frequently, then this will provide an estimator of the spot volatility (as our option measure will do). In our estimation, as most of the volatility forecasting literature, we will use daily integrated volatility.\(^3\) It can be considered as an estimator of spot volatility at market close in the asymptotic sense described above. The overnight period is non-trivial for the individual stocks while it is negligible for the S&P 500 market index as the futures written on it, which we will use in our analysis, trade for nearly 24 hours (except during the weekend).

### 2.2 Option-Based Measures

We continue with introducing our option-based measure of \( V_t. \) This measure is constructed from the prices of out-of-the-money (OTM) options written on the underlying asset in the following way. First, for a given point in time \( t, \) we denote with \( O_{t,\tau}(k) \) the theoretical price of an OTM option price written on the underlying asset which expires at time \( t + \tau \) and has a strike \( K = e^k \) (hence \( k \) is the log-strike). The OTM option would be worth zero if it were to expire today. It is a put if \( k \leq \log(F_{t,\tau}) \) and it is a call if \( k > \log(F_{t,\tau}) \), where \( F_{t,\tau} \) denotes the futures price written on the underlying asset at time \( t \) and which expires at \( t + \tau \). The available option prices are observed with error and are denoted with \( \hat{O}_{t,\tau}(k). \) The conditional characteristic function of the price increment

---

\(^3\)We also experimented with spot volatility estimators constructed from using only returns in a local window around market close, but the forecasting results were worse than the ones we report below for the daily \( TV. \)
under the risk-neutral measure $Q$, denoted by $L_{t,\tau}(u) = \mathbb{E}_t^Q(e^{iu(x_{t+\tau}-x_t)})$, can be spanned by a portfolio of options in the following way

$$L_{t,\tau}(u) = 1 - (u^2 + iu) \int_{\mathbb{R}} e^{(iu-1) k - iux_t} O_{t,\tau}(k) dk, \quad u \in \mathbb{R},$$  \hspace{1cm} (4)$$

where in the above we have set the risk-free interest rate to zero (this is inconsequential as we apply this formula only for $\tau$ very close to zero). We denote the feasible counterpart of $L_{t,\tau}(u)$ by $\hat{L}_{t,\tau}(u)$, which is formed by a Riemann sum approximation of the integral in (4) on the basis of the available options together with linear interpolation and extrapolation (in $BSIV$ space) to fill-in missing option observations for strikes within the available strike range and/or to extrapolate option prices outside the available strike range in certain cases. The details of these calculations are given in the Appendix.

Following [Todorov (2019)], the option estimator of volatility is given by

$$\hat{\sigma}_{t,\tau} = \frac{2}{\tau \hat{u}^2_{t,\tau}} \log |\hat{L}_{t,\tau}(\hat{u}_{t,\tau})|,$$  \hspace{1cm} (5)$$

where we set $\hat{u}_{t,\tau} = \hat{u}^{(1)}_{t,\tau} \land \hat{u}^{(2)}_{t,\tau}$ with

$$\hat{u}^{(1)}_{t,\tau} = \inf \left\{ u \geq 0 : |\hat{L}_{t,\tau}(u)| \leq 0.2 \right\}, \quad \hat{u}_{t,\tau} = \sqrt{\frac{2 \log(0.05)}{\tau \hat{\sigma}^2_{t,ATM}}}, \quad \hat{u}^{(2)}_{t,\tau} = \arg\min_{u \in [0,\pi]} |\hat{L}_{t,\tau}(u)|,$$  \hspace{1cm} (6)$$

for $\hat{\sigma}_{t,ATM}$ denoting the Black-Scholes implied volatility of the option with strike closest to the current stock price at time $t$. The above choice of $u$ aims to evaluate $\hat{L}_{t,\tau}(u)$ at a high value of its argument for which the inference for $\hat{L}_{t,\tau}(u)$ is still reliable. The need to use $u$ as high as possible is because for high values of $u$, jumps play only negligible role in $|L_{t,\tau}(u)|$. The choice of $u$ can be viewed as the counterpart of the choice of truncation level when computing the return-based TV$t$. More specifically, [Todorov (2019)] shows that

$$\hat{L}_{t,\tau}(u/\sqrt{\tau}) \approx e^{-u^2/2V_t}, \quad \text{as } \tau \downarrow 0 \text{ and for fixed } u > 0.$$  \hspace{1cm} (7)$$

Intuitively, the short maturity of the options minimizes the impact on the above option portfolio from time-varying volatility while the use of characteristic exponent $u$ strictly away from zero provides robustness against jumps. As a result $\hat{V}_{t,\tau}(\hat{u}_{t,\tau})$ provides an estimate of true spot diffusive volatility despite of the fact that it is computed from options which are risk-neutral expectations.
of their payoffs. We refer to Todorov (2019) for the formal statement of the above result.

In Figure 1, we illustrate the decomposition of the estimate of \( (u^2 + iu) \int_\mathbb{R} e^{(iu-1)k-iux}O_{t,\tau}(k)dk \), from which our volatility estimator is constructed. The example is based on options written on the S&P 500 index with time-to-maturity of 4 business days. As seen from the middle panel of the figure, the value of the real part of \( \hat{L}_{t,\tau}(\hat{u}_{t,\tau}) \) is formed from options whose strike is very close to the current price. Indeed, options with strikes more than 2 standard deviations from the current stock price receive almost zero weight. This is to be expected as there is very little signal in the prices of deep OTM options about the diffusive spot volatility. We also note that the portfolio of options that generates the real part of \( \hat{L}_{t,\tau}(\hat{u}_{t,\tau}) \) involves both long and short positions in options which are close to the money. The total value of the short position is relatively small and it effectively serves as a way to remove the contribution of jumps in the option prices. Turning to the imaginary part of \( \hat{L}_{t,\tau}(\hat{u}_{t,\tau}) \), we can see from the right panel of Figure 1 that the portfolio that generates it goes long near the money calls and short near the money puts. Since the option prices around the money are nearly symmetric, the total value of the imaginary part of \( \hat{L}_{t,\tau}(\hat{u}_{t,\tau}) \) is close to zero. This is in line with the approximation in (7) which is behind the option-based volatility estimator.

Since on a given day we can have several short maturities, we average the volatility estimates from the two shortest available ones to arrive at the following option-based volatility estimator

\[
OV_t = \frac{1}{2} \left( \hat{V}_{t,\tau_1}(\hat{u}_{t,\tau_1}) + \hat{V}_{t,\tau_2}(\hat{u}_{t,\tau_2}) \right). \tag{8}
\]

We finally note that for individual stock options, pre-scheduled events such as earning and dividend announcements can introduce non-trivial upward bias in the option-based volatility measures as these jump events, with known arrival times, can generate nontrivial volatility. Therefore, whenever an announcement is very close (but before) the expiration date of the options, we difference the volatility estimates constructed from the options with the two different maturity dates in order to annihilate the effect of the announcement jump. Details on this are provided in the Appendix.

2.3 Empirical Evidence

We compute the return- and option-based volatility measures for the S&P 500 market index and individual stocks in the Dow Jones Industrial Average (DJIA). Our sample for the S&P 500 index covers the period from the beginning of January 2008 till the end of December 2018, with a total

\footnote{Recall that no-arbitrage implies that the diffusion coefficient of \( X \) is the same under \( \mathbb{P} \) and \( \mathbb{Q} \).}
Figure 1: Characteristic Function SPX Option Portfolio on February 1, 2016. The left panel plots observed OTM option prices against their strikes. The next two panels plot their contributions to the estimate of $(u^2 + iu)\int_{R} e^{iu(k-1)} O_{t,\tau}(k) dk$. The maturity of the options is 4 trading days and the forward price level on the date was 1938.3.

of 2758 trading days. We consider 28 individual stocks in the DJIA with good option coverage. Due to the limited availability of short-dated option data on individual stocks in 2008 and 2009, our sample for individual stocks starts at the beginning of January 2010 and ends at the end of December 2019 with a total of 2494 trading days on average.\(^5\) We obtain option mid-quotes at market close from OptionMetrics. On each day, we keep up to two maturities, with the shortest being at least 2 trading days for the S&P 500 index options and at least 3 trading days for the single-name options.

We use E-mini S&P 500 futures for computing the return-based market volatility measures. The high-frequency data for the E-mini S&P 500 futures and individual stocks are obtained from TickData and TAQ, respectively. We use five minute sampling frequency in order to minimize the potential effect from presence of microstructure noise in observed prices. This is particularly relevant for individual stock data.\(^6\) Additional details on the processing of the raw data and the computation of the volatility measures are provided in the Appendix.

\(^5\)Due to different availability of short-dated options for individual stocks during the sample, the dates of available option-based volatility estimates differ slightly across stocks.

\(^6\)We also experimented with realized market volatility measures based on one minute data, with the results being qualitatively the same as the ones reported in the paper that are based on five minute data.
In Figure 2, we plot the return- and option-based volatility measures for the S&P 500 index and in Panel A of Table 1 we report summary statistics for them. The $OV$ and $TV$ measures are in general very close, which manifests in very similar time-series medians for the two series. Nevertheless, we can see that the mean of $TV$ exceeds that of $OV$. A potential reason for this is that the separation of volatility from jumps from high-frequency returns is more difficult in periods of high volatility where one needs to use a higher threshold. We further note that in the early part of the sample until around the end of 2011, the maturity of the available options is longer than for the rest of the sample, see the Appendix for the details. For longer-dated options the effect of the volatility and jump risk premium is larger and this might generate some upward bias in the resulting option-based volatility measures.

From Figure 2 and Panel A of Table 1, we can observe that the option-based $OV$ volatility estimator is significantly less noisy than its return-based counterpart. In particular, the volatility of the first-difference of $OV$ is nearly one-fourth that of $TV$. This results in $OV$ series which is more persistent than $TV$. By comparing the medians of $OV$ with $RV$, we can conclude that jump risk contributes a nontrivial 30% in the total return variation. Finally, in Table 1 we also report summary statistics for the two popular option measures used in prior work, mainly the at-the-money $BSIV$ and the $VIX$ index. As seen from the reported results, the two measures are significantly upward biased measures of volatility, compared to the return-based $TV$ and our option-based $OV$ estimator. Not surprisingly, the bias in $VIX$ is higher than that in $BSIV$ as the former uses out-of-the-money options which have higher implied volatility. The bias in $BSIV$ and $VIX$ is rather nontrivial and hence they cannot be used directly for studying the volatility without removing the (jump) risk premium component in them.

