AARHUS
UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

The Fine Structure of Equity-Index Option Dynamics

Torben G. Andersen, Oleg Bondarenko, Viktor Todorov and George Tauchen

CREATES Research Paper 2013-52

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

The Fine Structure of Equity-Index Option Dynamics*

Torben G. Andersen† Oleg Bondarenko‡ Viktor Todorov§ George Tauchen¶

November 1, 2013

Abstract

We analyze the high-frequency dynamics of S&P 500 equity-index option prices by constructing an assortment of implied volatility measures. This allows us to infer the underlying fine structure behind the innovations in the latent state variables driving the movements of the volatility surface. In particular, we focus attention on implied volatilities covering a wide range of moneyness (strike/underlying stock price), which load differentially on the different latent state variables. We conduct a similar analysis for high-frequency observations on the VIX volatility index as well as on futures written on it. We find that the innovations in the risk-neutral intensity of the negative jumps in the S&P 500 index over small time scales are best described via non-Gaussian shocks, i.e., jumps. On the other hand, the innovations over small time scales of the diffusive volatility are best modeled as Gaussian with occasional jumps.

Keywords: high-frequency data, implied volatility, jump activity, Kolmogorov-Smirnov test, stable process, stochastic volatility, VIX index.

JEL classification: C51, C52, G12.

*Andersen gratefully acknowledges support from CREATES, Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. Todorov's work was partially supported by NSF grant SES-0957330. We are also grateful for support from a grant by the CME Group. We thank Makoto Takahashi for providing assistance with collecting and cleaning the VIX futures data.

†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; NBER, Cambridge, MA; and CREATES, Aarhus, Denmark; e-mail: t-andersen@northwestern.edu.

‡Department of Finance, University of Illinois at Chicago, Chicago, IL 60607; e-mail: olegb@uic.edu

§Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@northwestern.edu.

¶Department of Economics, Duke University, Durham, NC 27708; e-mail: george.tauchen@duke.edu.

1 Introduction

Financial asset prices exhibit time-varying volatility. Volatility risk has a nontrivial impact on the behavior of investors and they require compensation for bearing it. Over the last two decades trading activity in derivatives, that allow for speculation or hedging vis-a-vis the volatility risk, has grown dramatically. These instruments include the plain vanilla options on the underlying asset, such as the broad equity market index, but also more directly the so-called variance swaps, which are forward contracts on future realized volatility (which in turn is a nonparametric estimate for the unobserved quadratic variation). The price of the variance swap can be recovered in model-free fashion through the price of a portfolio of out-of-the-money options on the underlying asset. The Chicago Board Options Exchange (CBOE) uses this nonparametric approach to compute the well-known VIX volatility index based on S&P 500 index options. In recent years, the VIX index itself has become the underlying instrument for futures and options written on it, further expanding the opportunities for managing broad exposures to equity market volatility risk.

The abundance of reliable data on volatility-related derivative contracts enables us to take a closer look at the properties of the stochastic volatility, which is otherwise latent, or "hidden," within the stock returns. Todorov and Tauchen (2011a) show that, although the VIX index is a risk-neutral expectation of future realized volatility, under conventional and realistic model settings, the VIX index preserves important information about the behavior of the latent stochastic volatility over small time scales. In particular, the presence of VIX jumps can be traced back to discontinuities in the stochastic volatility process itself, although the magnitudes of the jumps in the two series generally differ. Similarly, the "concentration" of small jumps, known as the jump activity (Blumenthal-Getoor index), in (spot) volatility and the VIX index coincides. Finally, the presence of a diffusive component is preserved when going from stochastic volatility to the VIX. Using high-frequency VIX data, Todorov and Tauchen (2011a) find that the VIX index – and hence the stochastic volatility – contains jumps and is best described as a pure-jump process with infinite variation jumps.

The goal of this paper is to enrich the empirical evidence on the properties of the latent stochastic volatility over small time scales using high-frequency data across a much wider set of derivative contracts than used hitherto in the literature. Under the standard assumption that the (risk-neutral) jump intensity is a sole function of the stochastic volatility, Black-Scholes implied volatilities obtained from S&P 500 index options are also functions of the stochastic volatility and its components

alone (along with the characteristics of the option contracts). Hence, they likewise typically "inherit" the behavior of volatility (and its components) over small scales. Furthermore, the identical logic applies to derivatives written on the VIX index such as VIX futures. Therefore, data for an extended set of volatility derivatives will clearly aid efficiency and robustness in evaluating the findings of Todorov and Tauchen (2011a), based strictly on high-frequency VIX data.

More importantly, however, the additional derivatives data allow us to obtain qualitatively new insights. For example, many studies conclude that volatility has multiple factors. Further, recent evidence suggests that the dynamics of the risk-neutral negative jump intensity for the equity index cannot be captured fully by (the components of) stochastic volatility, see, e.g., Bollerslev and Todorov (2011) and Andersen et al. (2013a). In this case, VIX will be governed by the factors driving the jump intensity or the stochastic volatility, or both. The use of derivative data loading differentially on the above factors (state variables) helps us discern the properties of the different types of factors and thus fosters a deeper understanding of the fine structure of the VIX dynamics. For example, the Black-Scholes implied volatilities of short-maturity out-of-the-money puts load primarily on the risk-neutral intensity of negative jumps, so their high-frequency behavior speaks to the small scale behavior of the factors driving the jump intensity. Similarly, the implied volatility of short-maturity at-the-money options is mostly determined by the latent spot volatility and, hence, provides more direct evidence on the structure of the volatility process.

On a practical level, Andersen et al. (2013) document problems associated with the construction of the VIX at high frequencies. These are mostly related to the rules for truncating the deep out-of-the-money options in the computation of the index.[1] Using the S&P 500 index options data directly allows us to alleviate the impact of such errors on the high-frequency VIX increments.

In summary, our empirical analysis exploits the following high-frequency data series. First, we use short-maturity S&P 500 futures and futures options traded on the Chicago Mercantile Exchange (CME). Using the option prices, we construct one-month Black-Scholes implied volatility series with a fixed log moneyness (strike/futures price), measured relative to the at-the-money implied volatility. Second, we use high-frequency data on the S&P 500 index futures and the VIX index. Finally, we use data on the two nearby futures written on the VIX index. Combined, our data span the period from January 2007 till May 2012, but we filter out problematic observations and there is a bit of mismatch in the samples of our high-frequency data across the different types of derivatives contracts.

[1] Theoretically, the replication strategy behind the VIX index involves using options with log-moneyness covering the whole real line. In practice, we have options with finite moneyness only and this inevitably induces approximation errors.

We apply two very different procedures in our investigation of the fine structure of the asset price dynamics at small time scales. The first technique involves computing a ratio of power variations at two different frequencies, which allows us to estimate the activity index of the given process. Estimators of this type have been studied by Ait-Sahalia and Jacod (2010) and Todorov and Tauchen (2010, 2011a). The activity index for a process is two if it has a diffusion component. On the other hand, if no diffusion term is present, the activity index equals the jump activity index for the (alternative) pure-jump model, and it is then less than two. Our second econometric tool is the empirical cumulative distribution function (cdf) of the nonparametrically de-volatilized high-frequency increments of the process and was proposed by Todorov and Tauchen (2013). The de-volatilized high-frequency increments should be approximately Gaussian if the process is jump-diffusive, and they should follow a stable distribution if the process is of the pure-jump type.

Turning to our results, we document an intriguing pattern. The lowest point estimates for the activity, of around 1.6, are obtained for the implied volatility of deep out-of-the-money put options. As we move in the strike dimension from out-of-the-money puts to at-the-money options, the estimates for the activity index gradually increase to the maximal value of two, indicative of a jump-diffusive process. Our additional test based on the empirical cdf of the de-volatilized increments corroborates this finding, i.e., it also suggests the implied volatilities for deep out-of-the-money puts are pure-jump processes, while pointing towards the opposite conclusion for the implied volatility of near-the-money options. These results are consistent with the risk-neutral intensity of the negative price jumps being driven by state variable(s) of the pure-jump type and the stochastic volatility being governed by a jump-diffusion. The value of volatility related indices or contracts, such as the VIX index and the VIX futures, are functions of both the state variables driving the volatility as well as those determining the jump intensity. Hence, in finite samples, point estimates for the activity index of such series will typically fall between the lowest and highest activity indices of the state variables governing their dynamics. This is exactly what we find as the estimates for the activity indices of the VIX-based contracts consistently take values within the range of the activity estimates obtained across the moneyness dimension of the S&P 500 implied volatility surface.

The rest of the paper is organized as follows. Section 2 introduces notation and formally defines the option-based quantities that we study in the paper. In Section 3, we describe the separate data set used in our analysis and we conduct an initial analysis regarding the liquidity of the individual instruments. Section 4 reviews the econometric tools we use, while Section 5 contains the main empirical results, and Section 6 concludes.

2 Setup

We assume the underlying asset price process, X_t, is an Itô semimartingale under the statistical measure $\mathbb{P}$, characterized through the following return dynamics,

$$\frac{dX_t}{X_t} \;=\; \alpha_t\,dt \;+\; \sqrt{V_t}\,dW_t \;+\; \int_{\mathbb{R}} x\,\widetilde{\mu}^{\mathbb{P}}(ds,dx), \qquad \text{under } \mathbb{P}, \tag{1}$$

where W_t is a Brownian motion and μ is an integer-valued measure with compensator $\nu^{\mathbb{P}}$ and $\widetilde{\mu}^{\mathbb{P}} = \mu - \nu^{\mathbb{P}}$ is the associated martingale measure.

Under mild regularity conditions, no arbitrage implies the existence of a locally equivalent probability measure $\mathbb{Q}$, labeled risk-neutral, under which X_t evolves as follows,

$$\frac{dX_t}{X_t} \;=\; (r_t - \delta_t)\,dt \;+\; \sqrt{V_t}\,dW_t^{\mathbb{Q}} \;+\; \int_{\mathbb{R}} x\widetilde{\mu}^{\mathbb{Q}}(ds,dx), \qquad \text{under } \mathbb{Q}, \tag{2}$$

where r_t is the instantaneous risk-free interest rate and δ_t is the dividend yield; $W_t^{\mathbb{Q}}$ is a Brownian motion with respect to the probability measure $\mathbb{Q}$; $\widetilde{\mu}^{\mathbb{Q}} = \mu - \nu^{\mathbb{Q}}$ for μ being the integer-valued measure in equation (1) and $\nu^{\mathbb{Q}}$ being a jump compensator of the form $\nu^{\mathbb{Q}}(dt,dx) = a_t dt \otimes \nu^{\mathbb{Q}}(dx)$ for some Lévy measure $\nu^{\mathbb{Q}}(dx)$.

We further assume that $V_t = f(\mathbf{S}_t)$ and $a_t = g(\mathbf{S}_{t-})$ for some $n\times 1$ dimensional stochastic process $\mathbf{S}_t$ following a Lévy-driven SDE, and $f(\cdot)$ and $g(\cdot)$ being arbitrary smooth functions: $\mathbb{R}^n \to \mathbb{R}_+$.

We denote European-style out-of-the-money option prices for the asset X at time t by $O_{t,k,\tau}$. Assuming frictionless trading in the options market, the option prices are given as,

$$O_{t,k,\tau} \;=\; \begin{cases} \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau} r_s\,ds}\,(X_{t+\tau}-K)^+\right], & \text{if } \; K > F_{t,t+\tau}\;, \\ \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^{t+\tau} r_s\,ds}\,(K-X_{t+\tau})^+\right], & \text{if } \; K \le F_{t,t+\tau}\;, \end{cases} \tag{3}$$

where τ is time-to-maturity, or tenor, K is the strike price, $F_{t,t+\tau}$ is the forward price for X at time t with respect to date $t+\tau$, and $k = \ln(K/F_{t,t+\tau})$ is the log-moneyness.

