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

Viktor Todorov† and Yang Zhang‡

June 29, 2021

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. We would like to thank Eric Ghysels (the Editor), three anonymous referees as well as seminar participants at various conferences for many useful comments and suggestions.

†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).[1] 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 risk 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 it 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 information in short-dated options not spanned by spot diffusive volatility 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-

[1] 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).

Scholes model can be backed out from the price of a single option written on the underlying 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, 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 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. Given that the risk premium component of the VIX 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.

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 forecasting.

The use of alternative volatility proxies for improving the measurement and forecasting of volatility adopted here is in part inspired by the approach of Ghysels et al. (2020) who show that measurement and forecasting of volatility from high-frequency return data can be improved upon by integrating return data across days, provided the volatility process is significantly persistent. The two approaches (ours and that of Ghysels et al. (2020)) share the common feature that the "measurement errors" in the two volatility proxies are uncorrelated (asymptotically). They differ, however, in the bias in the volatility proxies. In our case, shrinking time-to-maturity for the option estimator and a shrinking time window for the return-based volatility estimator will guarantee that the volatility proxies have asymptotically negligible biases. In the setting of Ghysels et al. (2020), on the other hand, past volatility proxies will have asymptotically negligible bias for today's volatility under the additional assumption of near unit root type dynamics for the volatility.

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 data 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. 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 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 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 special case of a MIDAS regression model (Ghysels et al. (2007)).[2] 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%

[2] Recent 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.

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. The weight assigned to the option-based volatility proxies in the volatility forecasts of the mixture models is well above 50% in the case of the S&P 500 index and is around 50% in the case of the individual stocks.

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 risk-neutral 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 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. In Section 5 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 6 concludes. The Appendix contains additional details on the data, the construction of the volatility measures and tests as well as volatility forecast results on an individual stock level.

2 Nonparametric Measurement of Volatility

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

$$x_t = x_0 + \int_0^t \mu_s ds + \int_0^t \sqrt{V_s} dW_s + \sum_{s \le t} \Delta x_s, \tag{1}$$

where μ_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 $\Delta_{t,i}^n x = x_{t-1+\kappa+(1-\kappa)i/n} - x_{t-1+(1-\kappa)(i-1)/n}$, for $t \in \mathbb{N}_+$ and $i = 1, ..., n$. Our first high-frequency volatility estimator is the intraday realized volatility defined as

$$RV_t = \sum_{i=1}^{n} (\Delta_{t,i}^n x)^2, \ t \in \mathbb{N}_+. \tag{2}$$

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

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

with $BV_t = \frac{\pi}{2} \frac{n}{n-1} \sum_{i=2}^n \sqrt{|\Delta_{t,i-1}^n x||\Delta_{t,i}^n x|}$ 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 a consistent and asymptotically normal 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).[4]

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 $\widehat{O}_{t,\tau}(k)$. The conditional characteristic function of the price increment under the risk-neutral measure $\mathbb{Q}$, denoted by $\mathcal{L}_{t,\tau}(u) = \mathbb{E}_t^{\mathbb{Q}}(e^{iu(x_{t+\tau}-x_t)})$, can be spanned by a portfolio of options in the following way

$$\mathcal{L}_{t,\tau}(u) = 1 - (u^2 + iu)\int_{\mathbb{R}} e^{(iu-1)k - iux_t} O_{t,\tau}(k)dk, \; u \in \mathbb{R}, \tag{4}$$

where in the above we have set the risk-free interest rate to zero (this is inconsequential as we apply this formula only for τ very close to zero). We denote the feasible counterpart of $\mathcal{L}_{t,\tau}(u)$ by $\widehat{\mathcal{L}}_{t,\tau}(u)$, 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

$$\widehat{V}_{t,\tau}(\widehat{u}_{t,\tau}) = -\frac{2}{\tau \widehat{u}_{t,\tau}^2} \log|\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})|, \tag{5}$$

[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 generally worse than the ones we report below for the daily TV.

[4] Recently, Bondarenko and Muravyev (2020) find that a significant part of the market risk premium is realized during the opening hours of European market trading before regular US stock trading begins. Our data for the S&P 500 index futures covers that important period, see the Appendix for more detail.

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

$$\widehat{u}_{t,\tau}^{(1)} = \inf\left\{u \geq 0 : |\widehat{\mathcal{L}}_{t,\tau}(u)| \leq 0.2\right\}, \ \overline{u}_{t,\tau} = \sqrt{-\frac{2\log(0.05)}{\tau\widehat{\sigma}_{t,ATM}^2}}, \ \ \widehat{u}_{t,\tau}^{(2)} = \mathrm{argmin}_{u\in[0,\overline{u}]}|\widehat{\mathcal{L}}_{t,\tau}(u)|, \tag{6}$$

for $\widehat{\sigma}_{t,ATM}$ denoting the Black-Scholes implied volatility of the option with strike closest to the current futures price at time t. The above choice of u aims to evaluate $\widehat{\mathcal{L}}_{t,\tau}(u)$ at a high value of its argument for which the inference for $\widehat{\mathcal{L}}_{t,\tau}(u)$ 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 $|\mathcal{L}_{t,\tau}(u)|$. 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

$$\widehat{\mathcal{L}}_{t,\tau}(u/\sqrt{\tau}) \approx \mathbb{E}_t^{\mathbb{Q}}\left(e^{-\frac{u^2}{2\tau}\int_t^{t+\tau} V_s ds}\right) \approx e^{-\frac{u^2}{2}V_t}, \text{ as } \tau \downarrow 0 \text{ and for fixed } u > 0. \tag{7}$$

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. The Monte Carlo evidence in Todorov (2019) shows that the main source of the approximation error in (7) is due to the separation of diffusive volatility from jumps while the time variation in volatility plays a significantly smaller role. As a result of the above result, $\widehat{V}_{t,\tau}(\widehat{u}_{t,\tau})$ provides a consistent estimate of true spot diffusive volatility despite of the fact that it is computed from options which are risk-neutral expectations of their payoffs.[5] We refer to Todorov (2019) for the formal statement of the above result and an associated CLT for $\widehat{V}_{t,\tau}(\widehat{u}_{t,\tau})$.

In Figure 1, we illustrate the decomposition of the estimate of $(u^2+iu)\int_{\mathbb{R}} e^{(iu-1)k-iux_t}O_{t,\tau}(k)dk$, 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 $\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})$ is formed from options whose strike is very close to the current spot spot 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 $\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})$ 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

[5]Recall that no-arbitrage implies that the diffusion coefficient of x is the same under $\mathbb{P}$ and $\mathbb{Q}$.

to the imaginary part of $\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})$, 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 $\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})$ is close to zero. This is in line with the approximation in (7) which is behind the option-based volatility estimator.

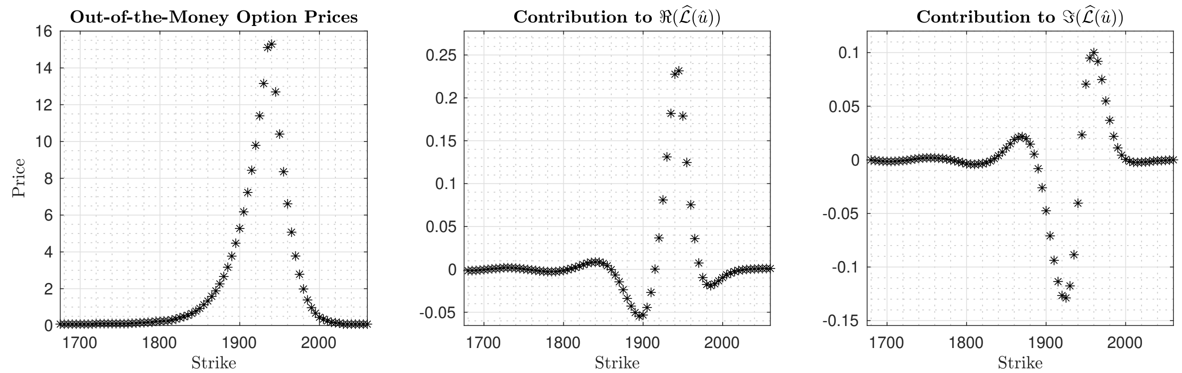

Figure 1: **Characteristic Function SPX Option Portfolio on February 1, 2016.** The left panel plots observed OTM option prices against their strikes. The next two panels plot their contributions to the estimate of $(u^2 + iu) \int_{\mathbb{R}} e^{(iu-1)k - iux_t} O_{t,\tau}(k) dk$. The time-to-maturity of the options is 4 trading days and the forward price level on the date was 1938.3.

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

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

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.[6]

[6]Pre-scheduled macroeconomic announcements tend to have much smaller impact on index options, for tenors considered here, compared to the effect earning announcements have typically on single-name equity options. For example, on days prior to FOMC announcements (with both expiration dates being after the announcement), the market index $\sqrt{\widehat{V}_{t,\tau_1}(\widehat{u}_{t,\tau_1})}$ has a mean which is only 5.5% higher than that of $\sqrt{\widehat{V}_{t,\tau_2}(\widehat{u}_{t,\tau_2})}$. By contrast, the mean of $\sqrt{\widehat{V}_{t,\tau_1}(\widehat{u}_{t,\tau_1})}$ is on average 22.5% higher than that of $\sqrt{\widehat{V}_{t,\tau_2}(\widehat{u}_{t,\tau_2})}$ on days prior to earning announcements for the individual stocks in our sample. For this reason, we do not perform the differencing of the near-by market index option-implied volatility estimates as such a procedure has poorer finite sample properties.

2.3 Combined Measures

Having two alternative estimates of spot volatility, it is natural to consider a combination of them to improve efficiency in the measurement of volatility. This will work if we have "diversification" of the measurement errors. This is indeed the case as shown formally in Todorov (2019). More specifically, we have the following hold approximately:

$$\widehat{V}_{t,\tau_i}(\widehat{u}_{t,\tau_i}) = V_t + \sqrt{\mathrm{Avar}(\widehat{V}_{t,\tau_i}(\widehat{u}_{t,\tau_i}))}\mathcal{Z}_{t,\tau_i},\ TV_t = V_t + \sqrt{\mathrm{Avar}(TV_t)}\mathcal{Z}_t,\ i = 1, 2, \tag{9}$$

where $\mathrm{Avar}(\widehat{V}_{t,\tau_i}(\widehat{u}_{t,\tau_i}))$ and $\mathrm{Avar}(TV_t)$ are $\mathcal{F}$-adapted (conditional) asymptotic variances in CLT-s for the two estimators (see Todorov (2019)) and $\{\mathcal{Z}_{t,\tau_1}\}_{t\geq 1}$, $\{\mathcal{Z}_{t,\tau_2}\}_{t\geq 1}$ and $\{\mathcal{Z}_t\}_{t\geq 1}$ are three independent i.i.d. sequences of standard normal random variables, defined on a product extension of the original probability space, and independent from $\mathcal{F}$ (and in particular from $\{V_t\}_{t\geq 1}$).[7] If we have estimates from the data for the asymptotic variances, denoted with $\{\widehat{\mathrm{Avar}}(\widehat{V}_{t,\tau_i}(\widehat{u}_{t,\tau_i}))\}_{i=1,2}$ and $\widehat{\mathrm{Avar}}(TV_t)$, then it is easy to see that the optimal estimate of spot diffusive volatility is

$$EV_t = \omega_t^{(1)}\widehat{V}_{t,\tau_1}(\widehat{u}_{t,\tau_1}) + \omega_t^{(2)}\widehat{V}_{t,\tau_2}(\widehat{u}_{t,\tau_2}) + \omega_t^{(3)}TV_t, \tag{10}$$

where

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

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

and $\omega_t^{(3)} = 1 - \omega_t^{(1)} - \omega_t^{(2)}$. The estimates for the asymptotic variances are given in the Appendix. 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.