The significant improvement in precision offered by the option data for the measurement of spot volatility should lead to nontrivial gains in the parametric and nonparametric study of the volatility and its dynamics. Since volatility is not directly observable from a discrete price record of the underlying asset, inference for it from returns is rather nontrivial. There is a large body of work that deals with this problem. The short-dated option data and the nonparametric method of Section 2.2 makes the spot volatility effectively observable and this should simplify significantly the inference for it. We illustrate this with the nonparametric estimation of the marginal distribution of $V_t$. Using $OV$, we can estimate this distribution by a standard kernel density estimator. The estimation of the marginal volatility distribution from the high-frequency returns, on the other hand, is more involved. We can use the daily $TV$ measure for this, but since $TV$ is estimated

---

7 We computed the $VIX$ index from our data following the CBOE white paper.
Table 1: Summary Statistics of Daily Return- and Option-based Variance Measures

Panel A: S&P 500 Index

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>STD</th>
<th>STD(∆V)</th>
<th>Q25</th>
<th>Median</th>
<th>Q75</th>
<th>AR(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RV</td>
<td>0.0396</td>
<td>0.0965</td>
<td>0.0724</td>
<td>0.0083</td>
<td>0.0161</td>
<td>0.0354</td>
<td>0.7188</td>
</tr>
<tr>
<td>TV</td>
<td>0.0281</td>
<td>0.0673</td>
<td>0.0458</td>
<td>0.0057</td>
<td>0.0113</td>
<td>0.0248</td>
<td>0.7693</td>
</tr>
<tr>
<td>OV</td>
<td>0.0222</td>
<td>0.0351</td>
<td>0.0121</td>
<td>0.0058</td>
<td>0.0108</td>
<td>0.0233</td>
<td>0.9404</td>
</tr>
<tr>
<td>BSIV</td>
<td>0.0363</td>
<td>0.0546</td>
<td>0.0185</td>
<td>0.0100</td>
<td>0.0188</td>
<td>0.0397</td>
<td>0.9427</td>
</tr>
<tr>
<td>VIX</td>
<td>0.0480</td>
<td>0.0714</td>
<td>0.0238</td>
<td>0.0140</td>
<td>0.0252</td>
<td>0.0501</td>
<td>0.9443</td>
</tr>
</tbody>
</table>

Panel B: Individual Stocks

<table>
<thead>
<tr>
<th>Ticker</th>
<th>#K</th>
<th>RV</th>
<th>TV</th>
<th>OV</th>
<th>BSIV</th>
<th>VIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPL</td>
<td>35</td>
<td>0.0394</td>
<td>0.0323</td>
<td>0.0393</td>
<td>0.0530</td>
<td>0.0638</td>
</tr>
<tr>
<td>AXP</td>
<td>19</td>
<td>0.0264</td>
<td>0.0212</td>
<td>0.0255</td>
<td>0.0379</td>
<td>0.0457</td>
</tr>
<tr>
<td>BA</td>
<td>22</td>
<td>0.0404</td>
<td>0.0310</td>
<td>0.0325</td>
<td>0.0464</td>
<td>0.0549</td>
</tr>
<tr>
<td>CAT</td>
<td>22</td>
<td>0.0526</td>
<td>0.0403</td>
<td>0.0411</td>
<td>0.0574</td>
<td>0.0662</td>
</tr>
<tr>
<td>CSCO</td>
<td>11</td>
<td>0.0483</td>
<td>0.0386</td>
<td>0.0286</td>
<td>0.0458</td>
<td>0.0531</td>
</tr>
<tr>
<td>CVX</td>
<td>19</td>
<td>0.0301</td>
<td>0.0249</td>
<td>0.0229</td>
<td>0.0344</td>
<td>0.0412</td>
</tr>
<tr>
<td>DIS</td>
<td>20</td>
<td>0.0310</td>
<td>0.0252</td>
<td>0.0253</td>
<td>0.0372</td>
<td>0.0458</td>
</tr>
<tr>
<td>GE</td>
<td>10</td>
<td>0.0411</td>
<td>0.0333</td>
<td>0.0296</td>
<td>0.0476</td>
<td>0.0594</td>
</tr>
<tr>
<td>GS</td>
<td>18</td>
<td>0.0401</td>
<td>0.0314</td>
<td>0.0377</td>
<td>0.0516</td>
<td>0.0599</td>
</tr>
<tr>
<td>HD</td>
<td>17</td>
<td>0.0281</td>
<td>0.0220</td>
<td>0.0232</td>
<td>0.0337</td>
<td>0.0403</td>
</tr>
<tr>
<td>IBM</td>
<td>16</td>
<td>0.0259</td>
<td>0.0207</td>
<td>0.0185</td>
<td>0.0267</td>
<td>0.0327</td>
</tr>
<tr>
<td>INTC</td>
<td>12</td>
<td>0.0465</td>
<td>0.0372</td>
<td>0.0346</td>
<td>0.0500</td>
<td>0.0583</td>
</tr>
<tr>
<td>JNJ</td>
<td>16</td>
<td>0.0159</td>
<td>0.0126</td>
<td>0.0119</td>
<td>0.0177</td>
<td>0.0224</td>
</tr>
<tr>
<td>JPM</td>
<td>20</td>
<td>0.0368</td>
<td>0.0294</td>
<td>0.0283</td>
<td>0.0417</td>
<td>0.0494</td>
</tr>
</tbody>
</table>

Notes: The table reports summary statistics of return- and option-based variance measures for S&P 500 Index from January 2008 to December 2018 (Panel A) and 28 individual stocks in the DJIA from January 2010 to December 2019 (Panel B). For the S&P 500 Index, we report the time-series mean (Mean), standard deviation (STD), and quantiles (Q25, Median, Q75) for each measure, together with the standard deviation of the first-order difference (STD(∆V)) and the first-order autocorrelation coefficient (AR(1)). For individual stocks, we report the time-series medians for each measure together with the median number of OTM strikes across the sample used for computing the option measures (second column). All variation measures are reported in annual variance units. The RV and TV series have been annualized using overnight adjustment factor based on the average variance of open-to-close versus close-to-open returns.

over one day over which spot volatility can change, a nonparametric kernel density estimator of TV will provide an estimate of the distribution of the daily integrated volatility instead of that of the spot volatility. Therefore, we can either shorten the window over which TV is computed as in Christensen et al. (2019) or compute the empirical characteristic function of the high-frequency returns and perform a regularized inversion as done in Todorov and Tauchen (2012). We do the latter here, and we note that this approach has the advantage of “built-in” robustness towards jumps, i.e., unlike the computation of TV it does not require explicit truncation of the returns, and the associated choice of a tuning parameter for it, to filter out the jumps. In the recovery of the density of the marginal law of $V_t$ from the returns, we set the regularization parameter as
high as possible to achieve better precision but so that the recovered density has a small number of violations of quasiconcavity, see Todorov and Tauchen (2012) for further details.

In Figure 3, we compare the density estimates of the spot volatility distribution from the high-frequency returns and the options. As seen from the figure, the two estimated densities have very similar modes. However, the return-based estimate imply a volatility density that is far more dispersed around its mode than the option-based one. The comparison between the two estimated volatility densities suggests that the primary reason for this is the lower precision in recovering volatility from returns. This is particularly true for the instances with high volatility realizations for which separation of volatility from jumps from returns is difficult. Overall, Figure 3 suggests volatility of volatility is far smaller than implied from return-based volatility estimates. We leave further analysis of the gains from using OV for studying the volatility dynamics for future work.

We turn next to the individual stock volatility estimates. In Panel B of Table 1 we provide summary statistics for the volatility measures of each of the individual stocks. Here, unlike the S&P 500 index case, the overnight period where there is no high-frequency record of the asset price, is rather nontrivial. Indeed, the variance of the overnight return is around half (or more) of the variance of the intraday return. Nevertheless, one can form a simple nonparametric estimate of $\kappa$ (the length of the overnight period) by simply comparing the average return variance during the overnight and intraday periods. After scaling up TV using such an estimate of $\kappa$, we can compare the two volatility proxies constructed from stock returns and options. As seen from Table 1, the levels of the two volatility measures (after scaling up appropriately TV to account for the overnight
Figure 3: **Spot Volatility Density Estimates for S&P 500 Index.** The solid line corresponds to kernel density estimate of annualized $\sqrt{OV_t}$, where we use Epanechnikov kernel and the Silverman’s automatic bandwidth of $h = 0.79 \times IQR(\sqrt{OV_t}) \times T^{-1/5}$, with IQR denoting the inter-quartile range. The dashed line is the spot volatility density estimate based on five-minute returns (adjusted for the diurnal intraday pattern) on E-mini S&P 500 futures and using the method developed in [Todorov and Tauchen (2012)](Todorov and Tauchen (2012)), with regularization parameter set to $R = 1.6$.

variance) are quite close for most of the stocks in the sample. This is remarkable given that the recovery of volatility from options is significantly more difficult for individual stocks than for the market index due to the more erratic individual stock behavior and the lower number of individual stock options. We also note that the option-based volatility estimates for the individual stocks are significantly noisier than their counterparts for the market index. This is mostly due to the fact that for many of the stocks, the number of available options per day is significantly lower than for the market index. In spite of that, even for individual stocks, unreported results suggest that $OV$ is slightly smoother than $TV$. Finally, we can see that, similar to the market index, $BSIV$ and the $VIX$ index are significantly upward biased measures of the spot volatility of individual stocks. For all stocks, $BSIV$ is roughly 50% higher than $TV$ and $VIX$ is roughly twice as high as $TV$ in terms of time-series medians. As such, these measures, unlike $OV$, are not direct measures of volatility and hence we will not use them in our analysis henceforth.

### 3 Forecasting Gains from Combining Noisy Volatility Proxies

Having a more precise estimate of spot volatility should help for forecasting future volatility. In fact, since the measurement errors in the return- and option-based volatility estimators are asym-
totically uncorrelated, see Todorov (2019), combining these two volatility proxies into a single one is the optimal thing to do. We now illustrate this point formally. Suppose that we are interested in forecasting a random variable $Z_{t+1}$ using two observable vectors $\hat{V}_{t,1}$ and $\hat{V}_{t,2}$ given by

$$\hat{V}_{t,1} = V_t + \Sigma_{t,1} \times \epsilon_{t,1} \quad \text{and} \quad \hat{V}_{t,2} = V_t + \Sigma_{t,2} \times \epsilon_{t,2},$$

where $V_t$, $\epsilon_{t,1}$ and $\epsilon_{t,2}$ are $k \times 1$ random vectors independent of each other, $\epsilon_{t,1}$ and $\epsilon_{t,2}$ satisfy

$$\mathbb{E}(\epsilon_{t,1}) = \mathbb{E}(\epsilon_{t,2}) = 0_k, \quad \text{var}(\epsilon_{t,1}) = \text{var}(\epsilon_{t,2}) = I_k,$$

and $\Sigma_{t,1}$ and $\Sigma_{t,2}$ are two diagonal matrices which are independent from $\epsilon_{t,1}$ and $\epsilon_{t,2}$. We will assume further that $\{(Z_t, V_t^\top)^\top \}_{t \geq 1}$ is a covariance stationary process, and for simplicity of exposition that $Z_t$ and $V_t$ have both mean zero (i.e., we have de-meaned the variables of interest). In addition, $\mathbb{E}(\Sigma_{t,1}^2) = \sigma_1^2 I_k$ and $\mathbb{E}(\Sigma_{t,2}^2) = \sigma_2^2 I_k$, for some constants $\sigma_1$ and $\sigma_2$. In our setting we can think of $Z_{t+1}$ as being the value of a volatility measure in the next time period which we are trying to forecast using two noisy measures of daily volatility estimates in the past and today - one from returns and one from options (i.e., $TV$ and $OV$ introduced in the previous section). As mentioned above, limit theory for $TV$ and $OV$, see Barndorff-Nielsen and Shephard (2004) and Todorov (2019), implies that they satisfy approximately the specification in (9)-(10).

Our goal will be to characterize the optimal linear forecast of $Z_{t+1}$, with respect to a square loss function, using the various volatility proxies constructed from $\hat{V}_{t,1}$ and $\hat{V}_{t,2}$. As is well known, this optimal forecast is given by the linear projection which we will denote with $\hat{E}(Z|X)$ when using a generic vector $X$ to predict a generic random variable $Z$. We will first show that

$$\hat{E}(Z_{t+1}|\hat{V}_{t,1}, \hat{V}_{t,2}) = \hat{E}(Z_{t+1}|\hat{V}_{t,m}),$$

where we denote

$$\hat{V}_{t,m} = (\sigma_1^2 + \sigma_2^2)^{-1} \sigma_2^2 \hat{V}_{t,1} + (\sigma_1^2 + \sigma_2^2)^{-1} \sigma_1^2 \hat{V}_{t,2}.$$  (12)

That is, the optimal way to combine the volatility proxies for the purposes of volatility forecasting is by weighting them optimally which leads to using the single predictor $\hat{V}_{t,m}$. Note that each element of $\hat{V}_{t,m}$ puts the same weight to the corresponding elements of $\hat{V}_{t,1}$ and $\hat{V}_{t,2}$. To establish the above result, lets suppose that $\hat{E}(Z_{t+1}|\hat{V}_{t,m})$ is given by

$$\hat{E}(Z_{t+1}|\hat{V}_{t,m}) = \alpha^\top \hat{V}_{t,m},$$  (13)
for some vector $\alpha$ of weights that satisfies

$$\mathbb{E} \left[ (Z_{t+1} - \alpha^\top \hat{V}_{t,m})\hat{V}_{t,m} \right] = 0_K. \quad (14)$$