We further denote the Black-Scholes implied volatility corresponding to the option price $O_{t,k,\tau}$ (as well as risk-free rate r_t and dividend yield δ_t) by $\kappa_{t,k,\tau}$. Provided the log-moneyness k is non-random or, more generally, $\mathcal{F}_t$-adapted, the Markovian assumption on the state vector, $\mathbf{S}_t$, implies that $\frac{e^{r_{t,t+\tau}}O_{t,k,\tau}}{F_{t,t+\tau}}$ is a function only of the tenor, the state vector, and the moneyness (and t, if $\mathbf{S}_t$ is not stationary under $\mathbb{Q}$), where $r_{t,t+\tau}$ denotes the risk-free interest rate for the period $[t,t+\tau]$. Of course, these relationships carry over to $\kappa_{t,k,\tau}$ as well.

The volatility VIX index, computed by the Chicago Board Option Exchange (CBOE) from short-maturity out-of-the-money options on the S&P 500 index, has a theoretical value of,

$$VIX_{t,\tau} = \sqrt{\frac{1}{\tau}\mathbb{E}_t^{\mathbb{Q}}\left(\int_t^{t+\tau} V_s ds + \int_t^{t+\tau}\int_{\mathbb{R}} (e^x - 1 - x)\mu(ds,dx)\right)}. \tag{4}$$

The futures price for the one-month VIX index is theoretically valued as,

$$VF_{t,\tau} = \mathbb{E}_t^{\mathbb{Q}} \left[VIX_{t+\tau,30} \right]. \tag{5}$$

We note that, given the assumed structure for the process X under the risk-neutral measure, all the above option-based measures are (different) functions of the state vector $\mathbf{S}_t$ and, of course, the characteristics of the derivative contract, such as tenor and moneyness. This means that while $\mathbf{S}_t$ is not in general directly observed – it is "hidden" in the characteristics of X – we can study the derivatives contracts by nonparametric means to gain knowledge about certain features of the state vector dynamics. We provide much more details regarding this issue later on.

3 Data

We deliberately draw on multiple sources of information regarding the very high frequency behavior of S&P 500 volatility. If our findings are consistent across different securities, traded on distinct exchanges, and over fairly long periods of time, then they are less likely to be driven by microstructure effects or other idiosyncracies that are specific to a given contract, exchange, or sample. We exploit tick data for the E-mini S&P 500 futures and options from the CME Group (CME). The Chicago Board of Options Exchange (CBOE) provides us with high-frequency data for the VIX index and the VIX futures. Finally, we rely on Treasury bill rates, obtained from the U.S. Federal Reserve, to proxy for the risk-free interest rate.

Obviously, the payoff, and thus pricing, of the various CME and CBOE derivative contracts above are all linked to the behavior of the underlying S&P 500 index. As such, they are necessarily interconnected. Figure 1 provides a flowchart introducing the relevant contracts and how they relate to each other within the two exchanges. The contracts in the oval gray boxes are the ones utilized in this paper. At the top of the diagram, we find the 500 individual stocks, each listed on the NYSE or NASDAQ, which are the constituents of the index. Below them, we find the S&P 500 (cash) index itself, obtained as a value-weighted average of the stock prices. This may be viewed as the first layer of derivative, or synthetic index, in the chain. Moving onto the right (CME) branch of the chart, we first encounter the E-mini S&P 500 futures contract. The E-mini futures price is tied closely to the underlying (non-traded) S&P 500 cash index by strong arbitrage forces. The CME E-mini futures contract is routinely viewed as the primary location for price discovery in the aggregate U.S. equity market. There is also a very active market for options written on the E-mini S&P 500 futures contract. We exploit these options to derive various (Black-Scholes) implied volatility measures across the spectrum of option moneyness. As the E-mini futures options

represent the third link in the chain, one may think of our derived implied volatilities as fourth level derivative measures constructed from the underlying portfolio of individual constituent stocks.

Below the S&P 500 index, on the top of the left (CBOE) branch of Figure 1, we find the SPX options. The CBOE has created the (model-free) VIX volatility index from the cross-section of SPX option prices across the strikes at a fixed 30-day maturity. While the VIX is not directly traded, VIX futures and VIX options have been introduced, and both contracts are now very liquid. We construct a synthetic 30-day fixed maturity VIX futures price from the two shortest, and most liquid, VIX futures maturities traded at a given point in time. Our analysis exploits both the VIX index itself as well as this 30-day VIX futures price. Since the VIX futures represent the fourth layer of derivatives for the S&P 500 stocks, our fixed maturity VIX futures price is likewise a high level derivative relative to the original stock portfolio.

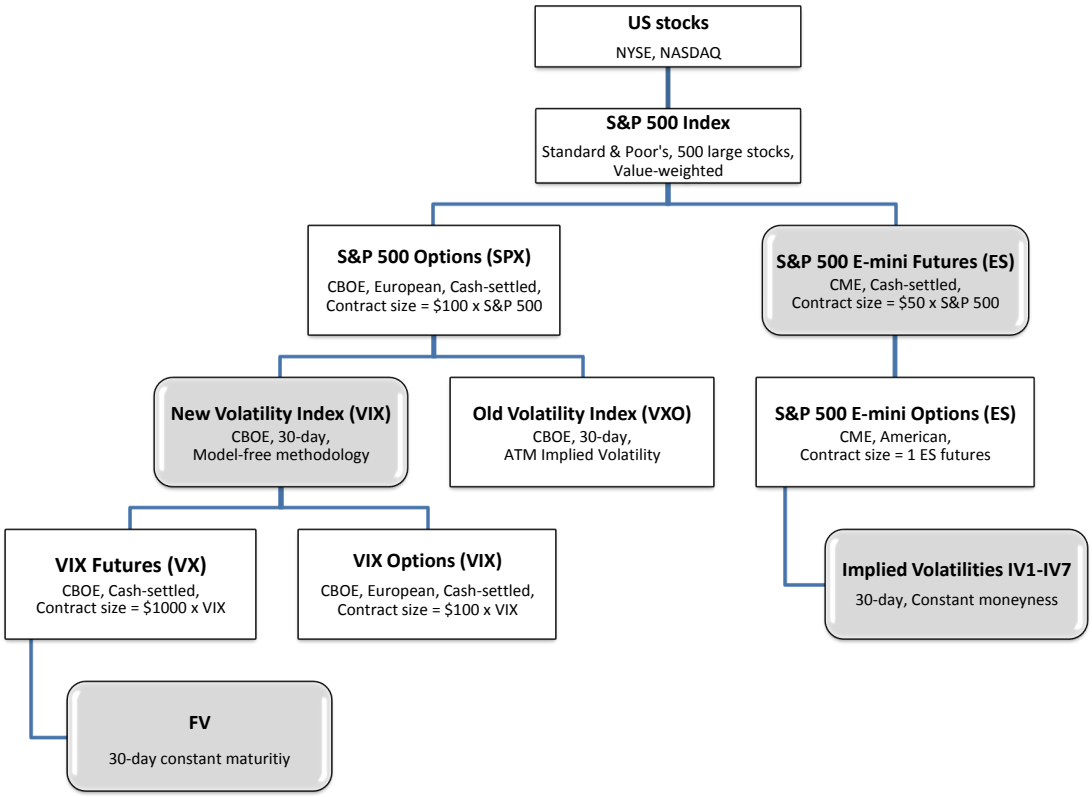

Figure 1: This flow chart illustrates the relationship between various S&P 500 volatility derivatives. The shaded boxes with rounded corners indicate series used in this study.

Figure 2 depicts the evolution of the S&P 500 index along with the VIX, 30-day VIX futures,

and three representative (Black-Scholes) implied volatility series. We discuss the construction of the individual series in detail below. We include the figure at this stage to provide an initial sense of the complementarity and the potential for incremental information obtained by exploring the full set of measures. As expected, the volatility series are highly correlated, but there are also noteworthy differences. For example, the VIX futures series is much less erratic than the VIX index, reflecting the longer effective maturity of the futures contracts (30-day to expiration plus the payoff reflecting the one-month forward-looking feature of the VIX). Likewise, the IV6 series for ATM options is dramatically less volatile than the corresponding IV2 series for deep out-of-the-money put options. We demonstrate below that these distinct patterns are aligned with corresponding discrepancies in the underlying fine structure of the price dynamics for the different measures.

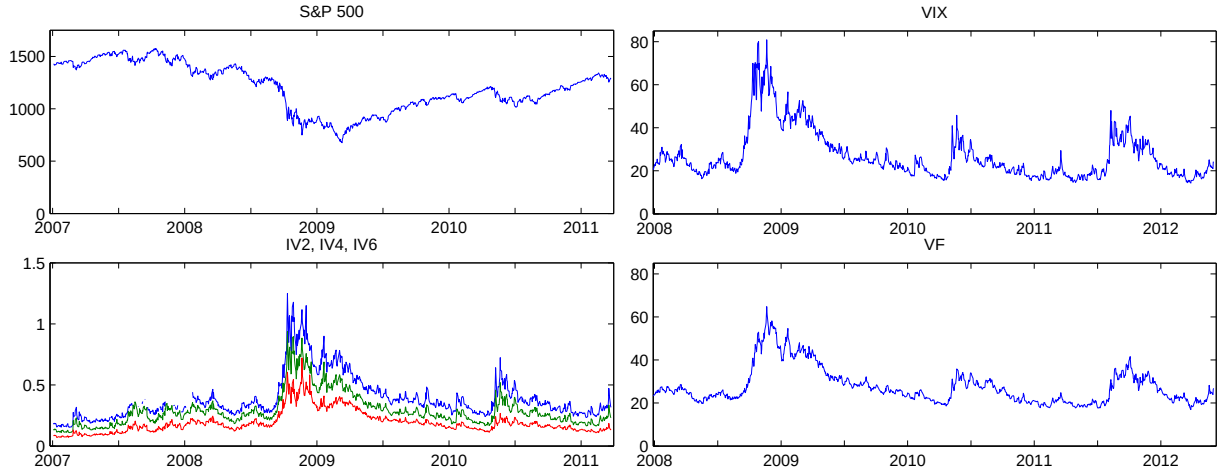

Figure 2: This figure plots daily values of S&P 500 in the top left panel, the implied volatilities IV2 (blue), IV4 (green), and IV6 (red) in the bottom left panel, the VIX index in the top right panel, and the 30-day VIX futures price, VF, in the bottom right panel.

3.1 E-mini S&P 500 Futures and Options

The E-mini S&P 500 futures contract and its options (commodity ticker symbol ES) are traded exclusively on the CME GLOBEX electronic platform. Trading takes place essentially 24 hours a day, five days a week. The markets for the ES futures and options are among the deepest and most liquid worldwide. Our sample period is January 3, 2007, to March 22, 2011, covering 1062 trading days. Appendix A provides additional information about these markets.

We construct 15-second series for both the futures and options using the "previous tick" method, implying we retain the last available quotes prior to the end of each 15 second interval. If no new quote arrives in a given 15-second interval, the last available quote prior to the interval is retained.

We focus on the main part of the regular trading hours, from 8:45 to 15:15 CT,[2] corresponding to 1560 15-second intervals per trading day. While the E-mini futures market is extremely liquid and generates fresh quotes at a much higher frequency than fifteen seconds, much less is known about the futures options, especially for strikes far from the current futures price, e.g., for the out-of-the-money options. To provide a meaningful concept of moneyness across time, we introduce a normalized moneyness measure, m. Letting σ_{BS} denote the ATM implied Black-Scholes volatility and F the forward price corresponding to tenor τ, we define,

$$m = \frac{\ln(K/F)}{\sigma_{BS}\sqrt{\tau}}\,. \tag{6}$$

Hence, in Figure 3, we plot the average values of various option liquidity measures versus moneyness, m. The top left panel depicts the average dollar bid-ask spread for the options, while the bottom left panel provides the normalized mid quote. It is evident that the bid-ask spread grows as the option prices increase, but much less than proportionally with price. In fact, for the deep out-of-the-money options, the spreads flatten so that (theoretically) very cheap options carry a non-trivial dollar spread. As a consequence, the upper right panel indicates a dramatic increase in the relative spread as we move away from the ATM options. Once we go beyond the moneyness of -4 to the left or 1.5 to the right, the spread exceeds 70% of the mid-quote value. Hence, transactions at the bid or ask will occur at prices that, on average, deviate from the quote midpoint by around 35% or more. For $m = -4$, the half-spread drops to about 25%, and it attains the much less elevated level of about 12% for $m = -3$ and $m = 1.5$. Finally, in the bottom right panel, we note a corresponding pattern for the liquidity expressed in terms of the average number of seconds between each quote update.[3] For $m = -4$ or $m = 2$, the quote frequency falls short of the 15 second mark, while for $m \in [-3, 1.5]$ the update frequencies are about 10 seconds or less, and around $m = 0$ they are about only one second. Finally, we face the issue that the strike range offered by the CME at times can be limited. For example, for $m = -4$, there is no strike available around or below this point for about 16% of our observations, so this region of moneyness is simply not covered during parts of our sample. Likewise, the corresponding non-coverage rates for $m = -3$ and $m = -2$ are 7.3% and 1.8%, respectively. For options around-the-money this problem vanishes, with only a fraction of a percent of observations missing for $m = -1$ or $m = 0.5$, and none for $m = 0$. However, for $m = 1$, non-coverage jumps to fully 47% of the observations, leading to

[2]Since the market occasionally faces irregularities in the quoting or transmission process during the initial period following the cash market opening, we are conservative and remove the first fifteen minutes of the regular trading day.