The idea of improving on the volatility measurement by combining volatility proxies with uncorrelated measurement errors used above is in the spirit of Ghysels et al. (2020) who show that one can improve on the measurement of volatility today by "borrowing" information from the past,

[7] The independence of $\{\mathcal{Z}_{t,\tau_1}\}_{t\geq 1}$, $\{\mathcal{Z}_{t,\tau_2}\}_{t\geq 1}$ and $\{\mathcal{Z}_t\}_{t\geq 1}$ from $\mathcal{F}$ is because the CLT-s for $\widehat{V}_{t,\tau_i}(\widehat{u}_{t,\tau_i})$ and TV_t hold stably, see Todorov (2019) for details.

provided the volatility process is persistent so that the potential bias in doing this is small. The difference between the current paper and Ghysels et al. (2020) is that here the volatility proxies use different sources of information, underlying asset returns and options, while in Ghysels et al. (2020) inference is based only on returns. In the same spirit as Ghysels et al. (2020), one can also consider potential further improvements of EV_t by utilizing past return and option data in the inference.

2.4 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 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.[8] 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.[9]

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.[10] Additional details on the processing of the raw data and the computation of the volatility measures are provided in the Appendix.

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

[8] 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.

[9] The individual stock options are American style while our derivation above of OV_t is based on European-style options. One can correct for the early exercise premium using various approaches, see e.g., Broadie and Detemple (1996) and references therein. We do not do this here, as the early exercise premium tends to be small for short-dated out-of-the-money options which are used in our analysis, see e.g., Bakshi et al. (2003).

[10] 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.

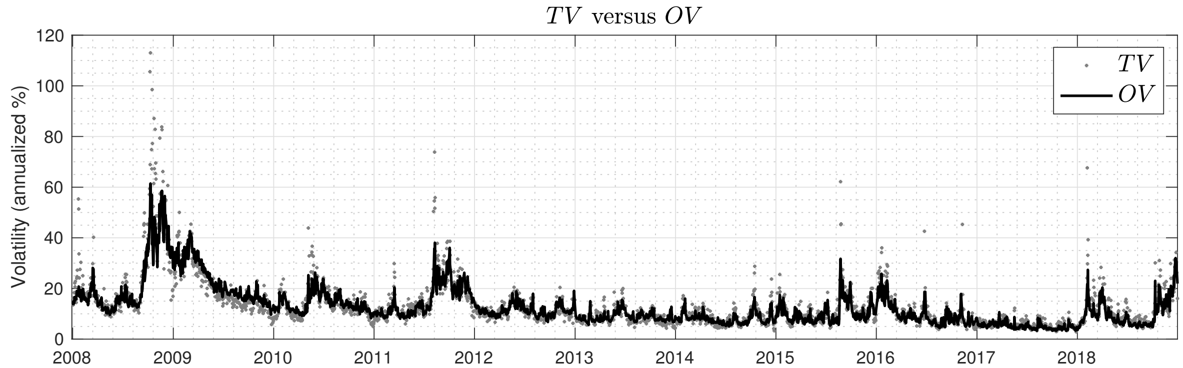

Figure 2: **Volatility Estimates for the S&P 500 Index.** The figure displays the daily OV and TV estimates for the S&P 500 index from 2008 to 2018, where OV is the solid line and the high-frequency measure is plotted with grey dots.

of high volatility where one needs to use a higher threshold.[11] 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 that 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 and the combined EV estimators are both significantly less noisy than TV. 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.[12] By comparing the medians of OV with RV, we can conclude that jump risk contributes a nontrivial 30% in the total return variation. The difference between OV and EV is relatively small, with EV having a smaller standard deviation than OV. 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.[13] 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

[11] The sensitivity of the truncated variation estimator to the choice of the threshold level suggests that realized jumps are of smaller size and their separation from diffusive moves is difficult at the frequency of five-minute.

[12] The approximation in (7) shows that OV is an expectation of a (function of) forward looking integrated volatility which will naturally make it smoother than the latent spot volatility itself due to the mean reversion in volatility. However, the time-to-maturity of the options used in our analysis is very short and therefore this smoothing effect should be relatively small.

[13] 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

	Mean	STD	STD(Δ)	Q25	Median	Q75	AR(1)
RV	0.0396	0.0965	0.0724	0.0083	0.0161	0.0354	0.7188
TV	0.0281	0.0673	0.0458	0.0057	0.0113	0.0248	0.7693
OV	0.0222	0.0351	0.0121	0.0058	0.0108	0.0233	0.9404
$BSIV$	0.0363	0.0546	0.0185	0.0100	0.0188	0.0397	0.9427
VIX	0.0480	0.0714	0.0238	0.0140	0.0252	0.0501	0.9443
EV	0.0211	0.0331	0.0123	0.0056	0.0104	0.0219	0.9304
JV	0.0237	0.0371	0.0169	0.0078	0.0131	0.0235	0.8964

Panel B: Individual Stocks

Ticker	#K	RV	TV	OV	$BSIV$	VIX	JV	Ticker	#K	RV	TV	OV	$BSIV$	VIX	JV
AAPL	35	0.0394	0.0323	0.0393	0.0530	0.0638	0.0232	KO	11	0.0172	0.0139	0.0134	0.0200	0.0239	0.0116
AXP	19	0.0264	0.0212	0.0255	0.0379	0.0457	0.0236	MCD	15	0.0189	0.0146	0.0149	0.0218	0.0271	0.0133
BA	22	0.0404	0.0310	0.0325	0.0464	0.0549	0.0243	MMM	14	0.0233	0.0184	0.0171	0.0290	0.0355	0.0205
CAT	22	0.0526	0.0403	0.0411	0.0574	0.0662	0.0265	MRK	16	0.0279	0.0226	0.0209	0.0303	0.0363	0.0176
CSCO	11	0.0483	0.0386	0.0286	0.0458	0.0531	0.0240	MSFT	17	0.0371	0.0297	0.0278	0.0400	0.0473	0.0198
CVX	19	0.0301	0.0249	0.0229	0.0344	0.0412	0.0188	NKE	19	0.0375	0.0291	0.0277	0.0390	0.0454	0.0200
DIS	20	0.0310	0.0252	0.0253	0.0372	0.0458	0.0200	PFE	10	0.0276	0.0225	0.0195	0.0295	0.0348	0.0186
GE	10	0.0411	0.0333	0.0296	0.0476	0.0594	0.0314	PG	17	0.0183	0.0148	0.0127	0.0197	0.0249	0.0137
GS	18	0.0401	0.0314	0.0377	0.0516	0.0599	0.0241	T	12	0.0234	0.0185	0.0172	0.0262	0.0316	0.0161
HD	17	0.0281	0.0220	0.0232	0.0337	0.0403	0.0188	UTX	16	0.0272	0.0214	0.0211	0.0326	0.0391	0.0222
IBM	16	0.0259	0.0207	0.0185	0.0267	0.0327	0.0153	VZ	15	0.0242	0.0189	0.0172	0.0256	0.0318	0.0171
INTC	12	0.0465	0.0372	0.0346	0.0500	0.0583	0.0249	V	17	0.0284	0.0221	0.0275	0.0395	0.0478	0.0217
JNJ	16	0.0159	0.0126	0.0119	0.0177	0.0224	0.0118	WMT	17	0.0220	0.0178	0.0163	0.0232	0.0276	0.0132
JPM	20	0.0368	0.0294	0.0283	0.0417	0.0494	0.0224	XOM	19	0.0249	0.0203	0.0184	0.0276	0.0330	0.0157

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 measures, together with the standard deviation of the first-order difference (STD(ΔV)) 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.

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. For the estimation of the marginal volatility distribution from the high-frequency returns, we compute the empirical characteristic function of the high-frequency returns and then perform a regularized inversion as done in Todorov and Tauchen (2012) (see Christensen et al. (2019) for an alternative approach). 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 implies 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. We do not compute EV for the individual stocks as the estimates of the asymptotic variances needed for it are very noisy, given the significantly smaller number of available short-dated options. Unlike the S&P 500 index case, the overnight period where there is no high-frequency record of the asset price, is rather nontrivial for individual stock data. 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 κ (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 κ, 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 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

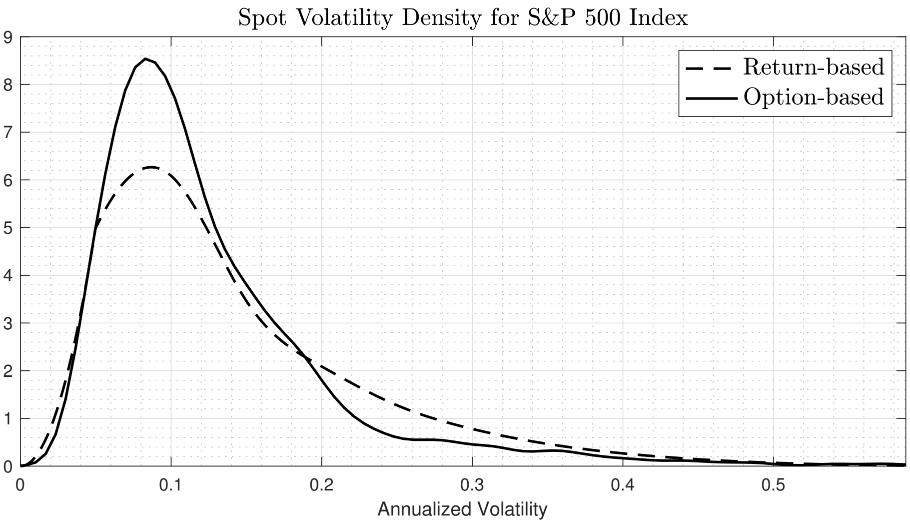

Figure 3: **Spot Volatility Density Estimates for S&P 500 Index.** The solid line corresponds to kernel density estimate of annualized $\sqrt{OV_t}$, where we use Epanechnikov kernel and the Silverman's automatic bandwidth of $h = 0.79 \times IQR(\sqrt{OV_t}) \times T^{-1/5}$, with IQR denoting the inter-quartile 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), with regularization parameter set to $R = 1.6$.