To show the result in (11), we need to establish that the following two conditions hold

$$\mathbb{E} \left[ (Z_{t+1} - \alpha^\top \hat{V}_{t,m})\hat{V}_{t,1} \right] = 0_K \text{ and } \mathbb{E} \left[ (Z_{t+1} - \alpha^\top \hat{V}_{t,m})\hat{V}_{t,2} \right] = 0_K. \quad (15)$$

This result, however, follows from the following easy to verify equalities (by making use of the fact that $\mathbb{E}((\epsilon_{t,1} - \epsilon_{t,2})(\hat{V}_{t,m} - V_t)^\top) = 0_{K \times K}$):

$$\mathbb{E} \left[ (Z_{t+1} - \alpha^\top \hat{V}_{t,m})(\hat{V}_{t,1} - \hat{V}_{t,m}) \right] = 0_K \text{ and } \mathbb{E} \left[ (Z_{t+1} - \alpha^\top \hat{V}_{t,m})(\hat{V}_{t,2} - \hat{V}_{t,m}) \right] = 0_K. \quad (16)$$

From here, for the expected losses from using $\hat{V}_{t,1}$ and $\hat{V}_{t,2}$ separately and in combination in constructing linear forecasts for $Z_{t+1}$, we can write

$$\mathbb{E} \left( Z_{t+1} - \hat{E}(Z_{t+1}|\hat{V}_{t,i}) \right)^2 = \mathbb{E}(Z_{t+1}^2) - \mathbb{E}(Z_{t+1}V_i) \left( \mathbb{E}(V_iV_i') + \sigma_i^2 I_k \right)^{-1} \mathbb{E}(Z_{t+1}V_i), \ i = 1, 2, \quad (17)$$

$$\mathbb{E} \left( Z_{t+1} - \hat{E}(Z_{t+1}|\hat{V}_{t,1}, \hat{V}_{t,2}) \right)^2 = \mathbb{E}(Z_{t+1}^2) - \mathbb{E}(Z_{t+1}V_i) \left( \mathbb{E}(V_iV_i') + \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} I_k \right)^{-1} \mathbb{E}(Z_{t+1}V_i). \quad (18)$$

Clearly, mixing the two volatility proxies in a linear way leads to reduction in the forecast error, with the size of this reduction depending on the noise in the two volatility proxies and the covariance between $V_t$ and $Z_{t+1}$. Can we do better? Suppose we know $\Sigma_{t,1}$ and $\Sigma_{t,2}$ (in practice we have estimates for $\Sigma_{t,1}$ and $\Sigma_{t,2}$ from the feasible CLT results for the volatility measures). In this case, we can construct the optimal estimate of $V_t$ on each day by weighting appropriately the two volatility proxies:

$$\hat{V}_{t,o} = \omega_t \hat{V}_{t,1} + (I_k - \omega_t) \hat{V}_{t,2}, \ \omega_t = (\Sigma_{t,1}^2 + \Sigma_{t,2}^2)^{-1} \Sigma_{t,2}^2. \quad (19)$$

The mean squared loss from using $\hat{V}_{t,o}$ for forecasting $Z_{t+1}$ is now given by

$$\mathbb{E} \left( Z_{t+1} - \hat{E}(Z_{t+1}|\hat{V}_{t,o}) \right)^2 = \mathbb{E}(Z_{t+1}^2) - \mathbb{E}(Z_{t+1}V_i') \left( \mathbb{E}(V_iV_i') + \mathbb{E}(\omega_t \omega_t^\top) \right)^{-1} \mathbb{E}(Z_{t+1}V_i). \quad (20)$$

\footnote{The estimator $\hat{V}_{t,o}$ is reminiscent of the optimal, in an asymptotically mean squared error sense, estimator of daily volatility constructed using return data across days by Ghysels et al. [2020].}
By Jensen’s inequality, the difference
\[ \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} I_k - E(\omega_t \omega_t^T), \] (21)
is positive semidefinite. Therefore, \( \hat{V}_{t,o} \) provides more efficient forecast than \( \hat{V}_{t,m} \), with the size of the efficiency gains depending on the time variation in \( \omega_t \). In fact, using exactly the same steps as above, we can show that \( \hat{V}_{t,o} \) is the optimal linear forecast when \( \Sigma_{t,1} \) and \( \Sigma_{t,2} \) are known. Obviously, the same result will continue to hold even when the vectors \( \hat{V}_{t,1} \) and \( \hat{V}_{t,2} \) are infinite dimensional (i.e., when we are forecasting using the whole history of volatility proxies).

Overall, our theoretical analysis shows that for the purposes of volatility forecasting and filtering, the optimal thing to do is to use the volatility proxies in a way that diversifies optimally the measurement error in them.

4 Market Volatility Forecasting

We now study empirically the gains from including options for the purposes of generating volatility forecasts. We start in this section with forecasting market volatility. We first introduce in Section 4.1 the models we use for generating the volatility forecasts as well as the estimation method and then in Section 4.2 we present the empirical results.

4.1 Volatility Forecasting Models and Inference

Our goal is to forecast total future return variance. In general, the optimal linear volatility forecast will be a moving average of the current and past realizations of the variable used in the forecasting. To achieve parsimony, we will follow the recent literature on volatility forecasting (see e.g., Bollerslev et al. (2016)) and consider the heterogeneous autoregressive (HAR) type models proposed by Corsi (2009), which are autoregressive models with coefficients that are suitably restricted. The generic HAR forecasting model for \( h \)-day ahead realized variance \( RV_{t,t+h} \) with predictor \( V_t \) is given by:

\[ RV_{t,t+h} = \beta_{0,h} + \beta_{1,h} V_{t,d} + \beta_{2,h} V_{t,w} + \beta_{3,h} V_{t,m} + \epsilon_{t+h}, \] (22)

where \( RV_{t,t+h} = \sum_{j=1}^{h} RV_{t+j} \), \( V_{t,d} = V_t \), \( V_{t,w} = \sum_{j=1}^{5} V_{t+1-j} \) and \( V_{t,m} = \sum_{j=1}^{22} V_{t+1-j} \). In addition to the past value of the predictor, the predictive model in (22) also includes the average values of the predictor over the past week and month. The weekly and monthly regressors allow to
capture in a parsimonious way the well-known long-memory feature of volatility. We refer to the predictive model for the future $RV$ based on the single variance predictor $V$ as HAR-$V$. For $V$ (the predictor), we consider four different measures of volatility, mainly the return-based $RV$ and $TV$, the option-based $OV$ and a combined one denoted with $EV$. Given the discussion in the previous section, $EV_t$ is constructed by optimally weighting the return- and option-based volatility proxies using estimates of the asymptotic variances for these measures on day $t$. The details about the constriction of $EV_t$ are provided in the Appendix.$^9$

We also look at a forecasting model that combines the return- and option-based volatility predictors $TV_t$ and $OV_t$, which we refer to as the HAR-MV:

$$RV_{t,t+h} = \beta_{0,h} + \beta_{1,h}OV_{t,d} + \beta_{2,h}OV_{t,w} + \beta_{3,h}OV_{t,m}$$
$$+ \gamma_{1,h}(TV_{t,d} - OV_{t,d}) + \gamma_{2,h}(TV_{t,w} - OV_{t,w}) + \gamma_{2,m}(TV_{t,m} - OV_{t,m}) + \epsilon_{t+h}. \tag{23}$$

As illustrated in Section 3, the inclusion of the two alternative volatility estimates helps “diversify” the measurement error in them, and this should be beneficial for the purposes of volatility forecasting. Unlike the HAR-EV model, however, in the HAR-MV model we do not restrict the weights assigned to the return- and option-based volatility estimators. On one hand, this provides an advantage as we do not need estimates of the asymptotic variances of the volatility estimators. On the other hand, however, we have more parameters to estimate in the predictive regression and we lose the advantage of assigning time-varying weights to the return- and option-based volatility estimates. This tradeoffs will be assessed empirically below.

Finally, we compare the forecasting performance of the above models with that of the HARQ model proposed recently by Bollerslev et al. (2016), which is given by

$$RV_{t,t+h} = \beta_{0,h} + (\beta_{1,h} + \beta_{1Q,h}RQ_{t}^{1/2})RV_{t,d} + \beta_{2,h}RV_{t,w} + \beta_{3,h}RV_{t,m} + \epsilon_{t+h}. \tag{24}$$

where $RQ_t$ is the so-called realized quarticity (estimate for the asymptotic variance of $RV_t$) constructed from the high-frequency data. The extra explanatory variable in the above predictive regression (relative to HAR-RV) is aimed at assigning time-varying weight to past volatility when forecasting future volatility according to a measure of the precision with each realized volatility is estimated from the data.$^{10}$

$^9$In the case of individual stocks, we can consider also adding past squared overnight returns in the predictive regressions because overnight periods generate nontrivial volatility for these assets. Similar to results in Bollerslev and Todorov (2011), we find that the gains from this are rather limited and therefore we do not do this here.

$^{10}$In an analogy to the HARQ model, we also experimented with augmenting the HAR-EV model by including the
Turning next to the estimation, we employ robust inference techniques to guard against the effect of outliers. This is very pertinent for the analysis here as, due to jumps, the RV series contains a lot of outliers, particularly on the individual stock level. To this end, we use the Huber loss function \cite{Huber1964}, which modifies the square loss function corresponding to OLS estimation by replacing it with linear function for large values of the losses. The use of this robust inference technique for the estimation of the model parameters, improves nontrivially the forecasting performance of all considered models.

We evaluate the forecasting performance of the different models using two out-of-sample approaches, with forecasts generated using only past information. One approach uses a rolling window (RW) for the model estimation and the other uses increasing window (IW) for this. In our implementation, we set the length of the rolling window to 1000 days. Our out-of-sample forecasting period is from December 2011 to December 2018, with a total of 1758 trading days.

4.2 Empirical Evidence

We assess the forecasting performance of the different models using the following loss functions:

$$\text{MSE}(Y, \hat{Y}) = (Y - \hat{Y})^2, \quad \text{QLIKE}(Y, \hat{Y}) = Y \left( \frac{Y}{\hat{Y}} \right) - \log \left( \frac{Y}{\hat{Y}} \right) - 1, \quad (25)$$

where $Y$ denotes the generic variable to be predicted and $\hat{Y}$ the candidate forecast.\footnote{When generating forecasts from each of the models, we apply an “insanity filter” for the forecasts, i.e., if a forecast is outside of the range of the values of the target realized variance observed in the estimation period, the forecast is winsorized by the minimum or maximum values of observed realized variance in the estimation period. Note that, even though the considered models are linear and the predictors are non-negative, negative parameter estimates for some of the coefficients in the forecasting models can nevertheless lead in certain cases to negative volatility forecasts.}

The results for one-day ahead volatility forecasts are presented in Panel A of Table 2. For ease of comparison, we normalize the forecast losses to the ones corresponding to the benchmark HAR-TV model. As seen from the table, the HAR-TV model performs very similar to the HAR-RV model in which the past values of RV are used instead of those of TV. We note in this regard that for models such as the Heston model in which the jump intensity is linear in the diffusive spot volatility, the optimal forecast of RV is a function of the diffusive volatility only, and hence it is beneficial to split return variation into diffusive and discontinuous one and use the former term $EV_t \times \sqrt{\widehat{A\text{var}}(EV_t)}$, for $\widehat{A\text{var}}(EV_t)$ denoting an estimate for the asymptotic variance of $EV_t$. For brevity, we do not report the estimation results for such an extension of the HAR-EV model as we found its performance to be very similar to that of HAR-EV.
only in the volatility forecasting (i.e., HAR-TV should perform better than HAR-RV). However, if the diffusive volatility jumps and its jumps are related to the price jumps, then it might be beneficial to use in addition to past TV also the past realized jump variation, i.e., the difference RV-TV, when forming the volatility forecast. Finally, the comparison between the HARQ model and our benchmark HAR-TV model provides mixed results. For some metrics, HARQ provides an improvement over HAR-TV but for other metrics it does worse.