[3]A quote update verifies that, at the given point in time, the option may be traded at the specified bid and ask prices, i.e., the quotes are "actionable."

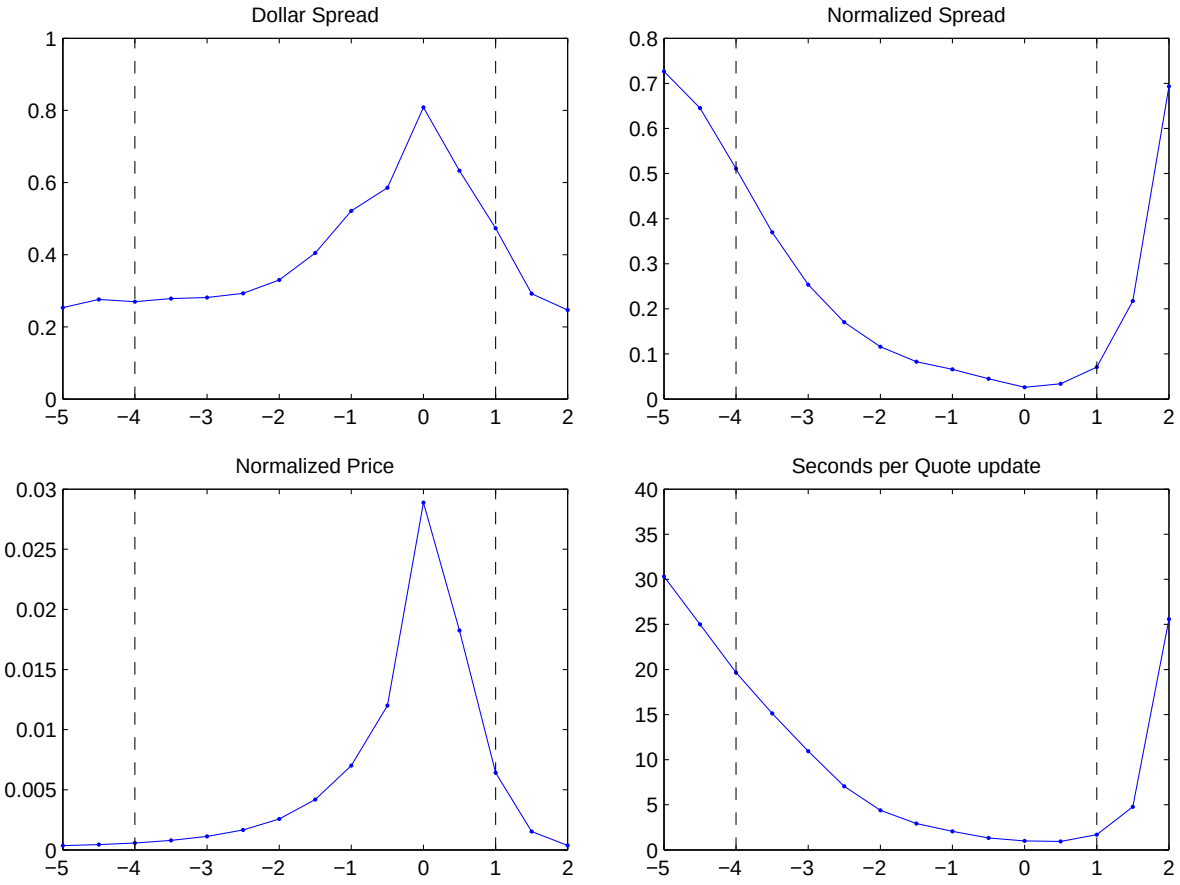

Figure 3: This figure plots average values of various option liquidity measures for different levels of the normalized moneyness m. For $m < 0$, puts are used; for $m > 0$, calls are used; and for $m = 0$, the average of puts and calls is computed. The dollar spread is $A - B$, where A and B are the dollar ask and bid prices; the normalized spread is $(A-B)/M$, where $M = \frac{1}{2}(A+B)$ is mid quote price; the normalized price is M/F. The last panel shows the average number of seconds per quote update. The dashed vertical lines indicate the range of moneyness m used in our analysis.

some concern about the integrity of our procedures due to potential endogenous sampling issues for this region of moneyness. The coverage deteriorates even more for further out-of-the-money call options, leading us to drop the $m = 1.5$ value from our analysis, even if – as is evident from Figure 3 – trading is very active for this degree of moneyness whenever coverage does exists.

In conclusion, we have a reasonable amount of timely quotes across a broad range of moneyness for the 30-day maturity E-mini futures options. However, at the 15-second frequency, there will, inevitably, be some staleness as well as price uncertainty associated with the bid-ask spread. Thus, we do not use the 15-second observations directly, but rely on quote mid-points from a lower frequency such as, e.g., five minutes. We do this in a couple of ways. The direct approach is to

use only every 20^{th} observation in the 15-second series. However, inspired by Podolskij and Vetter (2009), we can exploit additional information and simultaneously alleviate lingering microstructure effects by pre-averaging the 15-second quotes within five-minute blocks and use the resulting series as input to the subsequent analysis. The latter procedure is the one underlying our empirical results in Section 5. Since our theory is developed without consideration of microstructure noise, it is appealing with an implementation scheme that mitigates distortions arising from this source.

Given the above evidence concerning option liquidity, we construct seven separate 30-day implied (Black-Scholes) volatility measures IV1-IV7 for each 15-second interval, corresponding to the following levels of moneyness, $m = -4, -3, -2, -1$, 0, 0.5, and 1.[4] However, before including a specific observation in our analysis, we impose a number of checks regarding the validity of the underlying quotes. We defer the details of our construction to Appendix A.1. Lastly, when interpreting the final results, we avoid placing excessive weight on conclusions that stem from the furthest out-of-the-money options, as these are more prone to noise and measurement error.

3.2 VIX and VIX Futures

The volatility index VIX is disseminated in real time by the CBOE. The index is based on the model-free methodology and is equal to the square root of the par 30-day variance swap. The index is calculated from options prices across a range of strikes on the S&P 500 index and is quoted as an annualized standard deviation.[5] While the VIX index is not directly a traded asset, the VIX futures (ticker symbol VX) are traded on the CBOE Command trading platform, with trading from 7:00 am to 3:15 pm CT every weekday. Appendix B provides additional details about these markets.

The VIX futures contract started trading on March 26, 2004. Initially, the volume was low but it has since increased dramatically, and the contract is now considered one of the most successful CBOE product launches of all time. Our tick data for the VIX index and VIX futures were obtained directly from CBOE Market Data Express (MDX). Table 1 provides daily summary statistics for the two nearby (and most liquid) VIX futures contracts, denoted FV1 and FV2, with tenors spanning 0-30 and 30-60 days, respectively. Between 2006 and 2012, the volume grew more than 45-fold, the number of quotes increased more than one hundred-fold and the percentage spreads declined to less than a third of the original size. Certainly, by 2008, the liquidity of the VIX futures market is good enough to ensure that new quotes arrive in almost every 15 second interval during regular trading

[4]The implied volatility for a given m is obtained by linearly interpolating the IVs for strikes just below and above m. This is identical to the procedure used by the CBOE in obtaining the at-the-money Black-Scholes IV.

[5]The VIX should not be confused with the "old" VIX, now re-named VXO. The VXO index represents the Black-Scholes implied volatility of the 30-day ATM option on the S&P 100 index.

hours. Thus, we rely on the period January 2, 2008, to May 31, 2012, covering 1112 trading days, and we construct 15-second series using the "previous tick" method from 8:45 to 15:15 CT every trading day.[6] Moreover, like for the equity futures and IV series, we pre-average the 15-second quotes for the VIX and VIX futures within non-overlapping five-minute blocks.

Finally, as for the IV measures, we convert the VIX futures series with varying tenor into a fixed 30-day maturity series, denoted FV, through a weighted linear interpolation of the prices for the two nearby futures contracts.

Table 1: Daily Statistics for VIX Futures

First Maturity VF1							
	2006	2007	2008	2009	2010	2011	2012
Trading Volume, 000s	0.58	1.98	2.10	1.97	8.17	18.89	27.37
Number of Quotes, 000s	0.66	1.40	6.75	11.15	70.90	217.72	109.13
Normalized Spread, %	0.82	0.55	0.42	0.34	0.26	0.24	0.26
Second Maturity VF2							
	2006	2007	2008	2009	2010	2011	2012
Trading Volume, 000s	0.41	0.79	1.13	1.46	5.49	13.20	21.00
Number of Quotes, 000s	0.20	0.72	6.09	9.85	57.69	164.52	94.33
Normalized Spread, %	0.83	0.63	0.55	0.36	0.26	0.23	0.24

Note: VF1 and VF2 denote the two nearby maturities. The normalized spread is $(A-B)/M$, where A and B are the ask and bid prices and $M = \frac{1}{2}(A+B)$ is mid quote price.

4 Econometric Tools for the Empirical Analysis

4.1 The Formal Framework

In this section we present the various estimators and tests that we apply later in our empirical analysis. We denote the generic process under investigation by Z. This will be one of the option based measures or the S&P 500 index futures price. Our econometric analysis proceeds under the assumption that the dynamic evolution of Z may be captured by the following general specification,

$$dZ_t = \alpha_t dt + \sigma_{t-} dS_t + dY_t, \qquad 0 \leq t \leq T, \tag{7}$$

[6] For the VIX series we apply a very mild filter by removing observations that fall outside the daily high-low range, as reported at the end of trading. This correction reflects errors that have been recognized by the exchange.

where α_t and σ_t are processes with càdlàg paths; Y_t is a process whose second characteristic is identically zero, i.e., it is a process without a continuous local martingale component; S_t is a (strictly) stable process with characteristic function (see, e.g., Sato (1999)),[7]

$$\log\left[\mathbb{E}(e^{iuS_t}\right] = -t|cu|^{\beta}. \tag{8}$$

Note that Y_t need not be independent of S_t (or α_t and σ_t). Hence, Z_t does not necessarily "inherit" the tail properties of the stable process S_t. This implies, for example, that Z_t can be driven by a tempered stable process whose tail behavior is very different from that of the stable process. The model in equation (7) covers most, if not all, models routinely used for modeling financial asset prices. Key examples include the affine jump-diffusion model class of Duffie et al. (2000), the time-changed Lévy models of Carr et al. (2003), and the Lévy-driven SDEs, provided the Lévy density is locally stable (which is almost always true for parametric specifications of Lévy densities).