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 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 forecasting future volatility, a point which is also made in Ghysels et al. (2020). We now illustrate this formally. Suppose that we are interested in forecasting a random variable Z_{t+1} using two observable vectors $\widehat{\boldsymbol{V}}_{t,1}$ and $\widehat{\boldsymbol{V}}_{t,2}$ given by

$$\widehat{\boldsymbol{V}}_{t,1} = \boldsymbol{V}_t + \boldsymbol{\Sigma}_{t,1} \times \boldsymbol{\epsilon}_{t,1} \quad \text{and} \quad \widehat{\boldsymbol{V}}_{t,2} = \boldsymbol{V}_t + \boldsymbol{\Sigma}_{t,2} \times \boldsymbol{\epsilon}_{t,2}, \tag{11}$$

where $\boldsymbol{V}_t$, $\boldsymbol{\epsilon}_{t,1}$ and $\boldsymbol{\epsilon}_{t,2}$ are $N \times 1$ random vectors, and $\boldsymbol{\Sigma}_{t,1}$ and $\boldsymbol{\Sigma}_{t,2}$ are two diagonal $N \times N$ matrices. The observation error vectors, $\boldsymbol{\epsilon}_{t,1}$ and $\boldsymbol{\epsilon}_{t,2}$, are independent of each other as well as of $\boldsymbol{V}_t$, $\boldsymbol{\Sigma}_{t,1}$ and $\boldsymbol{\Sigma}_{t,2}$, and Z_{t+1}, and satisfy

$$\mathbb{E}(\boldsymbol{\epsilon}_{t,1}) = \mathbb{E}(\boldsymbol{\epsilon}_{t,2}) = \mathbf{0}_N, \ \operatorname{var}(\boldsymbol{\epsilon}_{t,1}) = \operatorname{var}(\boldsymbol{\epsilon}_{t,2}) = \mathbf{I}_N. \tag{12}$$

We will assume further that $\{(Z_t, \boldsymbol{V}_t^\top)^\top\}_{t \geq 1}$ is a covariance stationary process, and for simplicity of exposition that Z_t and $\boldsymbol{V}_t$ have both mean zero (i.e., we have de-meaned the variables of interest). In addition, $\mathbb{E}(\boldsymbol{\Sigma}_{t,1}^2) = \sigma_1^2 \mathbf{I}_N$ and $\mathbb{E}(\boldsymbol{\Sigma}_{t,2}^2) = \sigma_2^2 \mathbf{I}_N$, for some constants σ_1 and σ_2.

In our setting we can think of Z_{t+1} as being the value of a univariate 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) and equation (9) above, implies that they satisfy approximately the specification in (11)-(12). More specifically, consider $\widehat{\boldsymbol{V}}_{t,1} = (TV_t, TV_{t-1}, ..., TV_{t-N+1})^\top$ and $\widehat{\boldsymbol{V}}_{t,2} = (OV_t, OV_{t-1}, ..., OV_{t-N+1})^\top$, and recall that TV_t and OV_t are daily univariate estimates. In this specific case, we have $\boldsymbol{V}_t = (V_t, V_{t-1}, ..., V_{t-N+1})^\top$. Further, by noting that $\mathcal{Z}_{t,\tau_i}$ and $\mathcal{Z}_t$ in (9) are independent i.i.d. sequences, defined on an extension of the original probability space and independent from it, we have

$$\begin{aligned}
\boldsymbol{\Sigma}_{t,1}^2 &= \operatorname{diag}\left(\operatorname{Avar}(TV_t), ..., \operatorname{Avar}(TV_{t-N+1})\right), \\
\boldsymbol{\Sigma}_{t,2}^2 &= \frac{1}{2} \sum_{i=1}^{2} \operatorname{diag}\left(\operatorname{Avar}(\widehat{V}_{t,\tau_i}(\widehat{u}_{t,\tau_i})), ..., \operatorname{Avar}(\widehat{V}_{t-N+1,\tau_i}(\widehat{u}_{t-N+1,\tau_i}))\right).
\end{aligned} \tag{13}$$

Our goal is to characterize the optimal linear forecast of Z_{t+1}, with respect to a square loss function, using the various volatility proxies constructed from $\widehat{\boldsymbol{V}}_{t,1}$ and $\widehat{\boldsymbol{V}}_{t,2}$. We will first consider the forecasts based on linear combinations of $\widehat{\boldsymbol{V}}_{t,1}$ and $\widehat{\boldsymbol{V}}_{t,2}$, i.e., the case in which $\boldsymbol{\Sigma}_{t,1}$ and $\boldsymbol{\Sigma}_{t,2}$ are unknown. As is well known, this optimal forecast is given by the linear projection which we will denote with $\widehat{\mathbb{E}}(Z|\mathbf{X})$ when using a generic vector $\mathbf{X}$ to predict a generic random variable Z. We will first show that

$$\widehat{\mathbb{E}}(Z_{t+1}|\widehat{\boldsymbol{V}}_{t,1}, \widehat{\boldsymbol{V}}_{t,2}) = \widehat{\mathbb{E}}(Z_{t+1}|\widehat{\boldsymbol{V}}_{t,mix}), \tag{14}$$

where we denote

$$\widehat{\boldsymbol{V}}_{t,mix} = \left(\sigma_1^2 + \sigma_2^2\right)^{-1} \sigma_2^2 \widehat{\boldsymbol{V}}_{t,1} + \left(\sigma_1^2 + \sigma_2^2\right)^{-1} \sigma_1^2 \widehat{\boldsymbol{V}}_{t,2}. \tag{15}$$

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 $\widehat{\boldsymbol{V}}_{t,mix}$. Note that each element of $\widehat{\boldsymbol{V}}_{t,mix}$ puts the same weight to the corresponding elements of $\widehat{\boldsymbol{V}}_{t,1}$ and $\widehat{\boldsymbol{V}}_{t,2}$. To establish the above result, suppose that $\widehat{\mathbb{E}}(Z_{t+1}|\widehat{\boldsymbol{V}}_{t,mix})$ is given by

$$\widehat{\mathbb{E}}(Z_{t+1}|\widehat{\boldsymbol{V}}_{t,mix}) = \boldsymbol{\alpha}^\top \widehat{\boldsymbol{V}}_{t,mix}, \tag{16}$$

for some vector α of weights that satisfies

$$\mathbb{E}\left[(Z_{t+1} - \boldsymbol{\alpha}^\top \widehat{\boldsymbol{V}}_{t,mix})\widehat{\boldsymbol{V}}_{t,mix}\right] = \mathbf{0}_N. \tag{17}$$

To show the result in (14), we need to establish that the following two conditions hold

$$\mathbb{E}\left[(Z_{t+1} - \boldsymbol{\alpha}^\top \widehat{\boldsymbol{V}}_{t,mix})\widehat{\boldsymbol{V}}_{t,1}\right] = \mathbf{0}_N \text{ and } \mathbb{E}\left[(Z_{t+1} - \boldsymbol{\alpha}^\top \widehat{\boldsymbol{V}}_{t,mix})\widehat{\boldsymbol{V}}_{t,2}\right] = \mathbf{0}_N. \tag{18}$$

First, note that

$$\mathbb{E}((\boldsymbol{\Sigma}_{t,1}\boldsymbol{\epsilon}_{t,1} - \boldsymbol{\Sigma}_{t,2}\boldsymbol{\epsilon}_{t,2})(\widehat{\boldsymbol{V}}_{t,mix} - \boldsymbol{V}_t)^\top) = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\mathbb{E}(\boldsymbol{\Sigma}_{t,1}\boldsymbol{\Sigma}_{t,1}^\top) - \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\mathbb{E}(\boldsymbol{\Sigma}_{t,2}\boldsymbol{\Sigma}_{t,2}^\top) = \mathbf{0}_{N\times N}, \tag{19}$$

and

$$\widehat{\boldsymbol{V}}_{t,1} - \widehat{\boldsymbol{V}}_{t,mix} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}(\boldsymbol{\Sigma}_{t,1}\boldsymbol{\epsilon}_{t,1} - \boldsymbol{\Sigma}_{t,2}\boldsymbol{\epsilon}_{t,2}),\ \widehat{\boldsymbol{V}}_{t,2} - \widehat{\boldsymbol{V}}_{t,mix} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}(\boldsymbol{\Sigma}_{t,2}\boldsymbol{\epsilon}_{t,2} - \boldsymbol{\Sigma}_{t,1}\boldsymbol{\epsilon}_{t,1}). \tag{20}$$

Since $\boldsymbol{\epsilon}_{t,1}$ and $\boldsymbol{\epsilon}_{t,2}$ are both independent from Z_{t+1} and $\boldsymbol{V}_t$, we have $\mathbb{E}[(Z_{t+1} - \boldsymbol{\alpha}^\top \boldsymbol{V}_t)(\boldsymbol{\epsilon}_{t,1} - \boldsymbol{\epsilon}_{t,2})] = \mathbf{0}_N$. This, combined with the above two results, yields:

$$\mathbb{E}\left[(Z_{t+1} - \boldsymbol{\alpha}^\top \widehat{\boldsymbol{V}}_{t,mix})(\widehat{\boldsymbol{V}}_{t,1} - \widehat{\boldsymbol{V}}_{t,mix})\right] = \mathbf{0}_N \text{ and } \mathbb{E}\left[(Z_{t+1} - \boldsymbol{\alpha}^\top \widehat{\boldsymbol{V}}_{t,mix})(\widehat{\boldsymbol{V}}_{t,2} - \widehat{\boldsymbol{V}}_{t,mix})\right] = \mathbf{0}_N, \tag{21}$$

and hence (18), which in turn implies (14). From here, for the expected losses from using $\widehat{\boldsymbol{V}}_{t,1}$ and $\widehat{\boldsymbol{V}}_{t,2}$ separately and in combination in constructing linear forecasts for Z_{t+1}, we can write

$$\mathbb{E}\left(Z_{t+1} - \widehat{\mathbb{E}}(Z_{t+1}|\widehat{\boldsymbol{V}}_{t,i})\right)^2 = \mathbb{E}(Z_{t+1}^2) - \mathbb{E}(Z_{t+1}\boldsymbol{V}_t^\top)\left(\mathbb{E}(\boldsymbol{V}_t\boldsymbol{V}_t^\top) + \sigma_i^2\mathbf{I}_N\right)^{-1}\mathbb{E}(Z_{t+1}\boldsymbol{V}_t),\ i = 1, 2, \tag{22}$$

$$\mathbb{E}\left(Z_{t+1}-\widehat{\mathbb{E}}(Z_{t+1}|\widehat{\boldsymbol{V}}_{t,1},\widehat{\boldsymbol{V}}_{t,2})\right)^2=\mathbb{E}(Z_{t+1}^2)-\mathbb{E}(Z_{t+1}\boldsymbol{V}_t^\top)\left(\mathbb{E}(\boldsymbol{V}_t\boldsymbol{V}_t^\top)+\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mathbf{I}_N\right)^{-1}\mathbb{E}(Z_{t+1}\boldsymbol{V}_t). \tag{23}$$

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 $\boldsymbol{V}_t$ and Z_{t+1}. Can we do better? Suppose we know $\boldsymbol{\Sigma}_{t,1}$ and $\boldsymbol{\Sigma}_{t,2}$. In this case, exactly as in Section 2.3, we can construct the optimal estimate of $\boldsymbol{V}_t$ on each day by weighting appropriately the two volatility proxies:

$$\widehat{\boldsymbol{V}}_{t,opt}=\boldsymbol{\omega}_t\widehat{\boldsymbol{V}}_{t,1}+(\mathbf{I}_N-\boldsymbol{\omega}_t)\widehat{\boldsymbol{V}}_{t,2},\ \boldsymbol{\omega}_t=\left(\boldsymbol{\Sigma}_{t,1}^2+\boldsymbol{\Sigma}_{t,2}^2\right)^{-1}\boldsymbol{\Sigma}_{t,2}^2. \tag{24}$$

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

$$\mathbb{E}\left(Z_{t+1}-\widehat{\mathbb{E}}(Z_{t+1}|\widehat{\boldsymbol{V}}_{t,opt})\right)^2=\mathbb{E}(Z_{t+1}^2)-\mathbb{E}(Z_{t+1}\boldsymbol{V}_t^\top)\left(\mathbb{E}(\boldsymbol{V}_t\boldsymbol{V}_t^\top)+\mathbb{E}(\boldsymbol{\omega}_t\boldsymbol{\omega}_t^\top)\right)^{-1}\mathbb{E}(Z_{t+1}\boldsymbol{V}_t). \tag{25}$$

By Jensen's inequality, the difference

$$\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mathbf{I}_N-\mathbb{E}(\boldsymbol{\omega}_t\boldsymbol{\omega}_t^\top), \tag{26}$$

is positive semidefinite. Therefore, $\widehat{\boldsymbol{V}}_{t,opt}$ provides more efficient forecast than $\widehat{\boldsymbol{V}}_{t,mix}$, with the size of the efficiency gains depending on the time variation in $\boldsymbol{\omega}_t$. Obviously, the same result will continue to hold even when the vectors $\widehat{\boldsymbol{V}}_{t,1}$ and $\widehat{\boldsymbol{V}}_{t,2}$ 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 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 for the S&P 500 index with the results for the individual stocks reported in the Appendix.