Table 2: Forecasting Results for the S&P 500 Index

<table>
<thead>
<tr>
<th>Window</th>
<th>Criterion</th>
<th>HAR-RV</th>
<th>HAR-TV</th>
<th>HARQ</th>
<th>HAR-OV</th>
<th>HAR-MV</th>
<th>HAR-EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolling Window</td>
<td>MSE (Median)</td>
<td>0.9753</td>
<td>1.0000</td>
<td>0.8769</td>
<td>0.6857</td>
<td>0.6680</td>
<td>0.6182</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>0.9847</td>
<td>1.0000</td>
<td>1.0483</td>
<td>0.8927</td>
<td>0.8637</td>
<td>0.8605</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>0.9502</td>
<td>1.0000</td>
<td>0.8904</td>
<td>0.7315</td>
<td>0.6934</td>
<td>0.7107</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>0.9695</td>
<td>1.0000</td>
<td>1.4667</td>
<td>0.9267</td>
<td>0.8152</td>
<td>0.8528</td>
</tr>
<tr>
<td>Increasing Window</td>
<td>MSE (Median)</td>
<td>1.0317</td>
<td>1.0000</td>
<td>0.7435</td>
<td>0.6186</td>
<td>0.5797</td>
<td>0.5656</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>0.9646</td>
<td>1.0000</td>
<td>0.9655</td>
<td>0.9016</td>
<td>0.8799</td>
<td>0.8740</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>1.0065</td>
<td>1.0000</td>
<td>0.7935</td>
<td>0.7335</td>
<td>0.6925</td>
<td>0.7426</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>0.9980</td>
<td>1.0000</td>
<td>0.9005</td>
<td>1.0377</td>
<td>0.7729</td>
<td>0.9245</td>
</tr>
</tbody>
</table>

<p>| Panel B: One-week Ahead Relative Forecasting Loss |</p>
<table>
<thead>
<tr>
<th>Window</th>
<th>Criterion</th>
<th>HAR-RV</th>
<th>HAR-TV</th>
<th>HARQ</th>
<th>HAR-OV</th>
<th>HAR-MV</th>
<th>HAR-EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolling Window</td>
<td>MSE (Median)</td>
<td>1.0411</td>
<td>1.0000</td>
<td>0.8623</td>
<td>0.6522</td>
<td>0.6389</td>
<td>0.6133</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0161</td>
<td>1.0000</td>
<td>0.9715</td>
<td>0.8955</td>
<td>0.8728</td>
<td>0.8818</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>0.9686</td>
<td>1.0000</td>
<td>0.8531</td>
<td>0.7160</td>
<td>0.6953</td>
<td>0.7087</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>0.9966</td>
<td>1.0000</td>
<td>1.0357</td>
<td>0.9121</td>
<td>0.8019</td>
<td>0.8814</td>
</tr>
<tr>
<td>Increasing Window</td>
<td>MSE (Median)</td>
<td>0.9945</td>
<td>1.0000</td>
<td>0.8770</td>
<td>0.7246</td>
<td>0.6513</td>
<td>0.6559</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>0.9920</td>
<td>1.0000</td>
<td>0.9970</td>
<td>0.8432</td>
<td>0.8653</td>
<td>0.8227</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>0.9670</td>
<td>1.0000</td>
<td>0.9078</td>
<td>0.8794</td>
<td>0.7522</td>
<td>0.8205</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>0.9988</td>
<td>1.0000</td>
<td>0.9722</td>
<td>1.0954</td>
<td>0.8600</td>
<td>1.0007</td>
</tr>
</tbody>
</table>

<p>| Panel C: One-month Ahead Relative Forecasting Loss |</p>
<table>
<thead>
<tr>
<th>Window</th>
<th>Criterion</th>
<th>HAR-RV</th>
<th>HAR-TV</th>
<th>HARQ</th>
<th>HAR-OV</th>
<th>HAR-MV</th>
<th>HAR-EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolling Window</td>
<td>MSE (Median)</td>
<td>0.9759</td>
<td>1.0000</td>
<td>0.8593</td>
<td>0.7169</td>
<td>0.8331</td>
<td>0.7649</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0068</td>
<td>1.0000</td>
<td>0.9791</td>
<td>0.9851</td>
<td>0.9523</td>
<td>0.9856</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>0.9658</td>
<td>1.0000</td>
<td>0.8716</td>
<td>0.8533</td>
<td>0.8654</td>
<td>0.8641</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>1.0050</td>
<td>1.0000</td>
<td>0.9904</td>
<td>0.9867</td>
<td>0.9146</td>
<td>0.9903</td>
</tr>
<tr>
<td>Increasing Window</td>
<td>MSE (Median)</td>
<td>0.9878</td>
<td>1.0000</td>
<td>0.9245</td>
<td>0.6183</td>
<td>0.8846</td>
<td>0.6997</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>0.9903</td>
<td>1.0000</td>
<td>1.0051</td>
<td>0.9201</td>
<td>0.9713</td>
<td>0.9078</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>0.9407</td>
<td>1.0000</td>
<td>0.9031</td>
<td>0.7491</td>
<td>0.8974</td>
<td>0.8027</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>1.0007</td>
<td>1.0000</td>
<td>0.9950</td>
<td>1.0179</td>
<td>0.9796</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Note: The entries in the table are the ratios of the out-of-sample forecasting losses for different models relative to the ones for the HAR-TV model.

Turning next to a comparison between the return-based and option-based volatility forecasting
models, we note that the HAR-TV model is outperformed by the HAR-OV model for all metrics and estimation methods with the exception of the QLIKE criterion for IW case. The improvement offered by HAR-OV in terms of medians of the loss functions, in particular, is rather nontrivial. This result is indicative of the fact that OV is a more precise estimator of the spot diffusive volatility.

Recall from our theoretical analysis in the previous section that the best performing forecasting models should be the ones in which the return and option volatility proxies are combined, i.e., HAR-MV and HAR-EV should perform best. The results in Table 2 confirm that this is the case empirically. We note that, since the reported forecasting results are out of sample, the evidence in favor of HAR-MV is not due to overfitting as estimation of redundant coefficients will lead to noisier forecasts. Recall also that HAR-EV has the same number of predictors as HAR-TV. The reduction in the forecasting error from using HAR-MV over the benchmark return-based HAR-TV is rather nontrivial, particularly when looking at the medians of the forecasting losses with those of HAR-MV being more than 30% smaller than their counterparts for HAR-TV. This, of course, is in line with the “diversification” of the measurement error when combining TV and OV in the forecasting model, and our empirical results show that the gains from doing this are rather nontrivial.

Finally, we note that the forecasting model HAR-EV performs similar to its counterpart HAR-MV for some configurations and for others it performs somewhat worse. Recall from Section 3 that allowing for the weight assigned to return-based and option-based volatility to vary over time can provide additional efficiency gains, provided this time variation is non-trivial. Our empirical results indicate that such gains are small and even non-existent, which is probably due to the noise in estimating the asymptotic variances and the fact that time variation in the weight is not significant. Another possible explanation is that there is some discrepancy between the return- and option-based volatility proxies, i.e., a time-varying bias in one or both of them, which makes the separate inclusion of OV and TV beneficial. Overall, the results reported in Panel A of Table 2 suggest that short-dated options can help nontrivially in forecasting one-day ahead volatility. These gains are due to the more precise measurement of the true spot volatility and the “diversification” of the error in measuring volatility when combining option- and return-based measures of volatility.

We next compare the forecasting performance of the different models for horizons of one week and one month. The results are reported in Panels B and C of Table 2 and they are in line with those for the one-day ahead forecasts reported above. Mainly, the HAR-OV model tends to outperform the benchmark HAR-TV model. In addition, and as implied by the diversification of the measurement error in the volatility proxies, the mixture models outperform all other models. As expected, the gains decrease somewhat relative to the ones for the daily horizon, particularly
for the monthly horizon, but nevertheless they are still quite large.

In order to test formally whether the HAR-MV model provides a statistically significant improvement over the benchmark return-based HAR-TV model, we use the modification of the Diebold-Mariano test proposed by Clark and West (2007) for comparing forecasts of nested models. The null hypothesis of this test is that the HAR-TV and HAR-MV forecasting models generate the same forecasting loss while the alternative is that HAR-MV performs better. The results of the test for the three different forecasting horizons and the two estimation methods all indicate rejection of the null hypothesis at conventional significance levels, with p-values in all cases being below 0.01.

We end this section with comparing the forecasting performance of the best performing mixture models HAR-MV and HAR-EV and the benchmark HAR-TV model when using an alternative proxy of total return volatility. The forecasting gains of the mixture models over the benchmark HAR-TV model should be easier to see with a less noisy proxy for future volatility. Towards this end, we can utilize the efficient diffusive volatility estimator \( EV_t \) to construct a more efficient estimate of total quadratic variation over the day \([t-1, t]\) via

\[
ORV_t = EV_t + RV_t - TV_t. \tag{26}
\]

The above estimator uses the optimal \( EV_t \) for the diffusive part of the return variation and a return-based estimate of the realized jump variation. We note that options cannot help in recovering the realized jump variation. They can only be used to infer the conditional risk-neutral expectation of this quantity. As discussed earlier, \( EV_t \) is an estimate of the spot diffusive volatility at time \( t \). It can be viewed as an estimator of volatility over \([t-1, t]\), provided the stochastic variation in volatility over the interval is negligible. \( ORV \) reduces the measurement error in \( RV \), and therefore the reduction in the expected losses of the HAR-MV and HAR-EV models over the HAR-TV model should be bigger in percentage terms. The empirical results reported in Table 3 confirm this insight. Indeed, the gains of HAR-MV over HAR-TV increase uniformly, across forecasting horizons and estimation methods, relative to the case of forecasting the noisier \( RV \).

5 Volatility Forecasting for Individual Stocks

We continue in this section with evaluating the gains from using option data for volatility forecasting on individual stock level. We estimate the univariate volatility forecasting models for all 28 stocks in our sample. We do not construct the \( EV \) series, which combines the return- and option-based

\(^{12}\) An alternative is to use intraday option data and recover \( EV \) intraday.
Table 3: ORV Forecasting Results for the S&P 500 Index

<table>
<thead>
<tr>
<th>Window</th>
<th>Criterion</th>
<th>One-day Ahead</th>
<th>One-week Ahead</th>
<th>One-month Ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>HAR-MV</td>
<td>HAR-EV</td>
<td>HAR-MV</td>
</tr>
<tr>
<td>Rolling Window</td>
<td>MSE (Mean)</td>
<td>0.7317</td>
<td>0.7136</td>
<td>0.7395</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>0.6007</td>
<td>0.5766</td>
<td>0.7268</td>
</tr>
<tr>
<td>Increasing Window</td>
<td>MSE (Mean)</td>
<td>0.6778</td>
<td>0.6650</td>
<td>0.6842</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>0.5158</td>
<td>0.5073</td>
<td>0.6386</td>
</tr>
</tbody>
</table>

Note: The entries in the table are the ratios of the out-of-sample forecasting losses for different models relative to the ones for the HAR-TV model.

Volatility estimates, since for the individual stocks, as already discussed in Section 2, the number of available options is significantly smaller and this leads to unreliable estimates of the asymptotic variance of $OV$. Due to the shorter sample, our out-of-sample forecasting period is from January 2013 to December 2019. For each stock, we estimate the HAR models for generating one-day, one-week, and one-month ahead volatility forecasts. In order to save space, we report only the average relative forecasting loss ratios across all stocks for the different models, with the benchmark being again HAR-TV.

The volatility forecasting results are reported in Table 4. Starting with the one-day ahead volatility forecasts, we can see from Panel A of Table 4 some differences in the ranking of the models, according to their forecasting performance, relative to the case of forecasting market volatility. First, HAR-TV performs best among the return-based forecasting models according to any metric and estimation method. This is likely due to the fact that idiosyncratic jumps in individual stock prices can make raw realized volatility rather noisy. Second, in sharp contrast to market volatility forecasting, here the HAR-OV model is uniformly outperformed by HAR-TV. There are likely two reasons for this. One is that the individual stock option data is of significantly lower quality than the S&P 500 index option data (in terms of number of available options and the noise in them). Another is the possible biases in the option-based volatility measure. Recall from our theoretical discussion earlier, however, that even if $OV$ is noisier than $TV$, combining $OV$ and $TV$ should still offer gains for volatility forecasting. This is confirmed by our results here. Indeed, the HAR-MV model is the best among all considered models with gains of around 4-5% over HAR-TV. The test of Clark and West (2007) of equal performance of HAR-TV and HAR-MV rejects that null at the conventional 5% significance level, both when using rolling and increasing estimation windows.