For $\beta = 2$, the S_t process is a Brownian motion, and it constitutes the (first order) leading term driving Z_t in equation (7) at high frequencies. For $1 < \beta < 2$, S_t is a pure jump process, but it continues to (first order) dominate the drift term at high frequencies. Throughout, we restrict attention to this empirically relevant case, i.e., $1 < \beta \leq 2$. The self-similarity of the strictly stable process implies the following scaling property,

$$S_t - S_s \overset{d}{=} |t-s|^{1/\beta} S_1, \quad \forall\ 0 \leq s < t, \tag{9}$$

Finally, to retain the interpretation of S_t as the dominant (high frequency) term – and separately identify S_t from Y_t in this regard – we require Y_t to satisfy a scaling bound for high frequencies of the following type. There exists a sequence of stopping times increasing to infinity, T_p, and $\beta' < \beta$, such that for $\forall\ 0 \leq s < t$, we have,

$$\mathbb{E}|Y_{t\wedge T_p} - Y_{s\wedge T_p}|^q \leq K_p|t-s|, \quad \forall q > \beta', \tag{10}$$

where K_p is some constant depending on the sequence of stopping times.

These conditions imply that Y_t, in equation (7), plays the role of a high-frequency "residual" component. Combining the self-scaling result in equation (9) and the scaling bound on Y in equation (10), we obtain, under suitable regularity, for $s \in [0,1]$ and $\forall t \geq 0$ with $\sigma_t > 0$,

$$h^{-1/\beta}\frac{Z_{t+sh} - Z_t}{\sigma_t} \xrightarrow{\mathcal{L}} S'_{t+s} - S'_t, \quad \text{as } h \downarrow 0, \tag{11}$$

where S'_t is a Lévy process with a distribution identical to that of S_t, and the convergence is defined in the space of càdlàg functions equipped with the Skorokhod topology.

[7]This setting can be extended to allow for S_t being an asymmetric stable process. This is accommodated by differencing the increments in the statistics below, as proposed in Todorov (2013).

Intuitively, the result in equation (11) implies that each high frequency increment of Z behaves like that of a stable process with a constant scale σ_t known at the onset of the increment. If we directly observe the stochastic process σ_t, we can scale the increments of Z accordingly, and the limit result in (11) states that the scaled increments should be i.i.d. stable (in the limit for ever more frequent sampling). Separating the case $\beta = 2$ from $\beta < 2$ is of central importance as it is equivalent to distinguishing between Z being a jump diffusion or of the pure jump type.

4.2 Estimation and Inference Procedures

The goal of the empirical analysis in Section 5 is to test, using high-frequency observations, whether Z is a jump diffusion or a pure jump process and, furthermore, to estimate the activity index β. This section presents the econometric tools we apply for that task. We assume the process Z is observed at the equidistant grid points $0, \frac{1}{n}, ..., T$ with a fine mesh, $\frac{1}{n}$, that asymptotically approaches 0, and T is a fixed positive integer. We will use either the increments $\Delta_i^n Z = Z_{\frac{i}{n}} - Z_{\frac{i-1}{n}}$ or their so-called pre-averaged analogue (see Jacod et al. (2009) and Jacod et al. (2010)) constructed via additional sampling within each interval. The pre-averaged increments are defined as,

$$\widetilde{\Delta_i^n Z} = \sum_{\ell=1}^{L} \left\{ \left(X_{\frac{i-1}{n}+\frac{\ell}{nL}} - X_{\frac{i-1}{n}+\frac{\ell-1}{nL}} \right) \left(\frac{\ell}{L} \bigwedge \frac{L-\ell}{L} \right) \right\}, \quad i = 0, \frac{1}{n}, ..., T. \tag{12}$$

Given the available observation frequency of our data, we restrict attention to the case of L fixed. In this setting, the asymptotic behavior of our statistics is the same irrespective of whether $\widetilde{\Delta_i^n Z}$ or $\Delta_i^n Z$ is used. The purpose of the pre-averaging is simply to mitigate the potential impact of moderate microstructure noise on the estimation, as the formal asymptotic analysis below does not account for microstructure noise in the observations.

We consider two types of tests for the presence of a diffusion component, exploiting different aspects of the limiting result in equation (11). The first test is based on direct estimates of the index β for S_t. It is constructed from the ratio of power variations at two distinct frequencies. Formally, letting $1\{A\}$ be one for A true, and zero otherwise, our estimator of β is,

$$\widehat{\beta}(p) = \frac{p \log(2)}{\log\left(V_T^{n,2}(p,Z)\right) - \log\left(V_T^{n,1}(p,Z)\right)} \cdot 1\left\{ V_T^{n,2}(p,Z) \neq V_T^{n,1}(p,Z) \right\}, \tag{13}$$

where,

$$V_T^{n,1}(p,Z) = \sum_{i=1}^{nT} \left| \widetilde{\Delta_i^n Z} \right|^p, \qquad V_T^{n,2}(p,Z) = \sum_{i=2}^{nT} \left| \widetilde{\Delta_{i-1}^n Z} + \widetilde{\Delta_i^n Z} \right|^p. \tag{14}$$

This represents an extension of the activity estimator of Todorov and Tauchen (2011a,b), allowing for overlapping observations in the construction of $V_T^{n,2}(p,Z)$, as in Todorov (2013), to enhance

efficiency.[8] The estimator is consistent for β, when $p < \beta$. Moreover, under suitable conditions on Z and, importantly, $p < \beta/2$, we obtain the following asymptotic limit result,

$$\sqrt{n}\left(\widehat{\beta}(p) - \beta\right) \xrightarrow{\mathcal{L}-s} \frac{\sqrt{\int_0^T |\sigma_s|^{2p} ds}}{\int_0^T |\sigma_s|^p ds} \times \frac{\beta^2}{\mu_p(\beta) p \log(2)} \sqrt{\widetilde{\Xi}} \times \mathcal{N}. \tag{15}$$

$\mathcal{N}$ denotes a standard normal random variate which is defined on an extension of, and is independent from, the original probability space. Meanwhile, $\widetilde{\Xi}$ is given by,

$$\widetilde{\Xi} = \Xi^{(1,1)} - 2^{1-p/\beta}\Xi^{(1,2)} + 2^{-2p/\beta}\Xi^{(2,2)},$$

where $\Xi = \Sigma_0(p,\beta) + \Sigma_1(p,\beta) + \Sigma_1'(p,\beta)$ and $\Sigma_j(p,\beta) = \mathbb{E}(\mathbf{Z}_1\mathbf{Z}_{1+j}')$ for $j = 0, 1$ with $\mathbf{Z}_i = \left(|S_i|^p - \mu_p(\beta),\ \ |S_i + S_{i+1}|^p - 2^{p/\beta}\mu_p(\beta)\right)'$, $\mu_p(\beta) = \mathbb{E}|S_i|^p$, and $S_0, S_1, \ldots$ are i.i.d. β-stable random variables, defined through the characteristic function in equation (8) for $t = 1$. Consistent estimators for the integrated power variation terms that appear in the limiting distribution are readily obtained from corresponding realized power variation statistics; we refer to Todorov and Tauchen (2011a) for details on this as well as an enumeration of the requisite regularity conditions.

Our second test is based on the distributional implications of the small scale result in equation (11) and follows from the developments in Todorov and Tauchen (2013). In particular, we split the data into blocks of asymptotically shrinking length and form local volatility estimates for each block based only on observations from within the block. Next, we use these estimates to rescale (de-volatilize) the high-frequency increments within the block, in order to "annihilate" the stochastic volatility. Finally, we estimate the empirical cumulative distribution function (cdf) of the de-volatilized high-frequency increments. The test is then based on the distance between this empirical cdf and that of the standard normal, which constitutes the limiting distribution in equation (11) under the null of Z being a jump-diffusive process over the given time interval.

To formally define the procedure, we require some notation. Each block contains k_n high-frequency increments, each of length $1/n$, with $k_n \to \infty$ and $k_n/\sqrt{n} \to 0$. There are a total of $J_n = \lfloor Tn/k_n \rfloor$ separate blocks, $j = 1, \ldots, J_n$. The set of increments within block j is $\mathcal{I}_j = \{(j-1)k_n + 1, \ldots, jk_n\}$, while the same index set excluding element i is denoted $\mathcal{I}_j(i) = \mathcal{I}_j \setminus \{i\}$. We exploit the so-called Truncated Variation statistic of Mancini (2009) for estimating the quadratic variation of the diffusion component of Z. For each increment we define the following (noisy) volatility estimator, scaled to express the return variation per unit time interval,

$$\widehat{c}_i^n \ = \ n \cdot |\widetilde{\Delta_i^n Z}|^2 \cdot 1\left(|\widetilde{\Delta_i^n Z}| \leq \alpha n^{-\varpi}\right), \qquad i = \frac{1}{n}, \frac{2}{n}, \ldots, J_n. \qquad \alpha > 0,\ \varpi \in (0, 1/2). \tag{16}$$

[8] See Ait-Sahalia and Jacod (2010) for an estimator based on truncated power variation.

Our local volatility estimator for block j improves precision by averaging over the individual contributions and, again, scaling to reflect the return variation per time unit,

$$\widehat{C}_j^n = \frac{1}{k_n} \sum_{\iota \in \mathcal{I}_j} \widehat{c}_\iota^n, \qquad j = 1, \ldots, J_n \,. \tag{17}$$

However, it turns out that feasible inference requires we use a smaller number of increments within the block, $m_n \ll k_n$, in the construction of the empirical cdf. Furthermore, as we scale each high-frequency increment during the computation of the empirical cdf, we must ensure that the effect of that specific increment in the calculation of the local volatility estimator is dampened and, in fact, asymptotically vanishing. Thus, we introduce an alternative estimator for volatility (per unit time) over block j that excludes the impact from increment i,

$$\widehat{C}_{j(i)}^n = \frac{1}{k_n - 1} \sum_{\iota \in \mathcal{I}_j(i)} \widehat{c}_\iota^n, \qquad j = 1, \ldots, J_n \,. \tag{18}$$

The block volatility estimator in equation (17) is then modified as follows,

$$\widehat{C}_j^n(i) \;=\; \widehat{C}_j^n \,+\, \frac{\eta_n}{k_n - 1} \left(\widehat{C}_j^n \,-\, \widehat{c}_i^n \right) \;=\; \widehat{C}_{j(i)}^n \,+\, \frac{1 - \eta_n}{k_n} \left(\widehat{c}_i^n \,-\, \widehat{C}_{j(i)}^n \right), \tag{19}$$

where $\eta_n = 1 - \frac{m_n}{k_n}$. Since $0 < \eta_n < 1$ and $\eta_n \to 1$, the middle expression in equation (19) indicates the dampening of the contribution from the i$^{\text{th}}$ increment, while the expression to the right shows how the impact of $\widehat{c}_i^n$ is eliminated asymptotically, as all weight shifts to $\widehat{C}_{j(i)}^n$ in the limit.

Now, normalizing the increments by the modified local volatility estimator and exploiting m_n rather than k_n terms in each block, the empirical cdf of the de-volatilized high-frequency increments is given as,

$$\widehat{F}_n(\tau) \;=\; \frac{1}{N^n(\alpha, \varpi)} \sum_{j=1}^{J_n} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left\{ \frac{\sqrt{n} \widetilde{\Delta_i^n Z}}{\sqrt{\widehat{C}_j^n(i)}} \leq \tau \right\} 1 \left\{ |\widetilde{\Delta_i^n Z}| \leq \alpha n^{-\varpi} \right\}, \tag{20}$$

where, for the identical $\alpha > 0$ and $\varpi \in (0, 1/2)$ as in equations (17), we have defined,

$$N^n(\alpha, \varpi) = \sum_{j=1}^{\lfloor n/k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} 1 \left(|\widetilde{\Delta_i^n Z}| \leq \alpha n^{-\varpi} \right), \tag{21}$$

so $N^n(\alpha, \varpi)$ denotes the total number of increments used in the formation of the empirical cdf.

Under appropriate regularity, Todorov and Tauchen (2013) show that if Z is a jump-diffusion, i.e., S_t in equation (7) is a Brownian motion, we have,

$$\widehat{F}_n(\tau) \;\xrightarrow{\mathbb{P}}\; F(\tau), \tag{22}$$

where $F(\tau)$ is the cdf of a standard normal random variate and the convergence is uniform in $\tau \in \mathbb{R}$. Again, there is a CLT associated with this convergence in probability result, enabling us to construct a distribution-based test for whether Z is a jump diffusion via the Kolmogorov-Smirnov distance between the empirical cdf $\widehat{F}_n(\tau)$ and the cdf $F(\tau)$ of a standard normal variate. To conserve space, we refer to Todorov and Tauchen (2013) for details.