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:

$$RV_{t,t+h} = \beta_{0,h} + \beta_{1,h} V_{t,d} + \beta_{2,h} V_{t,w} + \beta_{3,h} V_{t,m} + \varepsilon_{t+h}, \tag{27}$$

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 (27) 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 the combined EV.[14]

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:

$$\begin{aligned} RV_{t,t+h} = \beta_{0,h} &+ \beta_{1,h} OV_{t,d} + \beta_{2,h} OV_{t,w} + \beta_{3,h} OV_{t,m} \\ &+ \gamma_{1,h} TV_{t,d} + \gamma_{2,h} TV_{t,w} + \gamma_{2,m} TV_{t,m} + \varepsilon_{t+h}. \end{aligned} \tag{28}$$

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 loose the advantage of assigning time-varying weights to the return- and option-based volatility estimates. This tradeoffs will be assessed empirically below.

[14] 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.

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

$$RV_{t,t+h} = \beta_{0,h} + (\beta_{1,h} + \beta_{1Q,h} RQ_t^{1/2}) RV_{t,d} + \beta_{2,h} RV_{t,w} + \beta_{3,h} RV_{t,m} + \varepsilon_{t+h}, \tag{29}$$

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.[15]

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 (Huber (1964)), 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:

$$\text{MSE}(Y, \widehat{Y}) = (Y - \widehat{Y})^2, \quad \text{QLIKE}(Y, \widehat{Y}) = \frac{Y}{\widehat{Y}} - \log\left(\frac{Y}{\widehat{Y}}\right) - 1, \tag{30}$$

[15]In an analogy to the HARQ model, we also experimented with augmenting the HAR-EV model by including the term $EV_t \times \sqrt{\widehat{Avar}(EV_t)}$, for $\widehat{Avar}(EV_t)$ 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.

where Y denotes the generic variable to be predicted and $\widehat{Y}$ the candidate forecast.[16] The above loss functions differ in the weight they put on forecast errors of different size and/or relation to the object that is forecasted.

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 stochastic volatility models with price jumps in which the jump intensity is linear in the diffusive spot volatility with no volatility jumps, 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 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.

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

[16] 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.

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

Panel A: One-day Ahead Relative Forecasting Loss

Window	Criterion	HAR-RV	HAR-TV	HARQ	HAR-OV	HAR-MV	HAR-EV
Rolling Window	MSE (Median)	0.9753	1.0000	0.8769	0.6857	0.6680	0.6182
	MSE (Mean)	0.9847	1.0000	1.0483	0.8927	0.8637	0.8665
	QLIKE (Median)	0.9502	1.0000	0.8904	0.7315	0.6934	0.7107
	QLIKE (Mean)	0.9695	1.0000	1.4667	0.9267	0.8152	0.8528
Increasing Window	MSE (Median)	1.0317	1.0000	0.7435	0.6186	0.5797	0.5656
	MSE (Mean)	0.9646	1.0000	0.9655	0.9016	0.8799	0.8740
	QLIKE (Median)	1.0065	1.0000	0.7935	0.7335	0.6925	0.7426
	QLIKE (Mean)	0.9980	1.0000	0.9005	1.0377	0.7729	0.9245

Panel B: One-week Ahead Relative Forecasting Loss

Window	Criterion	HAR-RV	HAR-TV	HARQ	HAR-OV	HAR-MV	HAR-EV
Rolling Window	MSE (Median)	1.0411	1.0000	0.8623	0.6522	0.6389	0.6133
	MSE (Mean)	1.0161	1.0000	0.9715	0.8955	0.8728	0.8818
	QLIKE (Median)	0.9686	1.0000	0.8531	0.7160	0.6953	0.7087
	QLIKE (Mean)	0.9966	1.0000	1.0357	0.9121	0.8019	0.8814
Increasing Window	MSE (Median)	0.9945	1.0000	0.8770	0.7246	0.6513	0.6559
	MSE (Mean)	0.9920	1.0000	0.9970	0.8432	0.8653	0.8227
	QLIKE (Median)	0.9670	1.0000	0.9078	0.8794	0.7522	0.8205
	QLIKE (Mean)	0.9988	1.0000	0.9722	1.0954	0.8600	1.0007

Panel C: One-month Ahead Relative Forecasting Loss

Window	Criterion	HAR-RV	HAR-TV	HARQ	HAR-OV	HAR-MV	HAR-EV
Rolling Window	MSE (Median)	0.9759	1.0000	0.8593	0.7169	0.8331	0.7649
	MSE (Mean)	1.0068	1.0000	0.9791	0.9851	0.9523	0.9856
	QLIKE (Median)	0.9658	1.0000	0.8716	0.8533	0.8654	0.8641
	QLIKE (Mean)	1.0050	1.0000	0.9904	0.9867	0.9146	0.9903
Increasing Window	MSE (Median)	0.9878	1.0000	0.9245	0.6183	0.8846	0.6997
	MSE (Mean)	0.9903	1.0000	1.0051	0.9021	0.9713	0.9078
	QLIKE (Median)	0.9407	1.0000	0.9031	0.7491	0.8974	0.8027
	QLIKE (Mean)	1.0007	1.0000	0.9950	1.0179	0.9796	1.0000

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.

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.

Overall, the results above show that for the purposes of volatility forecasting, it is best to combine the return- and option-based volatility proxies, with the size of the forecasting gains over return-based models being statistically significant. To get a sense of how much the option and return data contribute to the volatility forecast in the HAR-MV model, we compute the ratio $\frac{||\boldsymbol{\beta}_h||}{||\boldsymbol{\beta}_h||+||\boldsymbol{\gamma}_h||}$, where $\boldsymbol{\beta}_h = (\beta_{1,h}, \beta_{2,h}, \beta_{3,h})$ and $\boldsymbol{\gamma}_h = (\gamma_{1,h}, \gamma_{2,h}, \gamma_{3,h})$, and $||\cdot||$ denotes the Euclidian norm. For the horizon of one day and when using rolling estimation window, the mean value of this ratio is 0.74 (the results for the increasing estimation window and the weekly and monthly horizons are similar). That is, the option-based volatility proxy plays the leading role in generating the volatility forecast. This is in line with the evidence in Section 2.4 for the higher precision of OV versus TV. Similarly, the average weight in the EV variable, defined in (10), that is assigned to the option-based volatility proxy (i.e., $\omega_t^{(1)} + \omega_t^{(2)}$ in (10)) is 0.83.[17] We note that this weight,

[17]Note that option data enters slightly differently in OV and EV. Mainly, OV assigns equal weight to the volatility estimates from the two maturities while EV weights optimally these two option-based volatility proxies.

which recall is computed on a daily basis, varies a lot over the sample. This is in part a reflection of the time-varying quality of the option data as evident for example from the time-variation in the available options to compute the option-based volatility proxy from, see Figure 6 in the Appendix.

We end this section by 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 for future total return volatility which is the variable that is forecasted. That is, in the forecasting models introduced above, we replace the future $RV_{t,t+h}$ with $ORV_{t,t+h}$ which we introduce below. 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

$$ORV_t = EV_t + RV_t - TV_t. \tag{31}$$

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.[18] 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.

Table 3: ORV Forecasting Results for the S&P 500 Index

		One-day Ahead		One-week Ahead		One-month Ahead	
Window	Criterion	HAR-MV	HAR-EV	HAR-MV	HAR-EV	HAR-MV	HAR-EV
Rolling Window	MSE (Mean)	0.7317	0.7136	0.7395	0.7209	0.8959	0.9083
	QLIKE (Mean)	0.6007	0.5766	0.7268	0.7299	0.8916	0.9209
Increasing Window	MSE (Mean)	0.6778	0.6650	0.6842	0.6731	0.7832	0.7600
	QLIKE (Mean)	0.5158	0.5073	0.6386	0.6321	0.8232	0.8165

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 when forecasting future ORV.

[18] An alternative is to use intraday option data and recover EV intraday.

5 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.

5.1 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.[19] Suppose that the risk-neutral jump compensator of x is given by $\eta_t dt \otimes F(dx)$,[20] for some time-invariant measure F capturing the distribution of the jump size and a (predictable) stochastic process η_t capturing the time variation in the jump compensator.[21] 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 process η_t. That is, the options are functions of V_t and η_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 compensator. Since here we are interested only in the predictive content of the option data, it suffices to identify η_t up to a constant. 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:

$$\widehat{J}_{t,\tau} = \frac{2}{\tau}\sum_{j=2}^{N_\tau} e^{-k_{j-1}}(1-k_{j-1}+x_t)\widehat{O}_{t,\tau}(k_{j-1})(k_j-k_{j-1}) - \widehat{V}_{t,\tau}(\widehat{u}_{t,\tau}), \tag{32}$$

[19] Our focus in this section is the information content of the option data. For this reason, we do not consider forecasting models in which jump variation measures constructed from returns are used as predictors. The risk-neutral jump variation measures and the ones constructed from returns have different time series properties, see e.g., Andersen et al. (2015). Moreover, return-based jump variation measures have limited time series persistence and the gains from their inclusion in volatility forecasting regressions tend to be small, see e.g., Andersen et al. (2007). This is also indirectly confirmed here by the small differences in the performance of the HAR-RV and HAR-TV models.

[20] 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 stohastic process (possibly different for positive and negative jumps) that does not depend on the jump size and a time-invariant measure of the jump size.

[21] 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 $F(\mathbb{R}) = \infty$ (infinite activity jumps).

and from here by averaging across the two available tenors, we get $JV_t = \frac{1}{2}\left(\widehat{J}_{t,\tau_1} + \widehat{J}_{t,\tau_2}\right)$. The role of the deep out-of-the-money options in $\widehat{J}_{t,\tau}$ is bounded by the weight assigned to them in the portfolio spanning the second risk-neutral moment of returns given by the portfolio, $\frac{2}{\tau}\sum_{j=2}^{N_\tau} e^{-k_{j-1}}(1 - k_{j-1} + x_t)\widehat{O}_{t,\tau}(k_{j-1})(k_j - k_{j-1})$, see the proof of Theorem 1 in Todorov (2019).