Turning next to the weekly volatility forecasts, we see from Panel B of Table 4 that the gains from using option data get larger in relative terms than those for the daily horizon. Indeed, now HAR-MV provides more than 15% reduction in the forecasting loss over the HAR-TV model when
Table 4: Forecasting Results for DJIA Stocks

**Panel A: One-day Ahead Relative Forecasting Loss**

<table>
<thead>
<tr>
<th>Window</th>
<th>Criterion</th>
<th>HAR-RV</th>
<th>HAR-TV</th>
<th>HARQ</th>
<th>HAR-OV</th>
<th>HAR-MV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolling Window</td>
<td>MSE (Median)</td>
<td>1.0336</td>
<td>1.0000</td>
<td>1.0234</td>
<td>1.1385</td>
<td>0.9253</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0248</td>
<td>1.0000</td>
<td>1.0219</td>
<td>1.0293</td>
<td>0.9658</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>1.0320</td>
<td>1.0000</td>
<td>1.0105</td>
<td>1.1655</td>
<td>0.9569</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>1.0224</td>
<td>1.0000</td>
<td>1.0413</td>
<td>1.1055</td>
<td>0.9414</td>
</tr>
<tr>
<td>Increasing Window</td>
<td>MSE (Median)</td>
<td>1.0279</td>
<td>1.0000</td>
<td>1.0084</td>
<td>1.0049</td>
<td>0.8617</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0246</td>
<td>1.0000</td>
<td>1.0083</td>
<td>1.0556</td>
<td>0.9697</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>1.0255</td>
<td>1.0000</td>
<td>1.0123</td>
<td>1.1324</td>
<td>0.9199</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>1.0130</td>
<td>1.0000</td>
<td>1.0361</td>
<td>1.1508</td>
<td>0.9481</td>
</tr>
</tbody>
</table>

**Panel B: One-week Ahead Relative Forecasting Loss**

<table>
<thead>
<tr>
<th>Window</th>
<th>Criterion</th>
<th>HAR-RV</th>
<th>HAR-TV</th>
<th>HARQ</th>
<th>HAR-OV</th>
<th>HAR-MV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolling Window</td>
<td>MSE (Median)</td>
<td>1.0178</td>
<td>1.0000</td>
<td>0.9941</td>
<td>0.8795</td>
<td>0.7921</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0265</td>
<td>1.0000</td>
<td>1.0114</td>
<td>0.9586</td>
<td>0.9140</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>1.0215</td>
<td>1.0000</td>
<td>0.9985</td>
<td>0.9288</td>
<td>0.8412</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>1.0192</td>
<td>1.0000</td>
<td>1.0242</td>
<td>0.9731</td>
<td>0.8797</td>
</tr>
<tr>
<td>Increasing Window</td>
<td>MSE (Median)</td>
<td>1.0200</td>
<td>1.0000</td>
<td>0.9766</td>
<td>0.7423</td>
<td>0.7122</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0204</td>
<td>1.0000</td>
<td>0.9985</td>
<td>0.9957</td>
<td>0.9270</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>1.0251</td>
<td>1.0000</td>
<td>0.9989</td>
<td>0.8611</td>
<td>0.8077</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>1.0106</td>
<td>1.0000</td>
<td>1.0078</td>
<td>1.0243</td>
<td>0.8963</td>
</tr>
</tbody>
</table>

**Panel C: One-month Ahead Relative Forecasting Loss**

<table>
<thead>
<tr>
<th>Window</th>
<th>Criterion</th>
<th>HAR-RV</th>
<th>HAR-TV</th>
<th>HARQ</th>
<th>HAR-OV</th>
<th>HAR-MV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rolling Window</td>
<td>MSE (Median)</td>
<td>0.9990</td>
<td>1.0000</td>
<td>0.9780</td>
<td>0.9429</td>
<td>0.8993</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0206</td>
<td>1.0000</td>
<td>0.9947</td>
<td>0.9925</td>
<td>0.9485</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>1.0170</td>
<td>1.0000</td>
<td>0.9906</td>
<td>1.0054</td>
<td>0.9330</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>1.0102</td>
<td>1.0000</td>
<td>0.9997</td>
<td>0.9856</td>
<td>0.9572</td>
</tr>
<tr>
<td>Increasing Window</td>
<td>MSE (Median)</td>
<td>0.9962</td>
<td>1.0000</td>
<td>0.9798</td>
<td>0.8349</td>
<td>0.8266</td>
</tr>
<tr>
<td></td>
<td>MSE (Mean)</td>
<td>1.0112</td>
<td>1.0000</td>
<td>0.9869</td>
<td>1.0170</td>
<td>0.9550</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Median)</td>
<td>1.0054</td>
<td>1.0000</td>
<td>0.9847</td>
<td>0.9298</td>
<td>0.8859</td>
</tr>
<tr>
<td></td>
<td>QLIKE (Mean)</td>
<td>0.9993</td>
<td>1.0000</td>
<td>0.9929</td>
<td>1.0244</td>
<td>0.9568</td>
</tr>
</tbody>
</table>

Note: The entries in the table are the ratios of the cross-sectional averages of the out-of-sample forecasting losses for different models relative to the ones for the HAR-TV model.

using time series medians of the losses in the comparison and around 10% reduction when using time series means of the losses. The reason for the larger gain for the weekly horizon is likely the fact that there is a lot of return variation at daily frequency in stocks that is hard to forecast only with past volatility proxies as in the HAR models considered here, and this tends to put all models closer together in terms of their forecasting performance. When looking at weekly and longer horizons, the role of such higher frequency components of the return variation diminishes and this makes the differences in the performance of the various forecasting models much bigger (in
relative terms). Finally, Panel C of Table 4 shows that gains from using options for the purposes of volatility forecasting exist for a monthly horizon as well. Moreover, almost uniformly across stocks (with the exception of two stocks for the monthly horizon with increasing window), these gains are statistically significant at 5% significance level.

6 The Information Content of Short-Dated Options

In the analysis so far, we have aggregated the available short-dated options into portfolios that allow us to estimate non-parametrically the spot diffusive volatility. Mainly, we have used \( OV \) for volatility measurement and forecasting, in addition to the return-based \( TV \), and we have argued above that this should always provide gains, regardless of the forecasting model, because of the diversification of the volatility measurement errors. It is natural to ask, however, whether there is more relevant information in the short-dated options for the purposes of forecasting future volatility. This is what we aim to answer in this section.

6.1 Risk-Neutral Jump Variation Measures and Their Information Content

We start with extracting jump variation measures from the options which will “span” the remaining information in them. Let’s suppose that the risk-neutral jump compensator of \( x \) is given by \( \eta dt \otimes F(dx) \) for some time-invariant measure \( F \) capturing the distribution of the jump size and a (predictable) stochastic process \( \eta_t \) capturing the jump intensity.\(^{13}\) Then, because of their short time-to-maturity, the option prices \( O_{t,\tau}(k) \) depend on the information at time \( t \), \( \mathcal{F}_t \), only through the spot diffusive variance \( V_t \) and the jump intensity \( \eta_t \). That is, the options are functions of \( V_t \) and \( \eta_t \). The estimator \( OV_t \) is a measure of \( V_t \) and hence the only remaining information to be extracted from the short-dated options is about the jump intensity. Towards this end, we introduce a measure of the risk-neutral jump variation. It is formed as a difference of total risk-neutral return variation and its contribution from the continuous part of the price:

\[
J V_{t,\tau} = \frac{2}{\tau} \sum_{j=2}^{N_x} e^{-k_{j-1}} (1 - k_{j-1} + x_t) \tilde{O}_{t,\tau}(k_{j-1})(k_j - k_{j-1}) - \tilde{V}_{t,\tau}(\tilde{u}_{t,\tau}),
\]

\(^{13}\)In its most general form, the jump compensator takes the form \( dt \otimes F_t(dx) \) which allows for different dynamics of the jumps of different size. However, in essentially all option pricing models considered in prior work, the compensator is split into a jump intensity (possibly different for positive and negative jumps) that does not depend on the jump size and a time-invariant measure of the jump size.

\(^{14}\)The measure \( F \) is a Lévy measure, i.e., it satisfies \( \int_{\mathbb{R}} (x^2 \wedge 1) F(dx) < \infty \), which allows for situations with \( \int_{\mathbb{R}} F(dx) = \infty \) (infinite activity jumps).
and from here by averaging across the two available tenors, we get $JV_t = \frac{1}{2} (JV_{t,\tau_1} + JV_{t,\tau_2})$. It is easy to show that

$$JV_t \approx \int_{\mathbb{R}} x^2 F(dx) \times \eta_t, \quad \text{as } \tau_1 \downarrow 0 \text{ and } \tau_2 \downarrow 0.$$  \hfill (28)

In figure 4, we plot the market index $OV$ and $JV$ series and the cross-sectional averages of these measures for the DJIA stocks. It is interesting to note that both for the individual stocks and the market index, the average values of $OV$ and $JV$ appear very similar. This means that jumps play far more prominent role under the risk-neutral measure than under the statistical one. This is of course not surprising and is manifestation of the large risk premium demanded by investors for bearing jump risk. In terms of time series properties, we can see that $OV$ and $JV$ have similar dynamics but differences in their behavior do appear. In particular, $JV$ tends to spike much higher during crisis episodes than $OV$. In addition, the market index $JV$ declined somewhat faster than market volatility after the crisis period in the beginning of 2009. Another notable and persistent difference between market $OV$ and $JV$ emerges during the quiet period of 2017 and in the aftermath of the market turmoil of February 2018 when $JV$ was consistently above $OV$. This suggests that for the market index, the relationship between $OV$ and $JV$ seems to be shifting over time, with periods in which $JV$ can play more prominent role clearly present. For the DJIA stocks, the differences between $OV$ and $JV$ seem less persistent and $JV$ appears slightly below $OV$ on average.

Figure 4: **Option-Based Volatility Measures.** Solid lines on the plots correspond to $OV$ while dotted lines correspond to $JV$. Top panel displays estimates for the S&P 500 index while bottom panel displays cross-sectional averages of the option estimates for the DJIA stocks.

With the estimate of the quadratic variation due to jumps $JV$, the question whether short-
dated options contain additional relevant information for the purposes of volatility forecasting can be equivalently stated as asking whether JV has incremental information, relative to OV, for future volatility. Figure 4 suggests that OV and JV have somewhat different dynamics and what we aim to investigate now is whether this difference contains signals for the value of future volatility. The answer to this question will of course depend on the dynamics of \((V_t, \eta_t)\). To fix ideas, let’s suppose that \(V_t = h_V(F_{t,1}, F_{t,2}, F_{t,3})\) and \(\eta_t = h_{\eta}(F_{t,1}, F_{t,2}, F_{t,3})\), for some time-invariant functions \(h_V, h_{\eta}\) and a three-dimensional Markov process \((F_{t,1}, F_{t,2}, F_{t,3})\) with independent components. Virtually all parametric continuous-time option pricing models, considered in prior work, have such dynamics for the diffusive volatility and the jump intensity (and in most cases the functions \(h_V\) and \(h_{\eta}\) are linear). In this setup, if both \(V_t\) and \(\eta_t\) load positively on the three latent factors, then adding JV to OV in forming future volatility forecasts can provide gains. On the other hand, if \(V_t\) is loading only on the first two latent factors while \(\eta_t\) can be represented as a function of \(V_t\) and the third factor (similar to the models of Andersen et al. (2015) and Li and Zinna (2018)), then JV does not offer additional gains (over OV) for forecasting future volatility.