Finally, our choice of tuning parameters for constructing $\widehat{F}_n(\tau)$ in equation (22) reflects standard practice as well as the underlying observation frequency. Letting one trading day represent our unit interval, we have $n = 78$ five-minute increments for most of our series. Moreover, we set $k_n = 26$, so that we have three blocks per day, and $m_n = 20$, i.e., our (finite sample) implementation exploits about 75% of the increments in constructing the empirical cdf. Finally, following prior work, we set the truncation parameters to $\varpi = 0.49$ and $\alpha = 3\sqrt{\widehat{B}_j^n}$, where,

$$\widehat{B}_j^n = \frac{\pi}{2}\frac{n}{k_n - 1}\sum_{i=(j-1)k_n+2}^{jk_n} |\widetilde{\Delta_i^n Z}||\widetilde{\Delta_i^n Z}|, \quad j = 1, ..., \lfloor n/k_n \rfloor, \tag{23}$$

is the Bipower Variation statistic of Barndorff-Nielsen and Shephard (2006). Under our null that Z is a jump diffusion, it consistently estimates the diffusive quadratic variation. That is, we apply a time-varying threshold to ensure a better finite sample separation of the "big" jumps from the purely continuous increments. Lastly, to account for the well-known diurnal pattern in volatility, the raw high-frequency returns are standardized by a time-of-day scaling factor prior to analysis.

5 Empirical Findings

This section implements the tests introduced in the previous section for the alternative data series introduced in Section 3. Before progressing to the empirical results, we briefly discuss how the activity indices, i.e., the "fine structure" of the return dynamics, should be interconnected, in theory, for the different derivatives instruments and the underlying asset.

5.1 Linkages in the Fine Structure of Returns across Assets

For clarity and reference, we initially assume that the latent state vector, $\mathbf{S}_t$, governing the stochastic volatility and (risk-neutral) jump intensity dynamics of the underlying equity index, X_t, introduced in Section 2, is a univariate Markov process. In this case, we may readily show that $VIX_{t,\tau}$ is a (smooth) function of the value of the latent state variable at time t, $\mathbf{S}_t$, as well t and τ. Similar reasoning implies that $VF_{t,\tau}$ are smooth functions of $\mathbf{S}_t$ and the tenor of the contracts. In addtion, the Markovian assumption implies that the ratio $\frac{e^{r_{t,t+\tau}} O_{t,k,\tau}}{F_{t,t+\tau}}$ is a function only of $\mathbf{S}_t$

(and the specific characteristics of the option contract), as long as the measure of log-moneyness, k, is either a fixed constant or a function solely of the state vector, i.e., no extraneous variables enter the expression. Moreover, these properties carry over to the Black-Scholes implied volatility corresponding to $O_{t,k,\tau}$, as the latter also is a smooth transform of the ratio $\frac{e^{r_{t,t+\tau}} O_{t,k,\tau}}{F_{t,t+\tau}}$. Hence, the implied volatility measures, constructed from out-of-the-money S&P 500 options in Section 3, and henceforth abbreviated IV1-IV7, are functions of $\mathbf{S}_t$ alone as well.

In summary, in the above scenario, all option-based derivative prices are directly related to the state variable, $\mathbf{S}_t$, through a smooth functional. It follows, via Itô calculus, that the local behavior of these series must "inherit" the qualitative features of the underlying process, i.e., they share the "fine structure" of the increments which, in turn, is determined by the activity index. In other words, they all contain a diffusive component or they are all pure (infinite activity) jump processes. Likewise, the jump activity of all our option-based quantities equals that of $\mathbf{S}_t$. We stress, however, that the actual jump sizes will differ as the valuation of the various derivative securities arise from very different (smooth) transformations of $\mathbf{S}_t$. What (typically) coincides is the timing of jumps across the series. In conclusion, under the univariate Markovian assumption, the ability to identify and estimate the activity index is greatly enhanced by exploiting information across the series.

Matters are less straightforward when the state vector, $\mathbf{S}_t$, is multivariate. While the option-based derivatives prices continue to be smooth functions of the vector $\mathbf{S}_t$, the features dominating the local behavior of the derivatives price process now depend critically on the relative sensitivity to distinct components of $\mathbf{S}_t$. For example, recent studies find that the equity-index option surface dynamics is governed by different factors, with some including a diffusive term and others being governed by pure jump processes.[9] In this case, even if the stochastic volatility is a jump diffusion, the price path for certain derivative securities will be dominated by the exposure to the pure jump component of the latent state vector $\mathbf{S}_t$. Under such circumstances, the activity index for the derivative security is determined by the largest such index across the subcomponents of $\mathbf{S}_t$ for which it has a meaningful exposure. In short, the fine structure of the processes governing the option-based derivative prices and jump activity indices will hinge on their specific type of exposure to the underlying asset dynamics. They will only share the identical activity index if they all remain sensitive to the full set of elements in the state vector $\mathbf{S}_t$. Since derivatives contracts, deliberately, are designed to offer unique exposures to, and thus span, distinct features of the risks driving the underlying asset, it is plausible we will encounter systematic differences in this regard across alternative derivatives with diverging exposures to underlying asset risks.

[9] See, e.g., Andersen et al. (2013b) for empirical support for this type of model.

5.2 Activity Estimates for our Equity-Index related Securities

Our estimators and test statistics are constructed from increments obtained by pre-averaging the series over five-minute blocks with an underlying frequency of 15 seconds, as described in Section 3. The lower frequency increments are then obtained by simply adding two consecutive increments based on the five-minute blocks. Table 2 reports estimates of the activity index for S_t in equation (7), along with a one-sided test for the null hypothesis $\beta = 2$, i.e., a diffusion component is present in the return dynamics, against the alternative that the process is of the pure jump type. The point estimates for the various securities are obtained via equation (13), exploiting the ratio of the p^{th} order power variations for a few different values of p. We obtain separate estimates across the sample for 22 non-overlapping trading day periods, i.e., months. The tables summarize the point estimates through robust statistics, indicating the median monthly estimate, $\text{med}(\hat{\beta}(p))$, as well as the median deviation from this median value, MAD. Likewise, we provide the rejection frequencies for the one-sided test of $\beta = 2$ at the five and one percent levels.

In Table 2 there is a sharp contrast between our findings for the equity (futures) returns and the VIX index. The point estimates suggest that the S&P 500 index is a jump diffusion while the VIX measure is best characterized as a pure jump process with jumps of infinite variation and an activity index in the range 1.5-1.6. Moreover, for the VIX series, the rejection rates obtained using $p = 0.90$ are consistent with those we would expect if the null hypothesis is false, but the driving jump process is highly active with an index between 1 and 2; see the simulation evidence in Todorov and Tauchen (2010). We also note that the estimates for the S&P 500 and VIX indices are similar to those reported in Todorov and Tauchen (2011b) although the sample periods, 2007-2012 versus 2003-2008, are quite different, thus providing a useful robustness check.[10] Turning to the activity estimates for our fixed 30-day maturity VIX futures series, we find the evidence supportive of our findings regarding the VIX index, as it also seems to be of the pure-jump type. Nonetheless, the estimates for the activity index are now somewhat larger, falling in the 1.64-1.81 range.

Moving on to the results for the implied volatility measures based on S&P 500 options in Table 2, we observe an interesting pattern. The implied volatilities corresponding to the out-of-the-money put options all have activity index estimates close to those for the VIX index, although the point estimates increase as we move from the far out-of-the-money towards the at-the-money measures. On the other hand, the implied volatilities for (normalized) log-moneyness $m = 0.5$ yield estimates slightly in excess of 2, and the test is consistent with the presence of a diffusion term for this series.

[10] Another distinction between the current activity estimates and Todorov and Tauchen (2011b) is that we now use overlapping observations in the construction of $V_T^{n,2}(p, Z)$ in (14), enhancing efficiency.

Obviously, this runs counter to the estimates for all our other option-based derivatives contracts.[11] We defer the discussion of these features to Section 5.3, after we have presented the full range of evidence.

Finally, we note that our estimator for the activity index in equation (13) depends solely on the ratio of power variation measures obtained at two distinct sampling frequencies. As such, the performance of the estimator hinges on our ability to estimate this ratio with reasonable precision. While we do not observe actual realizations of the power variation statistics, we may check some of the implied properties of our procedures. Specifically, the fine structure of the increments for the various processes should be invariant to the level of volatility. On the other hand, if microstructure effects or other issues associated our empirical power variation measures induce distortions, we may expect to observe systematic biases as a function of the volatility level. Thus, in Figure 4, we plot our activity estimates $\widehat{\beta}(0.7)$ against the corresponding (nonparametric) realized volatility measure, $\log(RV)$, for each month in the sample. The figure displays only the estimates for the S&P 500 futures and the IV3 series, as they are representative of measures generating starkly different implications for the activity index. In both cases, the plots reveal no significant dependence of the activity estimates on the level of RV, thus confirming that our empirical activity measures are largely invariant with respect to the fluctuations in volatility. The corresponding plots for the remaining series (not displayed) produce identical conclusions.

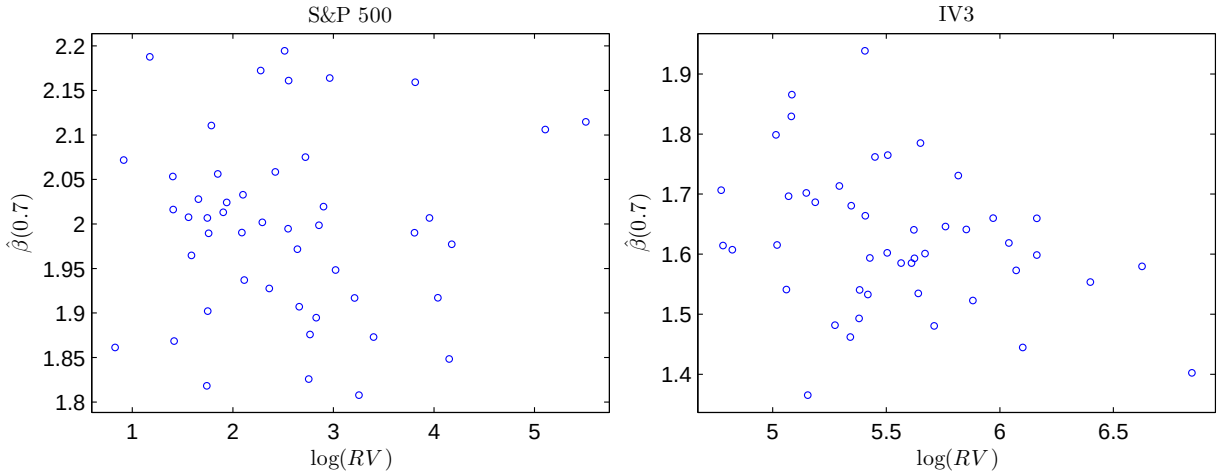

Figure 4: *Scatter Plots for* $\widehat{\beta}(0.7)$. *RV stands for the realized volatility over the period over which* $\widehat{\beta}(p)$ *is computed.*

[11] The estimated activity index may exceed 2 due to either regular sampling error or the presence of microstructure noise. In fact, the latter causes our activity estimator to diverges as we approach continuous sampling.