It is easy to show that

$$JV_t \approx \int_{\mathbb{R}} x^2 F(dx) \times \eta_t, \quad \text{as } \tau_1 \downarrow 0 \text{ and } \tau_2 \downarrow 0. \tag{33}$$

Note that, in general, we cannot identify separately η_t and the measure F as they enter in a product form in the jump compensator. Therefore, some normalization of F is needed. A common one would be to assume that $F(\mathbb{R}) = 1$, i.e., that F is a probability distribution and in this case η_t can be interpreted as the jump intensity capturing the probability of jump arrival. However, this normalization is obviously not appropriate when jumps are of infinite activity, i.e., when $F(\mathbb{R}) = \infty$. An alternative one that will work regardless of the jump activity is to set $\int_{\mathbb{R}} x^2 F(dx) = 1$ in which case JV_t is an estimate of η_t. For the analysis here, we only need η_t up to a constant as a multiple of η_t will obviously carry the same predictive content. That said, the short-dated options contain information to identify completely the jump compensator $\eta_t dt \otimes F(dx)$, see Qin and Todorov (2019).

In Figure 4, we plot the market index OV and JV series and the cross-sectional averages of these measures for the DJIA stocks while summary statistics for JV are reported in Table 1. 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.[22] This is in line with parametric models in which the jump intensity is modeled as an affine function of the volatility (and its factors), see e.g., Bates (2000). Nevertheless differences in the time series behavior of OV and JV do appear and our formal results later on will confirm that. 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

[22]Note that our separation of the volatility from jumps is based on their different role in the prices of short-dated options with various strikes. This separation does not rely on differences in time series behavior of the jump compensator and the spot diffusive volatility.

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.

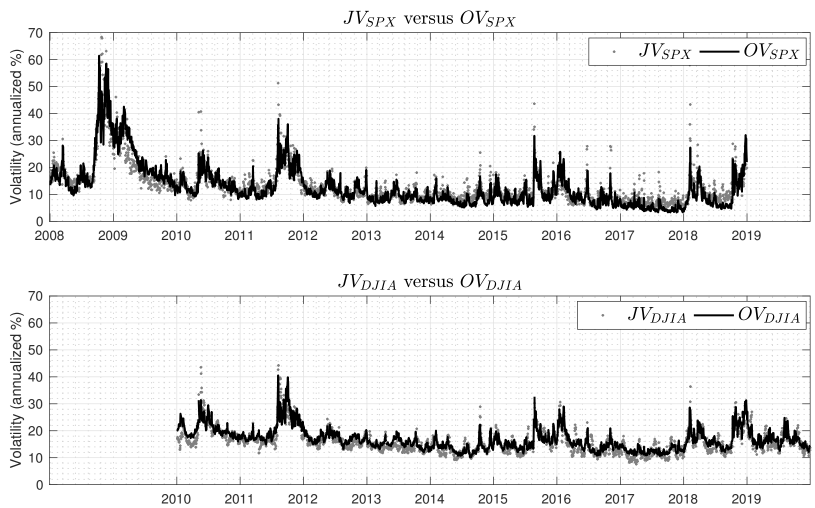

Figure 4: **Option-Based Volatility Measures.** Solid lines on the plots correspond to OV while dotted lines correspond to JV. Top panel displays estimates for the S&P 500 index while bottom panel displays cross-sectional averages of the option estimates for the DJIA stocks.

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, η_t). To fix ideas, let us 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_η 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_η are linear). In this setup, if both V_t and η_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 η_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

$$RV_{t,t+h} \perp \eta_t \big| V_t. \tag{34}$$

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

$$\mathbb{E}\left(e^{iu\log(RV_{t,t+h})+iz\log(\eta_t)}\big|\log(V_t)\right) = \mathbb{E}\left(e^{iu\log(RV_{t,t+h})}\big|\log(V_t)\right)\mathbb{E}\left(e^{iz\log(\eta_t)}\big|\log(V_t)\right), \tag{35}$$

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, following Bartlett (1938). More specifically, with the notation $\psi_{RV,V,\eta}(u,v,z) = \mathbb{E}(e^{iu\log(RV_{t,t+h})+iv\log(V_t)+iz\log(\eta_t)})$, for $u, v, z \in \mathbb{R}$, we have

$$\mathbb{E}\left(e^{iu\log(RV_{t,t+h})+iz\log(\eta_t)}\big|\log(V_t)\right) = \frac{\int_{\mathbb{R}} e^{-iv\log(V_t)}\psi_{RV,V,\eta}(u,v,z)dv}{\int_{\mathbb{R}} e^{-iv\log(V_t)}\psi_{RV,V,\eta}(0,v,0)dv}. \tag{36}$$

From here, (35) leads to

$$\begin{aligned}\int_{\mathbb{R}} e^{-iv\log(V_t)}\psi_{RV,V,\eta}(u,v,z)dv \int_{\mathbb{R}} e^{-iv\log(V_t)}\psi_{RV,V,\eta}(0,v,0)dv \\ = \int_{\mathbb{R}} e^{-iv\log(V_t)}\psi_{RV,V,\eta}(u,v,0)dv \int_{\mathbb{R}} e^{-iv\log(V_t)}\psi_{RV,V,\eta}(0,v,z)dv,\end{aligned} \tag{37}$$

and upon taking expectations, we have the following moment conditions involving only unconditional characteristic functions:

$$\begin{aligned}\int_{\mathbb{R}}\int_{\mathbb{R}} \psi_{RV,V,\eta}(u,v,z)\psi_{RV,V,\eta}(0,w,0)\psi_{RV,V,\eta}(0,-w-v,0)dvdw \\ = \int_{\mathbb{R}}\int_{\mathbb{R}} \psi_{RV,V,\eta}(u,v,0)\psi_{RV,V,\eta}(0,w,z)\psi_{RV,V,\eta}(0,-w-v,0)dvdw,\ u,z \in \mathbb{R}.\end{aligned} \tag{38}$$

Note that (38) implies (35) but the reverse does not need to hold. That is, (35) contains more information than (38) regarding the conditional independence hypothesis in (34). Since testing

(38) 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 η_t but instead we have estimates of them which satisfy approximately the following

$$\log(OV_t) = \log(V_t) + \epsilon_{t,1}, \ \ \log(JV_t) = \log\left(\eta_t \int_{\mathbb{R}} x^2 F(dx)\right) + \epsilon_{t,2}, \tag{39}$$

where $(\epsilon_{t,1}, \epsilon_{t,2})^\top \big| RV_{t+h}, V_t, \eta_t \sim N(\mathbf{0}_{2\times 1}, \mathbf{\Xi}_t)$, for some variance-covariance matrix $\mathbf{\Xi}_t$ that is given explicitly in the Appendix. Then, it is easy to see that the moment condition (38) is equivalent to

$$\begin{aligned} &\int_{\mathbb{R}}\int_{\mathbb{R}} \psi_{RV,OV,JV}(u, v, z)\psi_{RV,OV,JV}(0, w, 0)\psi_{RV,OV,JV}(0, -w-v, 0)dvdw \\ &\quad = \int_{\mathbb{R}}\int_{\mathbb{R}} \psi_{RV,OV,JV}(u, v, 0)\psi_{RV,OV,JV}(0, w, z)\psi_{RV,OV,JV}(0, -w-v, 0)dvdw, \ \ u, z \in \mathbb{R}, \end{aligned} \tag{40}$$

where we denote $\psi_{RV,OV,JV}(u, v, z) = \mathbb{E}(e^{iu\log(RV_{t,t+h}) + iv\log(OV_t) + iz\log(JV_t) + (1/2)(v\, z)\Xi_t(v\, z)^\top})$. From here, developing a test for (40) is relatively easy. Mainly, we can form sample moment conditions for pairs of (u, z), based on (40), 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 (40) are replaced with finite, but asymptotically expanding ones, based on the assumption of smoothness of the probability density of $(RV_{t,t+h}, V_t, \eta_t)$, see the Appendix for the details. If the sample moment condition vector is denoted with $\widehat{\mathbf{m}}$, which is of dimension $J \times 1$, our test will measure the distance of $\widehat{\mathbf{m}}$ from $\mathbf{0}_{J\times 1}$. The vector $\widehat{\mathbf{m}}$ is asymptotically normally distributed with mean zero under the null hypothesis. If we denote the length of the time-series with T and an estimate of the asymptotic variance with $\widehat{\mathbf{\Sigma}}_{\mathbf{m}}$, then our test statistics is given by

$$\widehat{W}_h = T\widehat{\mathbf{m}}^\top \widehat{\mathbf{\Sigma}}_{\mathbf{m}}^{-1}\widehat{\mathbf{m}}, \tag{41}$$

and should have an asymptotic $\chi^2(J)$ distribution if the hypothesis in (34) is true. The details on the construction of the test and its asymptotic behavior 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 $RV_{t,t+h} \perp V_t \big| \eta_t$. Since for the individual stocks, the computation of $\mathbf{\Xi}$ 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 4.0, 15.0 and 8.3, 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 19.7, 23.2 and 15.6, for horizons $h = 1$, $h = 5$ and $h = 22$, respectively. This is evidence against the hypothesis $RV_{t,t+h} \perp V_t\big|\eta_t$. Taken together, these two sets of test results suggest that η_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 component that is unrelated to diffusive volatility on the basis of a parametric option pricing model fitted to market index options.

5.2 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 η_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 4, 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. The results for the the individual stocks are in general similar to those for the S&P 500 index and are reported in the Appendix.

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.[23]

[23] 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.

Table 4: Forecasting Results using Multiple Option Predictors for the S&P 500 Index

		One-day Ahead		One-week Ahead		One-month Ahead	
Window	Criterion	HAR-MV	HAR-MV-JV	HAR-MV	HAR-MV-JV	HAR-MV	HAR-MV-JV
Rolling	MSE	0.8637	0.8235	0.8728	0.7979	0.9523	0.9335
	QLIKE	0.8152	0.7615	0.8019	0.7428	0.9146	0.9089
Increasing	MSE	0.8799	0.7760	0.8653	0.8141	0.9713	0.9535
	QLIKE	0.7729	0.7038	0.8600	0.8226	0.9796	0.9766

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.

5.3 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

$$VRP_t(1) = VIX_t^2 - RV_{t,m}, \tag{42}$$

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

$$VRP_t(2) = VIX_t^2 - ORV_{t,m}, \tag{43}$$

where we recall that our efficient estimate of the realized volatility, ORV, is defined in (31). 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

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.

consider a jump risk premium predictor defined as

$$JRP_t = JV_{t,m} - (RV_{t,m} - TV_{t,m}). \tag{44}$$

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.