Motivated by the above discussion, we will now test whether future volatility (measured by realized volatility) and current JV are independent conditional on the current value of OV. That is, we will test whether the following is true

\[ RV_{t,t+h} \perp \eta_t | V_t. \] (29)

This, in turn, can be equivalently formulated using conditional characteristic functions and using log-transforms:

\[
E\left(e^{iu \log(RV_{t,t+h}) + iz \log(\eta_t)} \big| \log(V_t)\right) = E\left(e^{iu \log(RV_{t,t+h})} \big| \log(V_t)\right) E\left(e^{iz \log(\eta_t)} \big| \log(V_t)\right),
\] (30)

for \(u, z \in \mathbb{R}\). The above expressions involve conditional expectations and therefore, without additional assumptions, one would need to perform nonparametric regressions as in Wang and Hong (2018). Alternatively, we can express the conditional characteristic functions as integrals involving \(V_t\) and unconditional characteristic functions using Fourier inversion, see Bartlett (1938). This will imply the following moment conditions involving only unconditional characteristic functions:

\[
\int_\mathbb{R} \int_\mathbb{R} \psi_{RV,V,\eta}(u, v, z)\psi_{RV,V,\eta}(0, w, 0)\psi_{RV,V,\eta}(0, -w - v, 0)dvdw = \int_\mathbb{R} \int_\mathbb{R} \psi_{RV,V,\eta}(u, v, 0)\psi_{RV,V,\eta}(0, w, z)\psi_{RV,V,\eta}(0, -w - v, 0)dvdw, \ u, z \in \mathbb{R},
\] (31)
where we denote $\psi_{RV,V,\eta}(u,v,z) = \mathbb{E}(e^{iu \log(RV_{t,t+h})+iv \log(V_t)+iz \log(\eta_t)})$, for $u,v,z \in \mathbb{R}$. Note that (33) implies (30) but the reverse does not need to hold. That is, (30) contains more information than (33) regarding the conditional independence hypothesis in (29). Since testing (33) is relatively simple, we will proceed here with a test based on it for a few pairs of $(u,z)$.

Of course, as already discussed above, we do not have direct observations of $V_t$ and $\eta_t$ but instead we have estimates of them which satisfy approximately the following

$$\log(OV_t) = \log(V_t) + \epsilon_{t,1}, \log(JV_t) = \log(\eta_t \int_{\mathbb{R}} x^2 F(dx)) + \epsilon_{t,2},$$

where $(\epsilon_{t,1},\epsilon_{t,2}) \mid RV_{t+h},V_t,\eta_t \sim N(0_{2\times 1}, \Xi_t)$, for some variance-covariance matrix $\Xi_t$ that is given explicitly in the Appendix. Then, it is easy to see that the moment condition (33) is equivalent to

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \psi_{RV,OV,JV}(u,v,z) \psi_{RV,OV,JV}(0,w,0) \psi_{RV,OV,JV}(0,-w-v,0) dv dw = \int_{\mathbb{R}} \int_{\mathbb{R}} \psi_{RV,OV,JV}(u,v,0) \psi_{RV,OV,JV}(0,w,z) \psi_{RV,OV,JV}(-w-v) dv dw,$$

where we denote $\psi_{RV,OV,JV}(u,v,z) = \mathbb{E}(e^{iu \log(RV_{t,t+h})+iv \log(OV_t)+iz \log(JV_t)+(1/2)(vz)\Xi_t(vz)^T})$. Form here, developing a test for (33) is relatively easy. Mainly, we can form sample moment conditions for pairs of $(u,z)$, based on (33), and using the empirical characteristic function of $(RV_{t,t+h},OV_t,JV_t)$. The infinite regions of integration in the finite sample counterparts of (33) are replaced with finite but asymptotically expanding ones (based on the assumption of smoothness of the probability density of $(RV_{t,t+h},OV_t,JV_t)$). If the sample moment condition vector is denoted with $\hat{m}$, which is of dimension $J \times 1$, our test will measure the distance of $\hat{m}$ from $0_{J\times 1}$. The vector $\hat{m}$ is asymptotically normally distributed with mean zero under the null hypothesis. If we denote an estimate of the asymptotic variance with $\hat{\Sigma}_m$, then our test statistics is given by

$$\hat{T}_h = T \hat{m}^T \hat{\Sigma}_m^{-1} \hat{m},$$

and should have an asymptotic $\chi^2(J)$ distribution if the hypothesis in (29) is true. The details on the construction of the test are provided in the Appendix.

We implemented the test on the data for horizons corresponding to one day, one week and one month. Along with the above test, we also performed its counterpart in which $OV_t$ and $JV_t$ have been swapped, i.e., we performed a test for the hypothesis $RV_{t,t+h} \perp \eta_t | V_t$. Since for the individual stocks, the computation of $\Xi$ is unreliable (due to the fewer available strikes for most of the sample
period), we implemented the tests only for the market index. The value of the test statistic for the hypothesis $RV_{t,t+h} \perp \eta_t | V_t$ is 4.5, 12.5 and 8.4, for horizons $h = 1$, $h = 5$ and $h = 22$, respectively. The critical value of the test for 5% significance level is 15.5. This means that the null hypothesis cannot be rejected at this significance level for any of the three horizons. On the other hand, the value of the test statistic for the hypothesis $RV_{t,t+h} \perp V_t | \eta_t$ is 20.1, 26.0 and 15.7, for horizons $h = 1$, $h = 5$ and $h = 22$, respectively. This is evidence against the hypothesis $RV_{t,t+h} \perp V_t | \eta_t$. Taken together, these two sets of test results suggest that $\eta_t$ has dynamics which differs from the diffusive volatility $V_t$ but this difference does not seem to be related to future volatility in a significant way. This is consistent with the empirical finding of [Andersen et al. (2015)] who find that jump intensity contains a large component that is unrelated to diffusive volatility on the basis of a parametric option pricing model fitted to market index options.

### 6.2 Augmented Volatility Forecasting Regressions

While the above results suggest that jump intensity contains no significant additional information for forecasting future volatility over what is contained in the current level of diffusive volatility (at least for the market), there can be nevertheless gains from including $JV$ in a volatility forecasting model. The reason is the measurement error. Mainly, the measurement error in $OV$ and $JV$ is negatively correlated empirically and this means that if the information content of $V_t$ and $\eta_t$ is similar, then one can benefit from combining $OV$ and $JV$ in the forecasting model. This diversification of errors argument is exactly the same as the one discussed in Section 3.

With the above theoretical considerations and theoretical results in mind, we now augment the HAR-MV model by including the past day value of $JV$. We refer to this model as HAR-MV-JV. The reason for including only past day $JV$ is that the gains from diversifying the measurement errors are largest for the daily estimate. In Table 5, we compare the forecasting performance of HAR-MV and HAR-MV-JV models. As seen from the table, their performance is similar. Nevertheless, adding $JV$ in the forecasting model does offer relatively small forecasting improvements. For the case of the S&P 500 index, the adjusted Diebold-Mariano test of equal forecasting performance yields p-values ranging from 0.24% to 6.13%, with the strongest statistical support for adding $JV$ being for the monthly forecasting horizon. For the individual stocks, the cross-sectional median of the test has p-value ranging from 0.27% to 16.54%, with the highest gains (statistically) from adding $JV$ being for daily and weekly horizons.

Overall, the above results show that the relevant information in the short-dated options for forecasting future volatility is contained in a portfolio of options that spans (recovers) the diffusive
spot volatility. Additional (but relatively small) gains can be obtained by adding a measure of the jump variation in the forecasting model.\footnote{The focus of this paper has been on the information in short-dated options as they allow for direct measurement of the spot characteristics of the underlying asset, i.e., the diffusive volatility and the risk-neutral jump intensity. Longer-dated options contain additional information about the volatility and jump risk dynamics but under the risk-neutral probability. Whether and to what extent this information can help in volatility forecasting will depend on the properties of the volatility and jump risk premia dynamics. If the latter can be disentangled from the volatility dynamics under the statistical probability, then one would expect additional volatility forecasting gains from using longer-dated options in forming the forecast. We leave the analysis of this question for future work.}

<table>
<thead>
<tr>
<th>S&amp;P 500 Index</th>
<th>One-day Ahead</th>
<th>One-week Ahead</th>
<th>One-month Ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window</td>
<td>Criterion</td>
<td>HAR-MV</td>
<td>HAR-MV-JV</td>
</tr>
<tr>
<td>Rolling</td>
<td>MSE</td>
<td>0.8637</td>
<td>0.8235</td>
</tr>
<tr>
<td></td>
<td>QLIKE</td>
<td>0.8152</td>
<td>0.7615</td>
</tr>
<tr>
<td>Increasing</td>
<td>MSE</td>
<td>0.8799</td>
<td>0.7760</td>
</tr>
<tr>
<td></td>
<td>QLIKE</td>
<td>0.7729</td>
<td>0.7038</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DJIA Stock</th>
<th>One-day Ahead</th>
<th>One-week Ahead</th>
<th>One-month Ahead</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window</td>
<td>Criterion</td>
<td>HAR-MV</td>
<td>HAR-MV-JV</td>
</tr>
<tr>
<td>Rolling</td>
<td>MSE</td>
<td>0.9658</td>
<td>0.9568</td>
</tr>
<tr>
<td></td>
<td>QLIKE</td>
<td>0.9414</td>
<td>0.9155</td>
</tr>
<tr>
<td>Increasing</td>
<td>MSE</td>
<td>0.9697</td>
<td>0.9600</td>
</tr>
<tr>
<td></td>
<td>QLIKE</td>
<td>0.9481</td>
<td>0.9220</td>
</tr>
</tbody>
</table>

Note: The entries in the table are the mean relative losses of HAR-MV and HAR-MV-JV models with respect to that of the HAR-TV model. For the DJIA stocks, we report cross-sectional averages.

6.3 Implications for Return Predictability

We end this section by documenting the gains offered by the short-dated options for return predictability. As we will show now, these gains stem from the higher precision in measuring volatility as well as from the extra information contained in the options (relative to that contained in the spot volatility). In recent work, Bollerslev et al. (2009) have shown that the market variance risk premium serves as a strong predictor of future returns. The variance risk premium measure of Bollerslev et al. (2009) is given by

$$VRP_t(1) = VIX_t^2 - RV_{t,m},$$

(35)
where \( VIX_t \) denotes the CBOE monthly volatility index, which is the risk-neutral expectation at time \( t \) of the volatility over the next month. Using our efficient volatility estimate \( EV_t \), we can improve on the measurement of the past month realized volatility and this leads to the following more precise measure of the variance risk premium

\[
VRP_t(2) = VIX_t^2 - OVR_{t,m}, \tag{36}
\]

where we recall that our efficient estimate of the realized volatility, \( ORV_t \), is defined in (26). Upon computing the two measures using the S&P 500 index data, we find that \( VRP(2) \) is significantly less noisy than \( VRP(1) \), with the standard deviation of \( VRP(2) \) being approximately 50% smaller than that of \( VRP(1) \). In addition to the above two measures of the variance risk premium, we also consider a jump risk premium predictor defined as

\[
JRP_t = JV_{t,m} - (RV_{t,m} - TV_{t,m}). \tag{37}
\]

Note that \( JV_t \) is a measure of instantaneous risk-neutral jump variation and for this reason we subtract from it the daily realized jump variation. Since the latter is quite noisy over one day, we average the jump risk premium estimate over one month.

We run the following return predictability regressions

\[
r_{t,t+h} = \alpha_{0,h} + \alpha_{1,h}X_t + \epsilon_{t+h}, \tag{38}
\]

for \( X_t \) being \( VRP_t(1) \), \( VRP_t(2) \) or \( JRP_t \). In Figure [5] we plot the \( t \)-statistics for the three predictors from the univariate predictive regressions along with the associated \( R^2 \). We can make several conclusions from the reported results. First, consistent with existing empirical evidence, the estimator \( VRP(1) \) of Bollerslev et al. (2009) works over short horizons of up to six months. For horizons longer or equal to six months, the \( t \)-statistic for \( VRP(1) \) is below the critical value of a one-sided test with size 5%. This is in sharp contrast to the reported results for \( VRP(2) \). Mainly, \( VRP(2) \) is a statistically significant return predictor for horizons of six months and above. These two results suggest that the measurement of the variance risk premium plays a critical role in the findings regarding its predictive content. In particular, our results suggest that the short term return predictability of the measure of Bollerslev et al. (2009) seems to be, at least in part, due to the short-term predictability of the measurement error in \( RV_{t,m} \). The latter will depend on realized jumps as well as volatility of volatility dynamics, so this type of return predictability seems
plausible. Finally, \( VRP(2) \) and \( JRP \) have similar predictive ability, with the latter performing slightly better than the former, both in terms of t-statistics as well as \( R^2 \).