Table 2: Activity Index Estimates and Diffusion Tests for the Implied Volatility Measures

	p	med$(\hat{\beta}(p))$	MAD	5%	1%
IV1	0.50	1.57	0.07	94.74	81.58
	0.70	1.63	0.07	84.21	55.26
	0.90	1.70	0.09	57.89	31.58
IV2	0.50	1.55	0.11	93.02	81.40
	0.70	1.60	0.13	86.05	69.77
	0.90	1.65	0.13	72.09	51.16
IV3	0.50	1.57	0.07	95.65	80.43
	0.70	1.61	0.07	89.13	69.57
	0.90	1.65	0.07	82.61	54.35
IV4	0.50	1.70	0.07	76.60	46.81
	0.70	1.72	0.08	65.96	36.17
	0.90	1.74	0.08	57.45	23.40
IV5	0.50	1.73	0.07	57.45	36.17
	0.70	1.74	0.08	59.57	34.04
	0.90	1.75	0.08	57.45	27.66
IV6	0.50	2.04	0.13	0.00	0.00
	0.70	2.03	0.11	0.00	0.00
	0.90	2.04	0.11	0.00	0.00
IV7	0.50	1.82	0.06	36.84	15.79
	0.70	1.88	0.08	21.05	5.26
	0.90	1.93	0.12	5.26	0.00
S&P 500	0.50	1.97	0.07	2.08	0.00
	0.70	2.00	0.07	0.00	0.00
	0.90	2.02	0.06	0.00	0.00
VIX	0.50	1.50	0.13	96.00	88.00
	0.70	1.56	0.11	96.00	82.00
	0.90	1.60	0.09	88.00	68.00
VF	0.50	1.64	0.11	74.00	64.00
	0.70	1.72	0.11	60.00	38.00
	0.90	1.81	0.11	42.00	6.00

Note: The estimator, $\widehat{\beta}(p)$, is defined in equation (13). The rejection rates in the last two columns are based on a one-sided test for $\beta = 2$, as well as the point estimate $\widehat{\beta}(p)$ and the corresponding estimate for the asymptotic variance provided in equation (15). The estimation and tests are performed over periods of 22 trading days.

We next present results from the alternative distribution-based test for a jump diffusion based on the empirical cdf of the de-volatilized high-frequency increments given in equation (20). As discussed in the previous section, this test is based on "distributional implications" stemming from

the presence of a diffusion term in the high frequency increments. As such, it serves as an additional robustness check for the above findings. This is useful as the simulation evidence in Todorov and Tauchen (2013) suggests that the activity estimator, unlike the distribution-based test, can be somewhat sensitive to microstructure noise and thus may, erroneously, signal the existence of a diffusive term in the price process. The findings from our distribution-based Kolmogorov-Smirnov test are reported in Table 3. To highlight potential deviations between the empirical distribution of the de-volatilized increments and the standard normal, we split the domain of the test into four regions, reflecting the quantiles of the standard normal distribution. A key advantage of the test is that the limiting distribution of the difference $\widehat{F}_n(\tau) - F(\tau)$ is independent of volatility, σ_t, and hence the critical values for the test are identical across assets and time periods. These critical values are reported in the bottom row of the table.

The results in Table 3 are largely in agreement with the point estimates for the activity index reported in Table 2. In particular, the test statistics for the S&P 500 futures and the implied volatility series IV6 are below the critical values. Hence, this test also is consistent with the presence of a diffusion term in these two series. Moreover, unlike the tests based on the activity index, the Kolmogorov-Smirnov test cannot reject a null hypothesis that IV5 contains a diffusion. For all remaining series, however, the test statistics uniformly exceed the critical values, thus rejecting the null of a jump-diffusion. Moreover, we again observe an interesting continuity as the implied volatility series "surrounding" IV5 and IV6, i.e., IV4 and IV7, display weaker rejections of the null hypothesis than the other series. Another interesting pattern is that the rejections of the null of a jump-diffusion model invariably are most significant for the de-volatilized increments falling within the 20-40 and 40-60 quantiles of the normal distribution. To visualize this observation, Figure 5 provides kernel estimates for the log-density of the de-volatilized high-frequency increments against the log-density of the standard normal for our representative series, the S&P 500 futures and IV3. Consistent with the value of the Kolmogorov-Smirnov test, as well as the activity estimates in Table 2, the empirical log-density for the adjusted intraday S&P 500 returns approximates the log-density of the standard normal closely. In contrast, for IV3, the empirical log-density exceeds that of the standard normal significantly in the center of the distribution as well as in the tails beyond -2 and 2, while in the intermediate ranges of $[-2, -0.5]$ and $[0.5, 2]$, the relative size of the two log-densities is reversed. This highlights the stark deviations from Gaussianity in the middle of the standardized distribution for the IV3 increments. Qualitatively similar displays (not provided) arise for the other series for which the Kolmogorov-Smirnov test signals an absence of a diffusive term.

Table 3: Kolmogorov-Smirnov Tests for Local Gaussianity

Series	Range of Test				Sample
	$Q_{0.01} - Q_{0.2}$	$Q_{0.2} - Q_{0.4}$	$Q_{0.6} - Q_{0.8}$	$Q_{0.8} - Q_{0.99}$	Size (in days)
IV1	6.28	10.59	10.59	6.23	844
IV2	4.90	8.57	9.30	5.29	955
IV3	3.87	6.62	7.12	3.86	1022
IV4	2.31	4.67	4.55	1.89	1048
IV5	1.64	1.44	1.86	1.60	1048
IV6	1.40	1.38	1.12	1.20	1048
IV7	2.50	4.23	4.76	2.50	419
S&P 500	1.14	0.42	0.53	1.47	1062
VIX	4.65	8.13	7.60	5.03	1112
VF	2.90	9.09	8.35	2.50	1112
Critical Value of Test					
$\alpha = 5\%$	2.52	1.99	1.99	2.52	
$\alpha = 1\%$	2.73	2.29	2.29	2.73	

Note: The Kolmogorov-Smirnov test is based on comparing $\widehat{F}_n(\tau)$ in (20) with $F(\tau)$ using the asymptotic distribution of $\widehat{F}_n(\tau) - F(\tau)$ derived in Todorov and Tauchen (2013). Q_p denotes the p-th quantile of the standard normal distribution and the range of the test indicates the range of values of τ over which the supremum of the difference $\widehat{F}_n(\tau) - F(\tau)$ is taken. The critical values are for $T = 4 \times 252$ days.

Can we rationalize the rather striking evidence in Figure 5? In fact, it turns out the findings are exactly what we would expect for a very closely related scenario. Todorov and Tauchen (2013) document that if the underlying process Z is either of pure-jump type or a jump-diffusion then, regardless of the value of β, $\widehat{F}_n(\tau)$ converges in probability to the cdf of the β-stable random variate with first absolute moment equal to $\sqrt{2/\pi}$, as long as we normalize the high-frequency increments by the bipower variation statistic, $\widehat{B}_j^n(i)$, rather than the truncated variation $\widehat{C}_j^n(i)$. However, in practice, the choice among these alternative techniques is immaterial, and normalization via the bipower statistic produces a plot very similar to Figure 5. Thus, on Figure 6, we compare the log-densities for the standard normal ($\beta = 2$) and the 1.7-stable distribution, but both scaled to

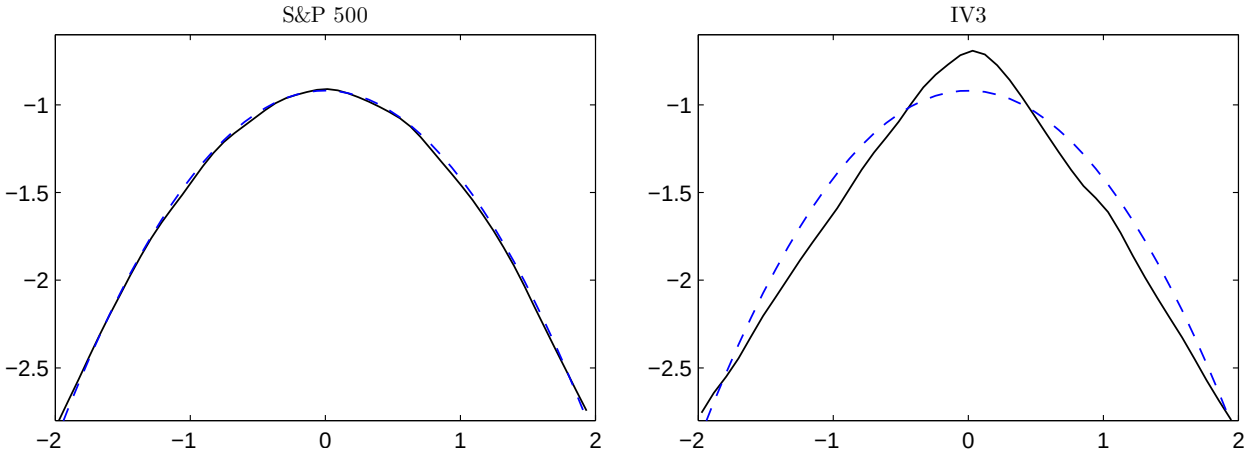

Figure 5: *Log-density of de-volatilized returns.* Solid lines on the plots are kernel-based (normal kernel) log-density estimates of the de-volatilized high-frequency increments $\frac{\sqrt{n}\widetilde{\Delta_i^n Z}}{\sqrt{\widehat{C}_j^n(i)}}$, for which $|\widetilde{\Delta_i^n Z}| \leq \alpha n^{-\varpi}$, and the dashed lines correspond to the log-density of the standard normal distribution. Left panel: S&P 500 Index, Right Panel: IV3.

have a first absolute moment of $\sqrt{2/\pi}$.[12] It is evident that Figure 6 portrays a qualitatively similar pattern to that observed for the log-density of the de-volatilized increments of IV3. Intuitively, since a stable distribution with index less than two is fat-tailed, it can only generate the same first absolute moment as the normal if the density around zero also exceeds that of the normal, inducing a typical leptokurtic shape on the density for the scaled non-Gaussian stable increments.

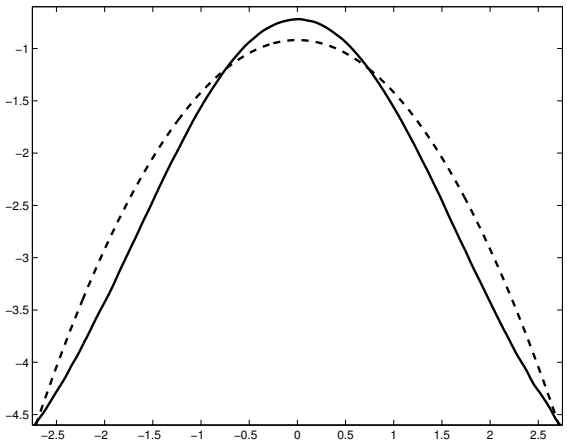

Figure 6: *Log-density of stable processes.* Solid line: log-density of 1.7-stable process with first absolute moment equal to $\sqrt{2/\pi}$; Dashed line: log-density of standard normal.

[12] $\beta = 1.7$ is roughly compatible with our point estimates for IV3 and IV4.

In summary, we find all our series, except for the S&P 500 futures and IV6 (and to some extent IV5), display properties that are consistent with them being of the pure jump type. We conclude this section with a few informal robustness checks regarding the finite sample behavior of the tests. In particular, the key idea behind our test based on the empirical cdf $\widehat{F}_n(\tau)$ is that normalization by the nonparametric volatility estimators $\widehat{C}_j^n(i)$ or $\widehat{B}_j^n(i)$ annihilate the effect of time-varying volatility which, otherwise, provides an alternative explanation for the presence of fat-tails in the (unconditional) return distribution. In essence, the test based on $\widehat{F}_n(\tau)$ seeks to separate this source of fat-tailedness from the impact of jumps which are the focus in the present work. Of course, volatility is known to be a highly persistent process. Hence, we can gauge whether we have been successful in "removing" the volatility from the high-frequency returns by checking for significant autocorrelation in the squared de-volatilized high-frequency increments, $\left(\frac{\sqrt{n}\widetilde{\Delta_i^n Z}}{\sqrt{\widehat{C}_j^n(i)}}\right)^2$. Figure 7 plots the serial correlation in the squared raw and de-volatilized high-frequency increments. It is evident that our scaling by the nonparametric volatility estimates $\widehat{C}_j^n(i)$ effectively annihilates the pronounced persistence in the second-moment of the raw returns. To further illustrate this, Figure 8 depicts the time series of raw and de-volatilized high-frequency increments. The figure indicates that the latter, unlike the former, appear consistent with the realization of an i.i.d. series, further corroborating the conclusion that the finite-sample performance of the Kolmogorov-Smirnov test based on $\widehat{F}_n(\tau)$ is satisfactory in the current context.