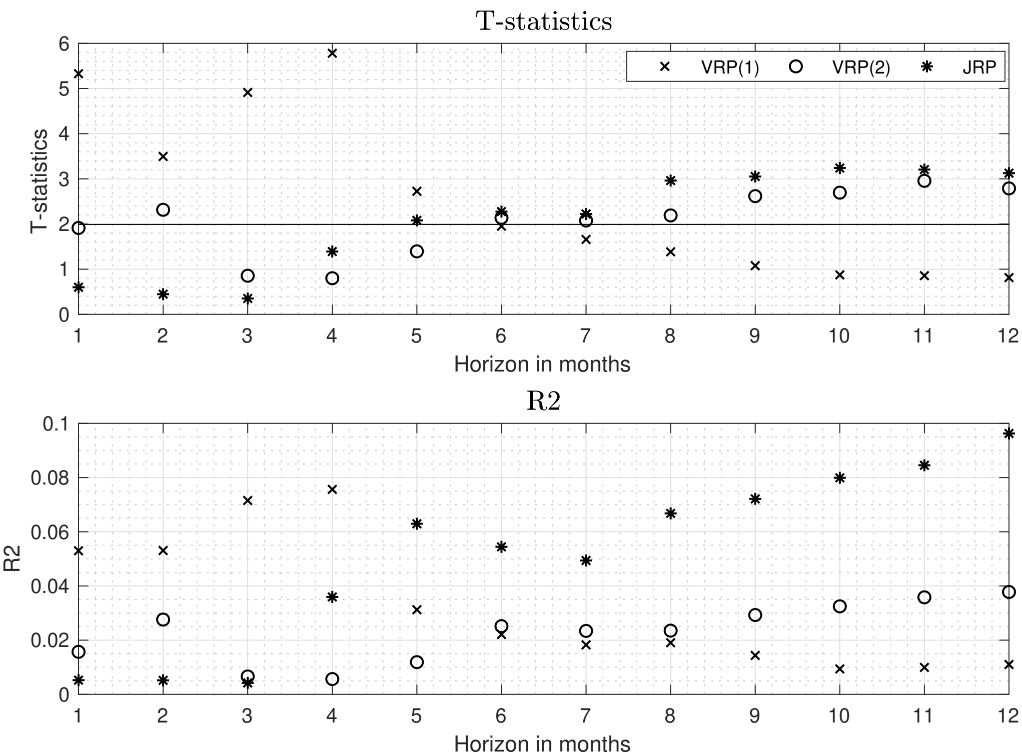

Figure 5: **Return Predictability Results.** Standard errors are computed using the Equal-Weighted Cosine estimator of the long-run variance with fixed-b critical values proposed in Lazarus et al. (2018) with degree of freedom equal to $0.4 \times T^{2/3}$. The critical value corresponding to the one-sided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

We run the following return predictability regressions

$$r_{t,t+h} = \alpha_{0,h} + \alpha_{1,h} X_t + \varepsilon_{t+h}, \tag{45}$$

for X_t being $VRP_t(1)$, $VRP_t(2)$ or JRP_t, and $r_{t,t+h}$ being the market excess return over the interval $[t, t+h]$ downloaded from Ken French's website. 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%.[24] 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.

To conclude, 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.

6 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. More specifically, we find that volatility forecasting models in which both past return- and option-based volatility proxies are used in generating the volatility forecast outperform significantly forecasting models based only on return or option data. The contribution of the option-based volatility proxy in the volatility forecast of the mixture models is on average well above 50% for the S&P 500 index and it is around 50% for the individual stocks. 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

[24]The t-statistic of the regression coefficient and the corresponding R^2 viewed as functions of the horizon of the predictive regressions reported here have very similar shape to those reported in Bollerslev et al. (2009) in spite of the different samples used in the two papers.

additional but small gains for volatility forecasting. Finally, the higher precision in measuring volatility offered by combining return and option data manifests itself in a more precise estimate of the market variance risk premium predictor used in prior work. We show that the latter exhibits different predictive behavior for future market returns than previously documented, highlighting further the importance of efficient volatility measurement proposed in the current paper.

References

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

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

Andersen, T. G., N. Fusari, and V. Todorov (2015). The risk premia embedded in index options. *Journal of Financial Economics 117*(3), 558–584.

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

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

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: Trades and quotes.

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

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

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

Bartlett, M. (1938). The characteristic function of a conditional statistic. *Journal of the London Mathematical Society 1*(1), 62–67.

Bates, D. S. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics 94*(1-2), 181–238.

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

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

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

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

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

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

Bollerslev, T., V. Todorov, and L. Xu (2015). Tail risk premia and return predictability. *Journal of Financial Economics 118*(1), 113–134.

Bollerslev, T. and H. Zhou (2002). Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics 109*(1), 33–65.

Bondarenko, O. and D. Muravyev (2020). Market return around the clock: A puzzle. *Available at SSRN*.

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

Broadie, M. and J. Detemple (1996). American option valuation: new bounds, approximations, and a comparison of existing methods. *The Review of Financial Studies 9*(4), 1211–1250.

Buccheri, G. and F. Corsi (2020). Hark the shark: Realized volatility modeling with measurement errors and nonlinear dependencies. *Journal of Financial Econometrics, forthcoming*.

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

Carr, P. and L. Wu (2008). Variance risk premiums. *The Review of Financial Studies 22*(3), 1311–1341.

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

Christensen, B. J. and N. R. Prabhala (1998). The relation between implied and realized volatility. *Journal of Financial Economics 50*(2), 125–150.

Christensen, K., M. Thyrsgaard, and B. Veliyev (2019). The realized empirical distribution function of stochastic variance with application to goodness-of-fit testing. *Journal of Econometrics 212*, 556–583.

Christoffersen, P., K. Jacobs, and B. Y. Chang (2013). Forecasting with option-implied information. In *Handbook of Economic Forecasting*, Volume 2, pp. 581–656. Elsevier.

Clark, T. E. and K. D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics 138*(1), 291–311.

Corradi, V. and W. Distaso (2006). Semi-parametric comparison of stochastic volatility models using realized measures. *The Review of Economic Studies 73*(3), 635–667.

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

Dobrev, D. and P. Szerszen (2010). The information content of high-frequency data for estimating equity return models and forecasting risk.

Feller, W. (1971). *An introduction to probability theory and its applications, vol 2.* John Wiley & Sons.

Fleming, J. (1998). The quality of market volatility forecasts implied by S&P 100 index option prices. *Journal of Empirical Finance 5*(4), 317–345.

Ghysels, E., P. Mykland, and E. Renault (2020). In-sample asymptotics and across-sample efficiency gains for high frequency data statistics.

Ghysels, E., A. Sinko, and R. Valkanov (2007). Midas regressions: Further results and new directions. *Econometric reviews 26*(1), 53–90.

Hansen, P. R., Z. Huang, and H. H. Shek (2012). Realized GARCH: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics 27*(6), 877–906.

Huber, P. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics 35*, 73–101.

Jacod, J. and P. Protter (2011). *Discretization of processes*, Volume 67. Springer Science & Business Media.

Jacod, J. and A. Shiryaev (2013). *Limit theorems for stochastic processes*, Volume 288. Springer Science & Business Media.

Jacod, J. and V. Todorov (2014). Efficient estimation of integrated volatility in presence of infinite variation jumps. *The Annals of Statistics 42*(3), 1029–1069.

Jacod, J. and V. Todorov (2018). Limit theorems for integrated local empirical characteristic exponents from noisy high-frequency data with application to volatility and jump activity estimation. *The Annals of Applied Probability 28*(1), 511–576.

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

Kambouroudis, D. S., D. G. McMillan, and K. Tsakou (2016). Forecasting stock return volatility: a comparison of garch, implied volatility, and realized volatility models. *Journal of Futures Markets 36*(12), 1127–1163.

Koopman, S. J., B. Jungbacker, and E. Hol (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance 12*(3), 445–475.

Koopman, S. J. and M. Schart (2013). The analysis of stochastic volatility in the presence of daily realized measures. *Journal of Financial Econometrics 11*, 76–115.

Lazarus, E., D. J. Lewis, J. H. Stock, and M. W. Watson (2018). Har inference: Recommendations for practice. *Journal of Business & Economic Statistics 36*(4), 541–559.

Li, J. and G. Zinna (2018). The variance risk premium: Components, term structures, and stock return predictability. *Journal of Business & Economic Statistics 36*(3), 411–425.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari LXIV*, 19–47.

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

Martin, G., A. Reidy, and J. Wright (2009). Does the option market produce superior forecasts of noise-corrected volatility measures? *Journal of Applied Econometrics 24*(1), 77–104.

Neely, C. J. (2009). Forecasting foreign exchange volatility: Why is implied volatility biased and inefficient? And does it matter? *Journal of International Financial Markets, Institutions and Money 19*(1), 188–205.

Oikonomou, I., A. Stancu, L. Symeonidis, and C. W. Simen (2019). The information content of short-term options. *Journal of Financial Markets*.

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

Qin, L. and V. Todorov (2019). Nonparametric implied lévy densities. *The Annals of Statistics 47*(2), 1025–1060.

Shephard, N. and K. Sheppard (2010). Realising the future: Forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics 25*, 197–231.

Takahashi, M., Y. Omori, and T. Watanabe (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics & Data Analysis 53*(6), 2404–2426.

Todorov, V. (2009). Estimation of continuous-time stochastic volatility models with jumps using high-frequency data. *Journal of Econometrics 148*(2), 131–148.

Todorov, V. (2019). Nonparametric spot volatility from options. *Annals of Applied Probability 29*, 3590–3636.

Todorov, V. (2020). Testing and inference for fixed times of discontinuity in semimartingales. *Bernoulli 26*(4), 2907–2948.

Todorov, V. and G. Tauchen (2012). Inverse realized laplace transforms for nonparametric volatility density estimation in jump-diffusions. *Journal of the American Statistical Association 107*(498), 622–635.

Wang, X. and Y. Hong (2018). Characteristic function based testing for conditional independence: a nonparametric regression approach. *Econometric Theory 34*(4), 815–849.

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 28 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.

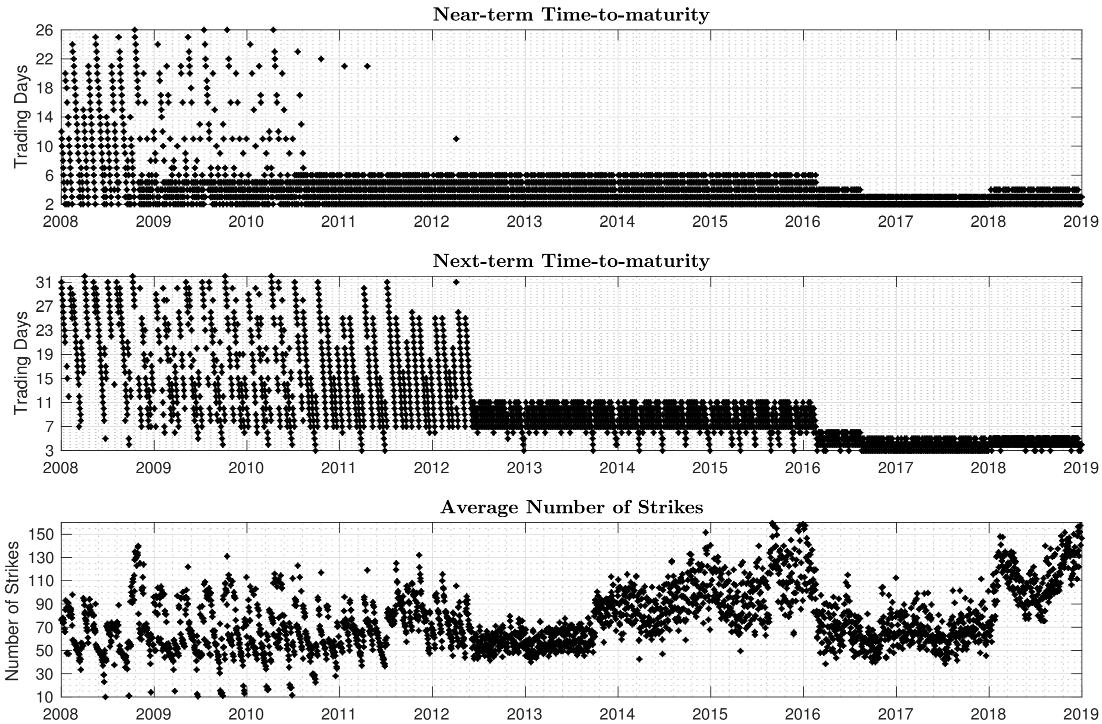

Figure 6: **Option Data Summary for SPX**.