![T-statistics and R2 plots]

Figure 5: **Return Predictability Results.** Standard errors are computed using the Equal-Weighted Cosine estimator of the long-run variance with fixed-b critical values proposed in Lazarus et al. (2018) with degree of freedom equal to \( 0.4 \times T^{2/3} \). The critical value corresponding to the one-sided 5% test is equal to 1.99 and is plotted as a solid line on the top panel.

Overall, our analysis here shows that a more precise measure of variance risk premium, based on the short-dated options, can have a big impact on the conclusion regarding its predictive content and the origins of the return predictability.

### 7 Conclusion

In this paper we study empirically the gains offered by short-dated options for the measurement and forecasting of volatility. Using the approach of Todorov (2019), we construct an option portfolio that provides a nonparametric estimate of the latent spot volatility. Since the measurement errors in the option-based volatility proxy and in a volatility estimate constructed from high-frequency returns on the underlying asset are (asymptotically) uncorrelated, their combined use should provide gains
both for measuring volatility as well as for its forecasting. We document that this is the case empirically for the S&P 500 market index and stocks in the Dow 30 index. We further show that in some cases the risk premium embedded in the option prices as well as the quadratic variation of the jumps in the underlying asset can provide additional but small gains for volatility forecasting.

References


Appendix

A. Details on the High-Frequency Data and Filters

For the return-based volatility measures for the S&P 500 index we use high-frequency trade data on the CME E-mini S&P 500 index futures (ES) over the trading hours Sunday to Friday 5:00 P.M. to 4:00 P.M. (Central Time). We obtain 5-minute high frequency trade series for ES futures from TickData and use the rollover method recommended by TickData (i.e. rollover on the 12th day of the expiration month). The return-based volatility measures for the individual stocks are constructed from their high-frequency price records over the trading hours Monday to Friday 8:30 A.M. to 3:00 P.M. (Eastern Time). We obtain Millisecond Trade intraday records for each stock from Trade And Quotation (TAQ) dataset. Following Barndorff-Nielsen et al. (2009), we construct five-minute trade-based price record for the individual stocks by applying the following filters:

1. Delete entries with a trade price of zero.
2. Delete entries with corrected trades (Trades with a nonzero Correction Indicator).
3. Delete entries with abnormal sale condition (Trades with a letter sale condition code).
4. If multiple transactions have the same time stamp: use the median price.
5. Delete entries for which the price deviated by more than 10 mean absolute deviations from a rolling centered median (excluding the observation under consideration) of 50 observations (25 observations before and 25 after).

B. Details on the Option Data and Filters

We obtain daily option data at market close for the S&P 500 index and 15 individual stocks in the Dow 30 index from OptionMetrics IvyDB US file. The underlying spot price is from the security file and the risk-free rate is from the 30-day maturity zero bond file. We download the earning announcement and pre-announcement dates from Zacks. Following Andersen et al. (2017), we apply a set of filters, retaining only trading days \((t)\), maturities \((T)\) and option observations according to the following criteria:
1. The trade date of the option is not an abbreviated trading day, a U.S. holiday, or a low-activity trading day just prior to U.S. holidays.
2. The close-ask and close-bid quotes are not missing and are strictly positive, with the ratio of close-ask to close-bid less than 10.
3. For the deep-out-of-money options, remove stale quotes by keeping the first call and put with the minimum tick size.
4. Drop \((t, T)\) pairs with less than 3 distinct OTM calls and puts.
5. For S&P 500 index options, keep options with 2 to 33 business days to expiration. For individual stock option, retain options with 3 to 45 business days to expiration.

The daily option-based volatility measures are calculated from up to two distinct short-dated options for S&P 500 index and single-name equity options, respectively. Figure 6 displays the time-to-maturity and the number of strikes per day in our sample for the options on the S&P 500 index.

![Near-term Time-to-maturity](image)

![Next-term Time-to-maturity](image)

![Average Number of Strikes](image)

Figure 6: Option Data Summary for SPX.
C. Details on the Option-based Volatility Measures

C.1. Time-to-maturity Convention and AM Settlement Adjustment

We count time-to-maturity in trading days (and exclude U.S. and exchange holidays). The traditional monthly SPX options are A.M. settled on the 3rd Friday of the calendar month and can be traded until the end of trading on the preceding Thursday. The settlement takes place in the morning of the expiration day. We calculate time-to-maturity for AM settlement SPX options using an overnight adjustment factor. This overnight adjustment factor is computed as the ratio between close-to-close full intraday return volatility and open-to-close intraday return volatility,

\[
ON = \frac{\sum_{t=1}^{T} (\sum_{i=1}^{n_{cc}} \Delta_{t,i}^{n_{cc}})^2}{\sum_{t=1}^{T} (\sum_{i=1}^{n_{oc}} \Delta_{t,i}^{n_{oc}})^2},
\]

(39)

where \(n_{cc}\) and \(n_{oc}\) represent the number of close-to-close and open-to-close observed price increments. The time-to-maturity for the AM settled SPX option is then computed as

\[
T = T_{AM} - 1 + \frac{ON - 1}{ON},
\]

(40)

where \(T_{AM}\) denotes the number of business days till expiration (including the expiration day as well). We calculate the overnight adjustment factor for SPX based on the 5-minute CME E-mini S&P 500 index futures data for our sample period from January 2008 till December 2018.

C.2. Determining Option Moneyness

We determine moneyness using the synthetic forward level which we compute from the put-call parity:

\[
F_{t,T} = K_{near,T} + e^{r_f \tau} \left[ \hat{O}_{t,T}(K_{near,T}, C) - \hat{O}_{t,T}(K_{near,T}, P) \right],
\]

(41)

where \(r_f\) is the risk-free rate and \(K_{near,T}\) represents the strike price with smallest put-call absolute price difference. In practice, we use up to 3 near-the-money put-call pairs with the smallest put-call price absolute difference to calculate the forward level and we then take the median of the resulting forward estimates.
C.3. Interpolation and Extrapolation for Missing Strikes

The integral in Equation (4) is approximated by a Riemann sum. In order to minimize the impact from uneven strike grid or from missing option quotes for very low or very high strikes, we perform interpolation and extrapolation which we now describe.

1. **Interpolation.** We fill strike gaps by linearly interpolating the Black-Scholes implied volatilities calculated from the observed option prices on a dense uniform strike price grid:

   \[ K = [K_{\text{low}} : \Delta K : K_{\text{high}}] \]  \hspace{1cm} (42)

   where \( K_{\text{low}} \) and \( K_{\text{high}} \) stands for the lowest and highest observed strike price in the data. We set the strike grid \( \Delta K \) to 1/8 of the minimum observed strike intervals. We then calculate option prices on the fine strike grid using the Black-Scholes formula and the interpolated Black-Scholes implied volatilities. The above-described interpolation is applied both to the S&P 500 index and individual stock options.

2. **Extrapolation.** The option-based volatility measure \( OV \) has its value determined mostly from the prices of near-the-money options. Nevertheless, for individual stocks and in relatively rare cases for the S&P 500 index options, we can have a lot of missing option observations even for moderate levels of moneyness. That is, for the lowest and/or highest available strike, the option price might be far from zero. This can have an adverse effect on the volatility extraction from the options. To prevent this, we perform a tail extension for all individual equity option pair \((t, T)\). We follow Bollerslev et al. (2015) and assume that the return distribution has regular variation in the tails. More specifically, we first compute the slopes for left and right tail decays from the quantiles of the observed OTM call and put option prices separately:

   \[ \hat{\lambda}^+ = \frac{\log(\hat{O}(K_{Q75}, C)/\hat{O}(K_{Q50}, C))}{\log(K_{Q50}^C/K_{Q75}^C)} \sqrt{5}, \] \hspace{1cm} (43)

   \[ \hat{\lambda}^- = \frac{\log(\hat{O}(K_{Q25}, P)/\hat{O}(K_{Q40}, P))}{\log(K_{Q25}^P/K_{Q40}^P)} \sqrt{5}, \] \hspace{1cm} (44)

   where \( K_{Q_p} \) is the \( p \)-th quantile of the observed strike prices in the data. We then extend both tails with strike increment \( \Delta K \) starting from the deepest OTM call and put strikes in the observed option data and iterating until the extrapolated call and put option prices fall
\[
\hat{O}_{t,T}(K,C) = e^{-\hat{\lambda}^+|\log((K_{\text{high}}+\Delta K)/K_{\text{high}})|} \hat{O}_{t,T}(K_{\text{high}}, C), \ K > K_{\text{high}}, \quad (45)
\]
\[
\hat{O}_{t,T}(K,P) = e^{-\hat{\lambda}^-|\log((K_{\text{low}}-\Delta K)/K_{\text{low}})|} \hat{O}_{t,T}(K_{\text{low}}, P), \ K < K_{\text{low}}. \quad (46)
\]

where \(K_{\text{low}}\) and \(K_{\text{high}}\) are the lowest and highest, respectively, observed strikes.

C.4. Volatility Estimation Prior to Earning Announcements

If on a given day prior to an earnings announcement, both available expiration dates are after the announcement, we need to modify the original estimator in (5) in order to account for the fixed time of discontinuity in the underlying asset price. More specifically, following Todorov (2020), we estimate spot volatility by differencing the two estimators of volatility formed from the two available tenors as follows

\[
\hat{V}_{t,T}(\hat{u}_t) = -\frac{2}{(T_2 - T_1)\hat{u}_t^2} \left( \log |\hat{L}_{t,T_2}(\hat{u}_t)| - \log |\hat{L}_{t,T_1}(\hat{u}_t)| \right), \quad (47)
\]

where \(\hat{u}_t = \hat{u}_t^{(1)} \wedge \hat{u}_t^{(2)}\) with

\[
\hat{u}_t^{(1)} = \inf \left\{ u \geq 0 : |\hat{L}_{t,T_2}(u)| \leq 0.95 \right\}, \quad \overline{u}_t = \sqrt{-\frac{2\log(0.95)}{T_2\hat{\sigma}_t^2_{ATM}}}, \quad \hat{u}_t^{(2)} = \arg\min_{u \in [0,\pi]} |\hat{L}_{t,T_2}(u)|. \quad (48)
\]

We note that the tuning parameter \(\hat{u}_t\) is set in a slightly different way than for the original estimator in (6). The reason for this is that in the presence of a fixed time to discontinuity prior to the expiration of the option contract, the price increment is no longer of order \(o_p(1)\) but is \(O_p(1)\). The estimator \(\hat{V}_{t,T}(\hat{u}_t)\) above is a valid volatility estimator on regular days (without fixed times of discontinuity in the underlying prior to expiration) as well but it is far less efficient than the original one in (5).

C.5. Efficient Volatility Estimator

We denote the optimal volatility estimator that combines the return- and option-based volatility proxies with \(EV\). For its construction, we need estimates for the asymptotic variances of each of the two volatility proxies.