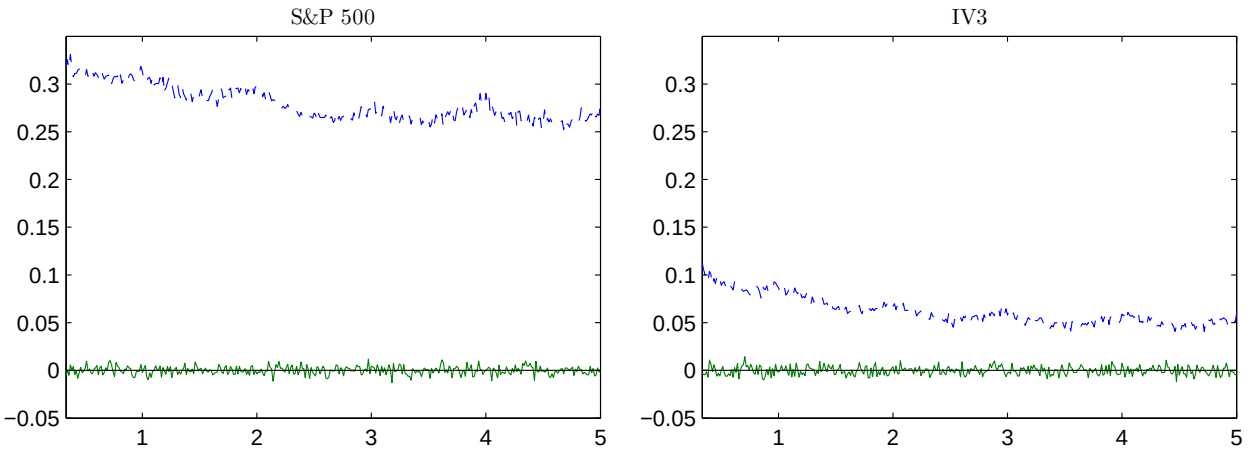

Figure 7: *Serial Correlation of Squared Increments.* Solid lines on the plot correspond to the de-volatilized high-frequency increments $\frac{\sqrt{n}\widetilde{\Delta_i^n Z}}{\sqrt{\widehat{C}_j^n(i)}}$, for which $|\widetilde{\Delta_i^n Z}| \leq \alpha n^{-\varpi}$, and the dashed lines to the raw high-frequency increments $\widetilde{\Delta_i^n Z}$.

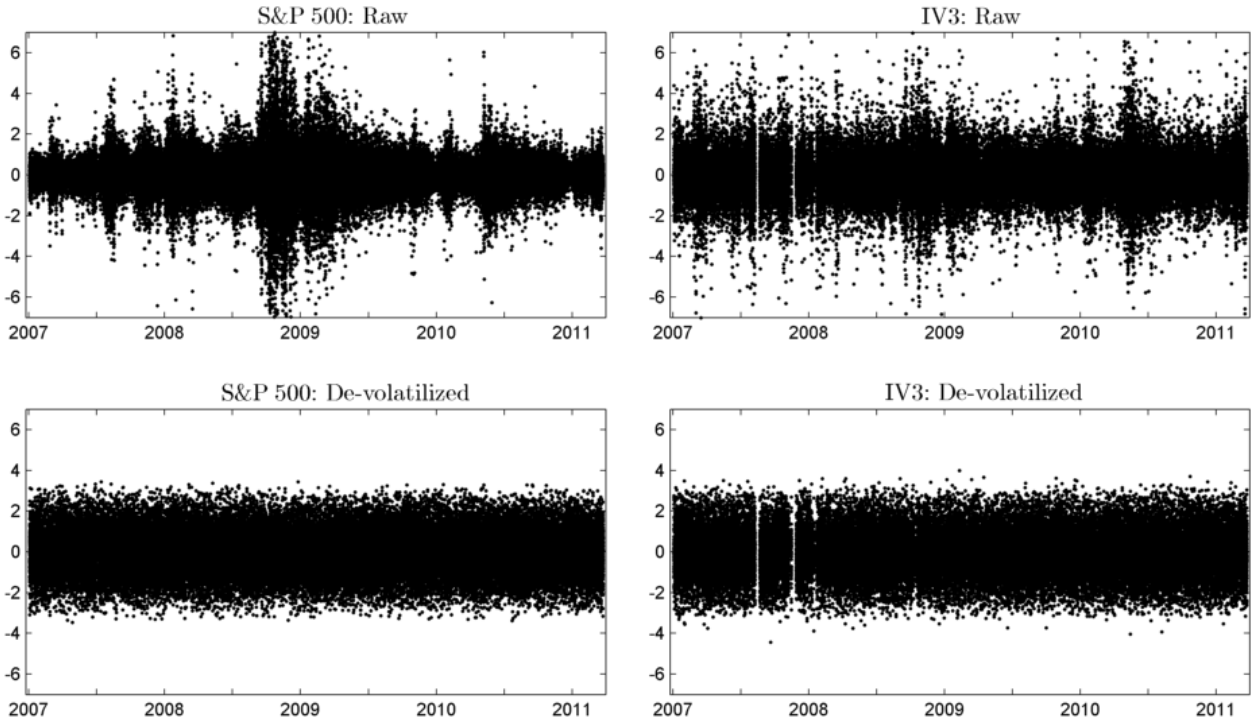

Figure 8: *Time Series of High Frequency Increments.* Top Panels: raw high-frequency increments $\widetilde{\Delta_i^n Z}$, standardized to have unit sample variance, Bottom Panels: the de-volatilized high-frequency increments $\frac{\sqrt{n}\widetilde{\Delta_i^n Z}}{\sqrt{\widehat{C}_j^n(i)}}$, for which $|\widetilde{\Delta_i^n Z}| \leq \alpha n^{-\varpi}$. Left panels: S&P 500 Index, Right Panels: IV3.

5.3 Interpreting the Results

As discussed in Section 5.2, if the stochastic volatility V_t is Markov relative to its natural filtration and the jump intensity is governed solely by volatility, then all our volatility-based measures must be of the same type as V_t, i.e., a jump-diffusion or of pure-jump type. Moreover, the jump activity should coincide for all the measures. However, the evidence in Tables 2 and 3 is at odds with this prediction. While most of the measures appear to be of pure jump type, the IV6 series along with the S&P 500 futures is best described via a jump-diffusive model. Consequently, we need a multivariate (latent) state vector for the underlying stock price X to potentially explain the empirical findings. Specifically, we may have different state variables driving different components of the volatility-based derivative measures. This is feasible if the implied volatility measures differ in their exposures to the jump and volatility risks in the underlying stock price process, X. For example, short-maturity deep out-of-the-money put options load mostly on the negative jump intensity and have hardly any exposure to the diffusive spot volatility, and analogous results hold for short-maturity deep out-of-the-money call options. At the same time, the Black-Scholes implied

volatility of the close-to-maturity at-the-money options are largely determined by the current (spot) value of the stochastic volatility. Thus, the deep out-of-the-money implied volatility measures IV1 and IV7 load predominantly on, respectively, the left and right jump intensity factors, while IV5 and IV6 load heavily on spot volatility. For the remaining volatility measures, we expect a relative loading on the volatility and jump intensity factors between that experienced by the deep out-of-the-money and the at-the-money implied option measures. As such, their fine structure, as characterized through the presence or absence of a diffusive term, will be determined by the empirically dominant high-frequency component.

For illustration, consider the simple case where the observed process is of the form $\alpha W_t + (1 - \alpha)S_t$, where W_t is a Brownian motion and S_t is a β-stable process with $\beta < 2$ and independent from W_t, so that $\alpha \in [0, 1]$ captures the relative weight of the two processes. In this scenario, for $\alpha = 0$, the observed process is of the pure jump type and its activity index coincides with the parameter of the stable process β. In contrast, as soon as $\alpha > 0$, the asymptotic small scale behavior of the process is determined by W_t, see, e.g., equation (11). However, this is an asymptotic result. In the presence of microstructure noise and sampling errors, the size of α will have a significant impact on the finite sample behavior of our estimators and test statistics. We illustrate this through Figure 9, which plots the mean of $\widehat{\beta}(0.7)$, obtained via a Monte Carlo simulation, for different values of α. As expected, if the Brownian motion receives considerable weight, the mean of $\widehat{\beta}(0.7)$ is close to the asymptotic limit of 2. On the other hand, for low values of α, the mean of $\widehat{\beta}(0.7)$ is quite close to β, rather than the asymptotic limit of 2, indicating that the finite sample behavior of the high-frequency increments is dominated by the β-stable process.

Clearly, the same type of factors will be operative when the diffusive and jump components display time-varying volatility and intensity, respectively. Specifically, the implied volatility measures are nonlinear functions of the state variables (factors) driving the volatility and jump intensity so the weight of the different factors in a locally linear expansion will fluctuate over time, depending on the realizations of the factors. Thus, if, e.g., the negative jump intensity factors are of pure jump type and the diffusive volatility evolves diffusively then, depending on the loadings of the implied volatility measures on the jump intensity and diffusive volatility factors, we should observe the estimates for the activity index fall between β and two. This is consistent with our findings in Table 2, where we noted the gradual increase in the estimates as we move across the moneyness spectrum from IV1 to IV6. A similar line of reasoning applies to the VIX index and the VIX futures. Since the VIX index may be viewed as a weighted average over the Black-Scholes implied volatilities for the out-of-the-money options, it too can be approximated by a weighted sum of the

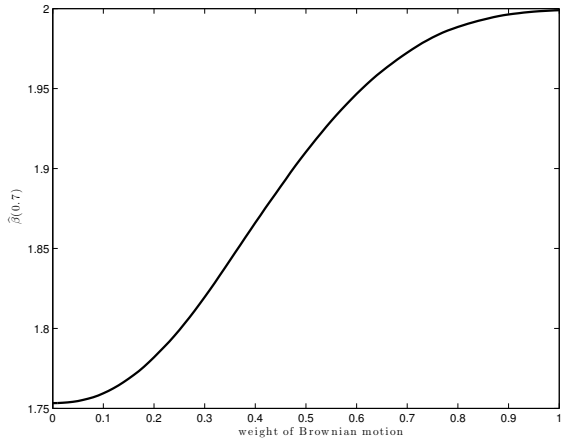

Figure 9: *Small sample behavior of* $\widehat{\beta}(0.7)$. The estimator is based on 22 days of data each of which contains 100 high-frequency returns simulated from the model $\alpha W_t + (1-\alpha)S_t$ for S_t being 1.75-stable with scale normalized such that $E|S_t| = \sqrt{\frac{2}{\pi}}t$. The figure plots the mean of $\widehat{\beta}(0.7)$ over 1000 Monte Carlo replications.

diffusive volatility and jump intensity factors.

The discussion also underscores the usefulness of including a diverse set of S&P 500 implied volatility measures in the analysis. The VIX index data, by construction, only speak to a certain weighted sum of the state variables driving the jump intensity and diffusive volatility. In contrast, by forming option portfolios that differ systematically in their degree of moneyness, we create a significant discrepancy in the portfolio loadings on the underlying factors which, in turn, should generate a sizeable spread in the estimates for the associated activity measures if, indeed, the underlying factors also differ substantially in this regard.

6 Conclusion

In this paper we conduct econometric analysis of the behavior over small time scales of various derivatives written on the S&P 500 index and the CBOE VIX volatility index. In particular, in our empirical analysis we study high-frequency intraday data on a wide range of options written on the S&P 500 index, the VIX index as well as a futures written on the latter. We conduct formal tests for deciding about the presence of a diffusive component in the dynamics of the series and further estimate the degree of jump activity in the absence of a diffusion. We analyze how the results regarding the different derivatives are theoretically connected within general no-arbitrage models for the underlying S&P 500 index. Our joint evidence suggests that the diffusive volatility and the (risk-neutral) intensity of the negative jumps in the S&P 500 index display different behavior

over small time scales: the increments of the former look locally Gaussian unlike those of the latter which are best described as locally stable with an activity index strictly less than two.