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,

$$\mathrm{ON} = \frac{\sum_{t=1}^{T}(\sum_{i=1}^{n^{cc}} \Delta_{t,i}^{n} x)^2}{\sum_{t=1}^{T}(\sum_{i=1}^{n^{oc}} \Delta_{t,i}^{n} x)^2}, \tag{46}$$

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

$$\tau = \tau_{\mathrm{AM}} - 1 + \frac{\mathrm{ON} - 1}{\mathrm{ON}}, \tag{47}$$

where τ_{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:

$$F_{t,\tau} = K_{near,\tau} + e^{r_f \tau} \left[\widehat{O}_{t,\tau}(K_{near,\tau}, C) - \widehat{O}_{t,\tau}(K_{near,\tau}, P) \right], \tag{48}$$

where r_f is the risk-free rate and $K_{near,\tau}$ 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:

$$\mathbf{K} = [K_{low} : \Delta K : K_{high}] \tag{49}$$

   where K_{low} and K_{high} stands for the lowest and highest observed strike price in the data. We set the strike grid ΔK 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, τ). 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:

$$\widehat{\lambda}^{+} = \frac{\log(\widehat{O}(K_{Q_{75}}, C)/\widehat{O}(K_{Q_{50}}, C))}{\log(K^C_{Q_{50}}/K^C_{Q_{75}})} \bigvee 5, \tag{50}$$

$$\widehat{\lambda}^{-} = \frac{\log(\widehat{O}(K_{Q_{50}}, P)/\widehat{O}(K_{Q_{25}}, P))}{\log(K^P_{Q_{50}}/K^P_{Q_{25}})} \bigvee 5, \tag{51}$$

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 ΔK 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 below the minimum tick size:[25]

$$\widehat{O}_{t,\tau}(K, C) = e^{-\widehat{\lambda}^{+}|\log((K_{high}+\Delta K)/K_{high})|}\widehat{O}_{t,\tau}(K_{high}, C), \ K > K_{high}, \tag{52}$$

$$\widehat{O}_{t,\tau}(K, P) = e^{-\widehat{\lambda}^{-}|\log((K_{low}-\Delta K)/K_{low})|}\widehat{O}_{t,\tau}(K_{low}, P), \ K < K_{low}. \tag{53}$$

where K_{low} and K_{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

$$\widehat{V}_{t,\tau}(\widehat{u}_t) = -\frac{2}{(\tau_2 - \tau_1)\widehat{u}_t^2}\left(\log|\widehat{\mathcal{L}}_{t,\tau_2}(\widehat{u}_t)| - \log|\widehat{\mathcal{L}}_{t,\tau_1}(\widehat{u}_t)|\right), \tag{54}$$

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

$$\widehat{u}_t^{(1)} = \inf\left\{u \geq 0 : |\widehat{\mathcal{L}}_{t,\tau_2}(u)| \leq 0.95\right\}, \ \overline{u}_t = \sqrt{-\frac{2\log(0.95)}{\tau_2\widehat{\sigma}^2_{t,ATM}}}, \ \ \widehat{u}_t^{(2)} = \text{argmin}_{u\in[0,\overline{u}]}|\widehat{\mathcal{L}}_{t,\tau_2}(u)|. \tag{55}$$

[25]The minimum tick size is set according to options specifications on Cboe website: `http://www.cboe.com/products/contract-specs-and-trading-hours/contract-specs`.

We note that the tuning parameter $\widehat{u}_t$ 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 $\widehat{V}_{t,\tau}(\widehat{u}_t)$ 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

For the construction of EV in Section 2.3, we need estimates for the asymptotic variances of each of the two volatility proxies. We start with the option-based one. Each volatility estimator $\widehat{V}_{t,\tau}(\widehat{u}_{t,\tau})$ is computed from options having time-to-maturity τ and whose log-strikes are given by

$$k_1 < k_2 < ... < k_{N_\tau},$$

where in order to simplify notation, we suppressed dependence on t and τ 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

$$\begin{aligned}\widehat{Avar}(\widehat{V}_{t,\tau}) = \frac{4}{\tau^2 \widehat{u}_{t,\tau}^4 |\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})|^4} \Bigg(& \Re\left(\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})\right)^2 A_{t,\tau}^{(1)} + \Im\left(\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})\right)^2 A_{t,\tau}^{(2)} \\ & + 2\Re\left(\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})\right) \Im\left(\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})\right) A_{t,\tau}^{(3)} \Bigg),\end{aligned} \tag{56}$$

where with the notation

$$f_t(u,k) = (u^2 + iu)e^{(iu-1)k - iux_t}, \tag{57}$$

we set

$$A_{t,\tau}^{(1)} = \sum_{j=2}^{N_\tau} \Re(f_t(\widehat{u}_\tau, k_{j-1}))^2 \widehat{\sigma}_{t,\tau}(k_{j-1})^2 (k_j - k_{j-1})^2, \; A_{t,\tau}^{(2)} = \sum_{j=2}^{N_\tau} \Im(f_t(\widehat{u}_\tau, k_{j-1}))^2 \widehat{\sigma}_{t,\tau}(k_{j-1})^2 (k_j - k_{j-1})^2,$$

$$A_{t,\tau}^{(3)} = \sum_{j=2}^{N_\tau} \Re(f_t(\widehat{u}_\tau, k_{j-1})) \Im(f_t(\widehat{u}_\tau, k_{j-1})) \widehat{\sigma}_{t,\tau}(k_{j-1})^2 (k_j - k_{j-1})^2,$$

and we denote

$$\widehat{\sigma}_{t,\tau}(k_j) = \sqrt{\frac{2}{3}}\left(\widehat{O}_{t,\tau}(k_j) - \frac{1}{2}\widehat{O}_{t,\tau}(k_{j-1}) - \frac{1}{2}\widehat{O}_{t,\tau}(k_{j+1})\right),\ j = 2, ..., N_\tau - 1, \tag{58}$$

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

$$\begin{aligned}
\widehat{\sigma}_{t,\tau}(k_{j^*}) &= \widehat{O}_{t,\tau}(k_{j^*}) - \widehat{O}_{t,\tau}(k_{j^*-1}) - (\widehat{O}_{t,\tau}(k_{j^*-1}) - \widehat{O}_{t,\tau}(k_{j^*-2}))\frac{K_{j^*} - K_{j^*-1}}{K_{j^*-1} - K_{j^*-2}}, \quad \text{if } k_{j^*} \le x_t,\\
\widehat{\sigma}_{t,\tau}(k_{j^*}) &= \widehat{O}_{t,\tau}(k_{j^*}) - \widehat{O}_{t,\tau}(k_{j^*+1}) - (\widehat{O}_{t,\tau}(k_{j^*+1}) - \widehat{O}_{t,\tau}(k_{j^*+2}))\frac{K_{j^*} - K_{j^*+1}}{K_{j^*+1} - K_{j^*+2}}, \quad \text{if } k_{j^*} > x_t.
\end{aligned} \tag{59}$$

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

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

D. Testing for Conditional Independence

D.1. Formulation of the Test and the Asymptotics for it

We start with forming the empirical characteristic function:

$$\widehat{\psi}_{RV,OV,JV}(u, v, z) = \frac{1}{T-h}\sum_{t=1}^{T-h} e^{iu\log(RV_{t,t+h}) + iv\log(OV_t) + iz\log(JV_t) + (1/2)(v\, z)\Xi_t(v\, z)^\top},\ u, v, z \in \mathbb{R}, \tag{60}$$

and using it, we define the following moment estimator

$$\begin{aligned}
\widehat{m}(u, z) = \int_{-v_T}^{v_T}\int_{-v_T}^{v_T} &(\widehat{\psi}_{RV,OV,JV}(u, v, z)\widehat{\psi}_{RV,OV,JV}(0, w, 0)\\
&- \widehat{\psi}_{RV,OV,JV}(u, v, 0)\widehat{\psi}_{RV,OV,JV}(0, w, z))\widehat{\psi}_{RV,OV,JV}(0, -w - v, 0)dvdw,
\end{aligned} \tag{61}$$

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 denote a vector formed of moment conditions $\widehat{m}(u, z)$, for several different values of the pair (u, z), with $\widehat{\mathbf{m}}$.

We next define $\widehat{\mathbf{\Sigma}}_{\mathbf{m}}$. For this, we first introduce the following notation

$$\zeta_t(u,v,z) = e^{iu\log(RV_{t,t+h})+iv\log(OV_t)+iz\log(JV_t)+(1/2)(v\,z)\Xi_t(v\,z)^\top} - \widehat{\psi}_{RV,OV,JV}(u,v,z), \tag{62}$$

and denote

$$\widehat{m}_t(u,z) = \int_{-v_T}^{v_T}\int_{-v_T}^{v_T} \widehat{m}_t(u,z;v,w)dvdw, \tag{63}$$

where

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

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

$$\widehat{\mathbf{\Sigma}}_{\mathbf{m}} = \widehat{\gamma}_m(0) + \sum_{l=1}^{L_T}\left(1 - \frac{k}{L_T+1}\right)(\widehat{\gamma}_m(l) + \widehat{\gamma}_m(l)^\top), \tag{64}$$

for a sequence L_T increasing to infinity as $T \to \infty$.

The behavior of the resulting test, given in (41), is not entirely standard because of the increasing range of integration over the products of the joint characteristic functions. Therefore, we need to ensure that the bias introduced in the moment condition introduced from that is asymptotically negligible. This will follow by choosing v_T appropriately and making use of an assumption for the smoothness for the joint probability density of (RV_{t+h}, V_t, η_t). More specifically, an assumption for the existence and integrability of the first three partial derivatives of this probability density with respect to the second element (V_t) ensures that (see e.g., Lemma XV.4.4 in Feller (1971)):

$$\psi_{RV,OV,JV}(u,v,z) = O(|v|^{-3}), \quad \text{for } |v| \to \infty \text{ and } u,z \in \mathbb{R} \text{ fixed.} \tag{65}$$

In addition, a standard ergodicity and mixing condition for the time series (RV_{t+h}, V_t, η_t) yields (see e.g., Theorem VIII.3.102 in Jacod and Shiryaev (2013)):

$$\mathbb{E}||\widehat{\psi}_{RV,OV,JV}(u,v,z) - \psi_{RV,OV,JV}(u,v,z)||^2 \le C/T, \quad \text{for } |v| \le \sqrt{T}, \tag{66}$$

and some constant C that does not depend on T as well as u, v, z. These two results can be used to reduce the asymptotics for the test to the derivation of a standard CLT for stationary sequences.