---

We start with the option-based one. Each volatility estimator \( \hat{V}_{t,\tau}(\hat{u}_{t,\tau}) \) is computed from options having time-to-maturity \( \tau \) and whose log-strikes are given by

\[
k_1 < k_2 < \ldots < k_{N_v},
\]

where in order to simplify notation, we suppressed dependence on \( t \) and \( \tau \) in the notation of the grid of log-strikes. Following Todorov (2019), the estimate for the asymptotic variance of the option-based volatility estimate is given by

\[
\hat{\text{Var}}(\hat{V}_{t,\tau}) = \frac{4}{\tau^2 \hat{u}_{t,\tau}^4 |\mathcal{E}_{t,\tau}(\hat{u}_{t,\tau})|^4} \left( \mathcal{R} \left( \mathcal{L}_{t,\tau}(\hat{u}_{t,\tau}) \right)^2 A_{t,\tau}^{(1)} + \mathcal{X} \left( \mathcal{L}_{t,\tau}(\hat{u}_{t,\tau}) \right)^2 A_{t,\tau}^{(2)} + 2 \mathcal{R} \left( \mathcal{L}_{t,\tau}(\hat{u}_{t,\tau}) \right) \mathcal{X} \left( \mathcal{L}_{t,\tau}(\hat{u}_{t,\tau}) \right) A_{t,\tau}^{(3)} \right),
\]

(49)

where with the notation

\[
f_t(u, k) = (u^2 + iu)e^{(iu-1)k-iux_t},
\]

(50)

we set

\[
A_{t,\tau}^{(1)} = \sum_{j=2}^{N_v} \mathcal{R}(f_t(\hat{u}_{\tau}, k_{j-1}))^2 \tilde{\sigma}_{t,\tau}(k_{j-1})^2 (k_{j-1})^2, \quad A_{t,\tau}^{(2)} = \sum_{j=2}^{N_v} \mathcal{X}(f_t(\hat{u}_{\tau}, k_{j-1}))^2 \tilde{\sigma}_{t,\tau}(k_{j-1})^2 (k_{j-1})^2,
\]

\[
A_{t,\tau}^{(3)} = \sum_{j=2}^{N_v} \mathcal{R}(f_t(\hat{u}_{\tau}, k_{j-1})) \mathcal{X}(f_t(\hat{u}_{\tau}, k_{j-1})) \tilde{\sigma}_{t,\tau}(k_{j-1})^2 (k_{j-1})^2,
\]

and we denote

\[
\tilde{\sigma}_{t,\tau}(k_j) = \sqrt{\frac{2}{3}} \left( \tilde{\sigma}_{t,\tau}(k_j) - \frac{1}{2} \tilde{\sigma}_{t,\tau}(k_{j-1}) - \frac{1}{2} \tilde{\sigma}_{t,\tau}(k_{j+1}) \right), \quad j = 2, \ldots, N_v - 1,
\]

(51)

with \( \sigma_{t,\tau}(k_1) = \sigma_{t,\tau}(k_2) \), and further for \( k_{j^*} \) being the log-strike closest (in absolute value) to \( x_t \), we set

\[
\tilde{\sigma}_{t,\tau}(k_{j^*}) = \tilde{\sigma}_{t,\tau}(k_{j^*}) - \tilde{\sigma}_{t,\tau}(k_{j^*}-1) - (\tilde{\sigma}_{t,\tau}(k_{j^*}-1) - \tilde{\sigma}_{t,\tau}(k_{j^*-2})) \frac{K_{j^*-1} - K_{j^*+2}}{K_{j^*+1} - K_{j^*+2}}, \quad \text{if } k_{j^*} \leq x_t,
\]

\[
\tilde{\sigma}_{t,\tau}(k_{j^*}) = \tilde{\sigma}_{t,\tau}(k_{j^*}) - \tilde{\sigma}_{t,\tau}(k_{j^*}+1) - (\tilde{\sigma}_{t,\tau}(k_{j^*+1}) - \tilde{\sigma}_{t,\tau}(k_{j^*+2})) \frac{K_{j^*} - K_{j^*+1}}{K_{j^*+1} - K_{j^*+2}}, \quad \text{if } k_{j^*} > x_t.
\]

(52)
Next, the estimate for the asymptotic variance of the return-based volatility estimate is given by (see e.g., Jacod and Protter (2011)):

$$\hat{Avar}(TV_t) = \frac{2}{3} \sum_{i=1}^{n} (\Delta_{t,i}^n x)^4 1\left(\left|\Delta_{t,i}^n x\right| \leq 3 \sqrt{BV_t \wedge RV_t \times n^{-1/2}}\right), t \in \mathbb{N}_+.$$ 

Using these estimates, the optimal volatility estimator $EV_t$ is given by (recall that we use options with two different times-to-maturity on each day):

$$EV_t = \omega_t^{(1)} \hat{V}_{t,\tau_1}(\hat{u}_{t,\tau_1}) + \omega_t^{(2)} \hat{V}_{t,\tau_2}(\hat{u}_{t,\tau_2}) + \omega_t^{(3)} TV_t,$$

where

$$\omega_t^{(1)} = \frac{\hat{Avar}(TV_t)\hat{Avar}(\hat{V}_{t,\tau_2})}{\hat{Avar}(\hat{V}_{t,\tau_1})\hat{Avar}(\hat{V}_{t,\tau_2}) + \hat{Avar}(TV_t)\hat{Avar}(\hat{V}_{t,\tau_1}) + \hat{Avar}(TV_t)\hat{Avar}(\hat{V}_{t,\tau_2})},$$

$$\omega_t^{(2)} = \frac{\hat{Avar}(TV_t)\hat{Avar}(\hat{V}_{t,\tau_1})}{\hat{Avar}(\hat{V}_{t,\tau_1})\hat{Avar}(\hat{V}_{t,\tau_2}) + \hat{Avar}(TV_t)\hat{Avar}(\hat{V}_{t,\tau_1}) + \hat{Avar}(TV_t)\hat{Avar}(\hat{V}_{t,\tau_2})},$$

and $\omega_t^{(3)} = 1 - \omega_t^{(1)} - \omega_t^{(2)}$. As seen from the above formula, higher asymptotic variance for a given volatility proxy naturally leads to smaller weight in the optimal volatility estimator $EV$. Importantly, the weights $\{\omega_t^{(j)}\}_{j=1,2,3}$ can vary over time, reflecting time-varying precision in the recovery of spot volatility from the various data sources.

**D. Testing for Conditional Independence**

For implementing the test for conditional independence, we form the empirical characteristic function:

$$\hat{\psi}_{RV,OV,JV}(u, v, z) = \frac{1}{T-h} \sum_{t=1}^{T-h} e^{iu \log(RV_t,t+h) + iv \log(OV_t) + iz \log(JV_t) + (1/2)(v z)\Xi_t(v z)^\top}, u, v, z \in \mathbb{R},$$

and using it, we define the following moment estimator

$$\hat{m}(u, z) = \int_{-v_T}^{v_T} \int_{-v_T}^{v_T} (\hat{\psi}_{RV,OV,JV}(u, v, z))\hat{\psi}_{RV,OV,JV}(0, w, 0) \psi_{RV,OV,JV}(0, w, 0) dv dw,$$

and

$$- \hat{\psi}_{RV,OV,JV}(u, v, 0)\hat{\psi}_{RV,OV,JV}(0, w, z)\hat{\psi}_{RV,OV,JV}(0, -w - v, 0) dv dw,$$

42
for $u, z \in \mathbb{R}$ and where $v_T$ is a sequence going to infinity that satisfies $v_T/\sqrt{T} \to 0$. We set $v_T$ in the following way

$$v_T = \frac{1}{10} \times \frac{T^{0.49}}{\text{std}(OV_t)}.$$

(56)

We denote a vector formed of moment conditions $\hat{m}(u, z)$, for several different values of the pair $(u, z)$, with $\hat{m}$. In implementing this test, we use two different values for $u$ and $z$, which results in 4 different vectors $(u, z)$ and a total of 8 moment conditions ($\hat{m}(u, z)$ are complex numbers). The two different values for $u$ and $z$ are such that the univariate characteristic functions of $RV_{t+h}$ and $JV_t$, respectively, reach values of 0.75 and 0.25.

We proceed with defining $\hat{\Sigma}_m$. We first introduce the following notation

$$\zeta_t(u, v, z) = e^{iu \log(RV_t, t+h) + iv \log(OV_t) + iz \log(JV_t)} \Xi_t(vz) - \hat{\psi}_{RV,OV,JV}(u, v, z),$$

(57)

and denote

$$\hat{m}_t(u, z) = \int_{-v_T}^{v_T} \int_{-v_T}^{v_T} \hat{m}_t(u, z; v, w) dv dw,$$

(58)

where

$$\hat{m}_t(u, z; v, w) = \zeta_t(u, v, z) \hat{\psi}_{OV}(w) \hat{\psi}_{OV}(-w - v) + \hat{\psi}_{RV,OV,JV}(u, v, z) \zeta_t(0, w, 0) \hat{\psi}_{OV}(-w - v)$$

$$+ \hat{\psi}_{RV,OV,JV}(u, v, z) \zeta_t(0, -w - v, 0) - \zeta_t(u, v, 0) \hat{\psi}_{OV}(w) \hat{\psi}_{OV}(-w - v)$$

$$- \hat{\psi}_{RV,OV}(u, v) \zeta_t(0, w, z) \hat{\psi}_{OV}(-w - v) - \hat{\psi}_{RV,OV}(u, v) \hat{\psi}_{OV,JV}(w, z) \zeta_t(0, -w - v, 0),$$

and we used the shorthand notation $\hat{\psi}_{OV}(v) = \hat{\psi}_{RV,OV,JV}(0, v, 0)$, $\hat{\psi}_{RV,OV}(u, v) = \hat{\psi}_{RV,OV,JV}(u, v, 0)$ and $\hat{\psi}_{OV,JV}(v, z) = \hat{\psi}_{RV,OV,JV}(0, v, z)$. Then, $\hat{\Sigma}_m$ is simply the long-run variance estimate of the moment vector, which we form using a Bartlett kernel:

$$\hat{\Sigma}_m = \hat{\gamma}_m(0) + \sum_{l=1}^{L_T} \left(1 - \frac{k}{L_T + 1}\right) (\hat{\gamma}_m(l) + \hat{\gamma}_m(l)^\top).$$

(59)

In our implementation, we set $L_T = 132$ (six months), which corresponds approximately to three months.
Finally, the matrix of asymptotic variances $\Xi_t$ is given by

$$\Xi_t = \frac{1}{4} \begin{pmatrix} \frac{1}{\hat{\sigma}_{t, t}} & 0 \\ 0 & \frac{1}{\hat{\sigma}_{t, \tau}} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \left( \Xi_{t, \tau_1} + \Xi_{t, \tau_2} \right) \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\hat{\sigma}_{t, \tau}} & 0 \\ 0 & \frac{1}{\hat{\sigma}_{t, \tau}} \end{pmatrix},$$

(60)

where $\Xi_{t, \tau}$ is a $2 \times 2$ matrix with the following entries

$$\Xi^{(1,1)}_{t, \tau} = \frac{1}{\tau^2} \sum_{j=2}^{N_\tau} g_t(k_{j-1})^2 \hat{\sigma}_{t, \tau}^2 (k_j - k_{j-1})^2, \quad \Xi^{(2,2)}_{t, \tau} = \widehat{\text{Var}}(\hat{\nu}_{t, \tau}),$$

(61)

and

$$\Xi^{(1,2)}_{t, \tau} = -\frac{2}{\tau^2 \hat{\sigma}_{t, \tau}^2 |\hat{L}_{t, \tau}(\hat{u}_{t, \tau})|^2} \left( \Re \left( \hat{L}_{t, \tau}(\hat{u}_{t, \tau}) \right) \hat{A}^{(1)}_{t, \tau} + \Im \left( \hat{L}_{t, \tau}(\hat{u}_{t, \tau}) \right) \hat{A}^{(2)}_{t, \tau} \right),$$

(62)

with the notation in (49) and (51)-(52) as well as

$$\begin{cases} 
\hat{A}^{(1)}_{t, \tau} = \sum_{j=2}^{N_\tau} \Re(f_t(\hat{u}_{t, \tau}, k_{j-1})) g_t(k_{j-1}) \hat{\sigma}_{t, \tau} (k_j - k_{j-1})^2 (k_j - k_{j-1})^2, \\
\hat{A}^{(2)}_{t, \tau} = \sum_{j=2}^{N_\tau} \Im(f_t(\hat{u}_{t, \tau}, k_{j-1})) g_t(k_{j-1}) \hat{\sigma}_{t, \tau} (k_j - k_{j-1})^2 (k_j - k_{j-1})^2,
\end{cases}$$

(63)

and

$$g_t(k) = 2e^{-k}(1 - k + x_t).$$

(64)

Since on a few days, the estimation of $\Xi_t$ can be unreliable (mainly due to low number of available strikes on the day), we exclude these days from the calculation of the test. More specifically, we drop days on which $v_T^2 \Xi^{(11)}_t > 2.$