In particular, lets define

$$
\begin{aligned}
\overline{m}(u,z) = \int_{\mathbb{R}}\int_{\mathbb{R}} [(&\widehat{\psi}_{RV,OV,JV}(u,v,z) - \psi_{RV,OV,JV}(u,v,z))\psi_{OV}(w)\psi_{OV}(-w-v) \\
&+ \psi_{RV,OV,JV}(u,v,z)(\widehat{\psi}_{OV}(w) - \psi_{OV}(w))\psi_{OV}(-w-v) \\
&+ \psi_{RV,OV,JV}(u,v,z)\psi_{OV}(w)(\widehat{\psi}_{OV}(-w-v) - \psi_{OV}(-w-v)) \\
&- (\widehat{\psi}_{RV,OV}(u,v) - \psi_{RV,OV}(u,v))\psi_{OV,JV}(w,z)\psi_{OV}(-w-v) \\
&- \psi_{RV,OV}(u,v)(\widehat{\psi}_{OV,JV}(w,z) - \psi_{OV,JV}(w,z))\psi_{OV}(-w-v) \\
&- \psi_{RV,OV}(u,v)\psi_{OV,JV}(w,z)(\widehat{\psi}_{OV}(-w-v) - \psi_{OV}(-w-v))]dvdw,
\end{aligned} \tag{67}
$$

where we use the shorthand notation $\psi_{OV}(v) = \psi_{RV,OV,JV}(0,v,0)$, $\psi_{RV,OV}(u,v) = \psi_{RV,OV,JV}(u,v,0)$ and $\psi_{OV,JV}(v,z) = \psi_{RV,OV,JV}(0,v,z)$. Then, using (65) and (66), it is easy to show under the null hypothesis in (40):

$$||\widehat{m}(u,z) - \overline{m}(u,z)|| = O_p\left(\frac{v_T}{T}\bigvee v_T^{-2}\right). \tag{68}$$

From here, by standard CLT for stationary and mixing sequences, one can establish under the null hypothesis (40), and provided $v_T/\sqrt{T} \to 0$ and $v_T/T^{1/4} \to \infty$:

$$\sqrt{T}\widehat{\mathbf{m}} \quad \xrightarrow{\mathcal{L}} \quad N(\mathbf{0}, \mathbf{\Sigma_m}), \tag{69}$$

where $\mathbf{\Sigma_m}$ is the long-run variance-covariance matrix corresponding to the moment vector with moment conditions given by $\overline{m}(u,z)$.

We proceed next with the consistency of $\widehat{\mathbf{\Sigma}}_{\mathbf{m}}$. We first define $\overline{m}_t(u,z;v,w)$ from $\widehat{m}_t(u,z;v,w)$ by replacing $\widehat{\psi}_{RV,OV,JV}$ with $\psi_{RV,OV,JV}$, and from here

$$\overline{m}_t(u,z) = \int_{\mathbb{R}}\int_{\mathbb{R}} \overline{m}_t(u,z;v,w)dvdw. \tag{70}$$

With this notation, we introduce the long-run covariance matrix based on $\overline{m}_t(u,z)$ as

$$\overline{\mathbf{\Sigma}}_{\mathbf{m}} = \overline{\gamma}_m(0) + \sum_{l=1}^{L_T}\left(1 - \frac{k}{L_T+1}\right)(\overline{\gamma}_m(l) + \overline{\gamma}_m(l)^{\top}), \tag{71}$$

where $\overline{\gamma}_m(l)$ is the sample autocovariance corresponding to $\overline{m}_t(u,z)$. Using (65) and (66), it is easy to show under the null hypothesis in (40):

$$||\widehat{\mathbf{\Sigma}}_{\mathbf{m}} - \overline{\mathbf{\Sigma}}_{\mathbf{m}}|| = O_p\left(\frac{v_T L_T}{\sqrt{T}} \bigvee v_T^{-2} L_T\right), \tag{72}$$

which is asymptotically negligible when $v_T/\sqrt{T} \to 0$ and $v_T/T^{1/4} \to \infty$, provided L_T grows at a slow rate (depending on the choice of v_T). The consistency of $\overline{\mathbf{\Sigma}}_{\mathbf{m}}$ for $\mathbf{\Sigma}_{\mathbf{m}}$ follows by standard arguments for stationary and mixing sequences.

D.2. Tuning Parameters

First, we set v_T in the following way

$$v_T = \frac{2}{3} \times \frac{T^{0.26}}{\mathrm{std}(OV_t)}. \tag{73}$$

Second, in implementing the 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 (recall that $\widehat{m}(u, z)$ 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.

Third, for the estimation of the long-run variance, we set $L_T = 132$ (six months), which corresponds approximately to three months.

Finally, the matrix of asymptotic variances $\mathbf{\Xi}_t$ is given by

$$\mathbf{\Xi}_t = \frac{1}{4}\begin{pmatrix} \frac{1}{OV_t} & 0 \\ 0 & \frac{1}{JV_t} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}\left(\overline{\Xi}_{t,\tau_1} + \overline{\Xi}_{t,\tau_2}\right)\begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}^{\top}\begin{pmatrix} \frac{1}{OV_t} & 0 \\ 0 & \frac{1}{JV_t} \end{pmatrix}, \tag{74}$$

where $\overline{\mathbf{\Xi}}_{t,\tau}$ is a 2×2 matrix with the following entries

$$\overline{\mathbf{\Xi}}_{t,\tau}^{(1,1)} = \frac{1}{\tau^2}\sum_{j=2}^{N_\tau} g_t(k_{j-1})^2 \widehat{\sigma}_{t,\tau}^2(k_{j-1})(k_j - k_{j-1})^2, \quad \overline{\mathbf{\Xi}}_{t,\tau}^{(2,2)} = \widehat{\mathrm{Avar}}(\widehat{V}_{t,\tau}), \tag{75}$$

and

$$\overline{\mathbf{\Xi}}_{t,\tau}^{(1,2)} = -\frac{2}{\tau^2 \widehat{u}_{t,\tau}^2 |\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})|^2}\left(\Re\left(\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})\right)\overline{A}_{t,\tau}^{(1)} + \Im\left(\widehat{\mathcal{L}}_{t,\tau}(\widehat{u}_{t,\tau})\right)\overline{A}_{t,\tau}^{(2)}\right), \tag{76}$$

with the notation in (56) and (58)-(59) as well as

$$\begin{cases} \overline{A}_{t,\tau}^{(1)} = \sum_{j=2}^{N_\tau} \Re(f_t(\widehat{u}_\tau, k_{j-1}))g_t(k_{j-1})\widehat{\sigma}_{t,\tau}(k_{j-1})^2(k_j - k_{j-1})^2, \\ \overline{A}_{t,\tau}^{(2)} = \sum_{j=2}^{N_\tau} \Im(f_t(\widehat{u}_\tau, k_{j-1}))g_t(k_{j-1})\widehat{\sigma}_{t,\tau}(k_{j-1})^2(k_j - k_{j-1})^2, \end{cases} \tag{77}$$

and

$$g_t(k) = 2e^{-k}(1 - k + x_t). \tag{78}$$

Since on a few days, the estimation of Ξ_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 $v_T^2 \Xi_t^{(11)} > 2$.

E. Volatility Forecasting for Individual Stocks

In this section we evaluate 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. 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 5. Starting with the one-day ahead volatility forecasts, we can see from Panel A of Table 5 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, e.g., due to the early exercise premium in the American style individual stock options. 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

Table 5: Forecasting Results for DJIA Stocks

Window	Criterion	HAR-RV	HAR-TV	HARQ	HAR-OV	HAR-MV
Panel A: One-day Ahead Relative Forecasting Loss						
Rolling Window	MSE (Median)	1.0336	1.0000	1.0234	1.1385	0.9253
	MSE (Mean)	1.0248	1.0000	1.0219	1.0293	0.9658
	QLIKE (Median)	1.0320	1.0000	1.0105	1.1655	0.9569
	QLIKE (Mean)	1.0224	1.0000	1.0413	1.1055	0.9414
Increasing Window	MSE (Median)	1.0279	1.0000	1.0084	1.0049	0.8617
	MSE (Mean)	1.0246	1.0000	1.0083	1.0556	0.9697
	QLIKE (Median)	1.0255	1.0000	1.0123	1.1324	0.9199
	QLIKE (Mean)	1.0130	1.0000	1.0361	1.1508	0.9481
Panel B: One-week Ahead Relative Forecasting Loss						
Window	Criterion	HAR-RV	HAR-TV	HARQ	HAR-OV	HAR-MV
Rolling Window	MSE (Median)	1.0178	1.0000	0.9941	0.8795	0.7921
	MSE (Mean)	1.0265	1.0000	1.0114	0.9586	0.9140
	QLIKE (Median)	1.0215	1.0000	0.9985	0.9288	0.8412
	QLIKE (Mean)	1.0192	1.0000	1.0242	0.9731	0.8797
Increasing Window	MSE (Median)	1.0200	1.0000	0.9766	0.7423	0.7122
	MSE (Mean)	1.0204	1.0000	0.9985	0.9957	0.9270
	QLIKE (Median)	1.0251	1.0000	0.9989	0.8611	0.8077
	QLIKE (Mean)	1.0106	1.0000	1.0078	1.0243	0.8963
Panel C: One-month Ahead Relative Forecasting Loss						
Window	Criterion	HAR-RV	HAR-TV	HARQ	HAR-OV	HAR-MV
Rolling Window	MSE (Median)	0.9990	1.0000	0.9780	0.9429	0.8993
	MSE (Mean)	1.0206	1.0000	0.9947	0.9925	0.9485
	QLIKE (Median)	1.0170	1.0000	0.9906	1.0054	0.9330
	QLIKE (Mean)	1.0102	1.0000	0.9997	0.9856	0.9572
Increasing Window	MSE (Median)	0.9962	1.0000	0.9798	0.8349	0.8266
	MSE (Mean)	1.0112	1.0000	0.9869	1.0170	0.9550
	QLIKE (Median)	1.0054	1.0000	0.9847	0.9298	0.8859
	QLIKE (Mean)	0.9993	1.0000	0.9929	1.0244	0.9568

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.

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 5 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 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 5 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.

Overall, as in the case of S&P 500 index, the best model for forecasting individual stock volatility is HAR-MV. For this model, we computed the ratio $\frac{||\beta_h||}{||\beta_h||+||\gamma_h||}$ in order to assess the relative weight assigned to the option-based volatility proxy in the volatility forecast. The cross-sectional mean of the time series mean of this ratio for the daily horizon and when using rolling estimation window is 0.48. This is somewhat lower than the corresponding number for the S&P 500 index and reflects the lower quality of the individual option data relative to market index option data.

Finally, in Table 6, we compare the forecasting performance of the HAR-MV model with that of the HAR-MV-JV model introduced in Section 5.2. Qualitatively, the results are similar to those reported for the S&P 500 index in Table 4. Mainly, the risk-neutral jump variation measure JV offers only minor forecasting gains (if any). The cross-sectional median of the adjusted Diebold-Mariano test for equal forecasting performance has p-value ranging from 0.27% to 16.54%, with the highest gains (statistically) from adding JV being for daily and weekly horizons.

Table 6: Forecasting Results using Multiple Option Predictors for DJIA Stocks

		One-day Ahead		One-week Ahead		One-month Ahead	
Window	Criterion	HAR-MV	HAR-MV-JV	HAR-MV	HAR-MV-JV	HAR-MV	HAR-MV-JV
Rolling	MSE	0.9658	0.9568	0.9140	0.8847	0.9485	0.9401
	QLIKE	0.9414	0.9155	0.8797	0.8279	0.9572	0.9426
Increasing	MSE	0.9697	0.9600	0.9270	0.9038	0.9550	0.9481
	QLIKE	0.9481	0.9220	0.8600	0.8963	0.9568	0.9447

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