Appendix

A E-mini S&P 500 Futures and Options

The E-mini S&P 500 futures and options trade almost continuously for five days a week. Specifically, from Monday to Thursday, the trading is from 3:30 pm to 3:15 pm of the following day, with a one-hour maintenance shutdown from 4:30 pm to 5:30 pm. On Sunday, trading is from 5:00 pm to 3:15 pm of the following day. We exploit the so-called best bid-offer (BBO) files, which, among other variables, provide the best bid, bid depth, ask, ask depth, last trade price, and last trade volume. Quotes and trades are time stamped to the second. However, the files also contain a sequence indicator that identifies the order in which quotes and trades arrive to the exchange. Thus, we know the exact sequence of orders within each second. The BBO files record every change in the best bid or offer. Importantly, the quote updates arrive in pairs, one for the bid and one for the ask, synchronized by the sequence variable.

The ES futures contract expires quarterly, on the March expiration cycle. The notional value of one contract is \$50 times the value of the S&P 500 stock index.[13] The ES futures contract has a tick size of 0.25 index points, or \$12.50. On any trading day, the ES options are available for seven maturity months: four months from the March quarterly cycle and three additional nearby months ("serial" options). The options expire on a third Friday of a contract month. However, the quarterly options expire at the market open, while the serial options expire at the market close. The option contract size is one S&P 500 E-mini futures. The minimum price movement is 0.05 index points. The strikes are multiples of 5 for near-term months and multiples of 25 for far months. If at any time the S&P 500 futures contract trades through the highest or lowest strike available, additional strikes are usually introduced. A disadvantage of the ES options is their American-style feature. However, we conduct our empirical analysis in such a way that the effect of the early exercise is minimal.

To guard against data errors, we apply the so-called "non-convexity" filter. The filter imposes a maximum threshold for the degree of no-arbitrage violations implied by the option mid-quotes. Apparent arbitrage violations could arise from staleness of quotes due to a temporary malfunction of the dissemination system and other issues. The procedures for implementing this filter is detailed in Appendix A.2. Whenever the "non-convexity" threshold is exceeded, we deem a cross-section too noisy and unreliable and do not use it in our analysis.

A.1 Constructing Implied Volatility Measures

There are 1620 15-second intervals per trading day. For each interval, we construct seven implied volatility measures IV1-IV7. These measures correspond to a fixed maturity of 30 calendar days and are constructed as follows.

1. In the dataset, we match all puts and calls by trading date t, time to maturity τ, and strike K. We take the mid-quote as the observed option prices and convert them into forward prices, where we approximate the risk-free rate r_f over $[t, t+\tau]$ by the rate of T-bills. The implied forward price $F_{t,t+\tau}$ is obtained via the put-call parity relationship for the at-the-money put and call options.

2. Because the ES options are American type, their prices $P_t^A(K,\tau)$ and $C_t^A(K,\tau)$ could be slightly

[13] The notional value of the original ("big") S&P 500 futures contract is \$250 times the values of the index.

higher than prices of the corresponding European options $P_t(K,\tau)$ and $C_t(K,\tau)$. The difference, however, is very small for short maturities that we focus on. This is particularly true for OTM an ATM options.[14]

To infer prices of European options $P_t(K,\tau)$ and $C_t(K,\tau)$ we proceed as follows. First, we discard all ITM options. Prices of OTM and ATM options are more reliable, have smaller bid-ask spreads, and are less affected by the early exercise feature. Second, we correct American option prices $P_t^A(K,\tau)$ and $C_t^A(K,\tau)$ for the value of the early exercise feature by using Barone-Adesi and Whaley (1987) approximation.[15] Third, we compute prices of ITM options through the put-call parity relationship.

3. From generated prices of European options, we compute Black-Scholes implied volatilities $IV(K,\tau)$ and linearly interpolate them to obtain implied volatilities for seven values of the normalized moneyeness $m = -4, -3, -2, -1, 0, 0.5, 1$, defined in (6).

4. We repeat the above steps for the two option maturities with τ_j, $j = 1, 2$ closest to 30 days, but excluding options with less than seven calendar days to expire. Finally, given two available maturities, τ_j, the seven constant moneyness implied volatilities IV1-IV7 are obtained computed for a fixed maturity of 30 calendar days via a linear interpolation.

A.2 Non-Convexity Filter

To preclude arbitrage opportunities, theoretical call and put prices must be monotonic and convex functions of the strike. In particular, the call prices must satisfy the following convexity restriction:

$$D_i \;=\; \frac{C(K_{i+1}) \,-\, C(K_i)}{K_{i+1} \,-\, K_i} \;-\; \frac{C(K_i) \,-\, C(K_{i-1})}{K_i \,-\, K_{i-1}} \;\geq\; 0\,,$$

and a similar restriction for the put prices. We obtain the option prices as the average of the bid and ask quotes and use the above restriction to identify "suspect" cross-section with apparent arbitrage violations, which could arise from recording errors, staleness, and other issues. Specifically, for each strike K_i, we compute the following measure of local non-convexity:

$$NC_i \;=\; -\,\min\{D_i, 0\}\,.$$

For low strikes ($K_i \leq F$), we compute NC_i from OTM puts and, for high strikes ($K_i > F$), we use OTM calls. We then average NC_i across all strikes to obtain the aggregate measure of non-convexity NC.

When $NC > 0.1$, we deem a cross-section unreliable and do not use it in our econometric analysis. In practice, we declare all IV series at that point in time to be NaN. For those cross-sections, the option prices indicate sizeable apparent arbitrage opportunities. Andersen et al. (2013) uses a similar non-convexity filter.

A.3 Pre-Averaging of the Option IV Increments

We group the 1620 15-second intervals from 8:30am-3:15pm into 162 2.5-minute bins (10 observations per bin). For each bin, we compute the average price, requiring that there are at least 6 (out of 10) non-NaN prices, otherwise the average price is marked as NaN.[16] From the 162 pre-averaged prices, we compute 81 non-overlapping 2.5-min returns, so that each non-overlapping 5-minute return block generates one pre-averaged return.

[14] As shown in Whaley (1986), the early exercise premium increases with the level of the risk-free rate, volatility, time to maturity, and degree to which an option is in-the-money.

[15] It is important to point out that this correction is always substantially smaller than typical bid-ask spreads. In particular, the correction generally does not exceed 0.2% of an option price.

[16] An IV observation can be labeled NaN due to a violation of the non-convexity filter, as explained above, or due to a lack of coverage for this strike range by the exchange. Since the moneyness varies with the at-the-money Black-Scholes implied volatility, an option can move from being covered to not being covered and vice versa on a given day, depending on the shifts in volatility as well as the existing coverage of the strike range by the exchange for that day. The lack of coverage is particularly problematic for the out-of-the-money call options.

Next, if all 81 returns are non-NaNs, we drop the first 3, corresponding to the first 15-minutes of trading, when most of the problems with the option quotes occur. If out of 81 returns, no more than 3 are NaNs, we drop the missing returns and possibly more from the start, so that we again are left with 78 returns. If there are more than 3 NaN returns, the entire day is deemed unusable and is dropped from our sample.

B The VIX Index and VIX Futures

The contract multiplier for the VIX futures contract is \$1000. The contract has a tick size of 0.05 points, or \$50.00. CBOE usually lists up to nine near-term serial months and five months on the February quarterly cycle. Normally, the settlement date is the Wednesday that is thirty days prior to the third Friday of the calendar month immediately following the month in which the contract expires. The VIX futures are cash settled. The settlement value is for the VIX futures is determined via Special Opening Quotation (SOQ), which calculates VIX from the sequence of opening prices of the options on the settlement date. The opening price for any series in which there is no trade is equal to the average of that option's bid price and ask price as determined at the opening of trading. The final settlement value is rounded to the nearest \$0.01.

References

Ait-Sahalia, Y. and J. Jacod (2010). Is Brownian Motion Necessary to Model High Frequency Data? *Annals of Statistics 38*, 3093–3128.

Andersen, T. G., O. Bondarenko, and M. Gonzalez-Perez (2013). Uncovering Novel Features of Equity-Index Return Dynamics via Corridor Implied Volatility. Working paper, Northwstern University and University of Illinois.

Andersen, T. G., N. Fusari, and V. Todorov (2013a). Parametric Inference and Dynamic State Recovery from Option Panels. Working paper, Northwstern University.

Andersen, T. G., N. Fusari, and V. Todorov (2013b). The Risk Premia Embedded in Index Options. Working paper, Northwstern University.

Barndorff-Nielsen, O. and N. Shephard (2006). Econometrics of Testing for Jumps in Financial Economics using Bipower Variation. *Journal of Financial Econometrics 4*, 1–30.

Bollerslev, T. and V. Todorov (2011). Tails, Fears and Risk Premia. *Journal of Finance 66*, 2165–2211.

Carr, P., H. Geman, D. Madan, and M. Yor (2003). Stochastic Volatility for Lévy Processes. *Mathematical Finance 13*, 345–382.

Duffie, D., J. Pan, and K. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica 68*, 1343–1376.

Jacod, J., Y. Li, P. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure Noise in the Continuous Case: The Pre-averaging Approach. *Stochastic Processes and their Applications 119*, 2249–2276.

Jacod, J., M. Podolskij, and M. Vetter (2010). Limit Theorems for Moving Averages of Discretized Processes plus Noise. *Annals of Statistics 38*, 1478–1545.

Mancini, C. (2009). Non-parametric Threshold Estimation for Models with Stochastic Diffusion Coefficient and Jumps. *Scandinavian Journal of Statistics 36*, 270–296.

Podolskij, M. and M. Vetter (2009). Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps. *Bernoulli 15*, 634–658.

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge, UK: Cambridge University Press.

Todorov, V. (2013). Power Variation from Second Order Differences for Pure Jump Semimartingales. *Stochastic Processes and their Applications 123*, 2829–2850.

Todorov, V. and G. Tauchen (2010). Activity Signature Functions for High-Frequency Data Analysis. *Journal of Econometrics 154*, 125–138.

Todorov, V. and G. Tauchen (2011a). Limit Theorems for Power Variations of Pure-Jump Processes with Application to Activity Estimation. *Annals of Applied Probability 21*, 546–588.

Todorov, V. and G. Tauchen (2011b). Volatility Jumps. *Journal of Business and Economic Statistics 29*, 356–371.

Todorov, V. and G. Tauchen (2013). Limit Theorems for the Empirical Distribution Function of Scaled Increments of Ito Semimartingales at High Frequencies. *Annals of Applied Probability, forthcoming*.

Research Papers 2013

2013-35: Bent Jesper Christensen, Robinson Kruse and Philipp Sibbertsen: A unified framework for testing in the linear regression model under unknown order of fractional integration

2013-36: Niels S. Hansen and Asger Lunde: Analyzing Oil Futures with a Dynamic Nelson-Siegel Model

2013-37: Charlotte Christiansen: Classifying Returns as Extreme: European Stock and Bond Markets

2013-38: Christian Bender, Mikko S. Pakkanen and Hasanjan Sayit: Sticky continuous processes have consistent price systems

2013-39: Juan Carlos Parra-Alvarez: A comparison of numerical methods for the solution of continuous-time DSGE models

2013-40: Daniel Ventosa-Santaulària and Carlos Vladimir Rodríguez-Caballero: Polynomial Regressions and Nonsense Inference

2013-41: Diego Amaya, Peter Christoffersen, Kris Jacobs and Aurelio Vasquez: Does Realized Skewness Predict the Cross-Section of Equity Returns?

2013-42: Torben G. Andersen and Oleg Bondarenko: Reflecting on the VPN Dispute

2013-43: Torben G. Andersen and Oleg Bondarenko: Assessing Measures of Order Flow Toxicity via Perfect Trade Classification

2013-44: Federico Carlini and Paolo Santucci de Magistris: On the identification of fractionally cointegrated VAR models with the F(d) condition

2013-45: Peter Christoffersen, Du Du and Redouane Elkamhi: Rare Disasters and Credit Market Puzzles

2013-46: Peter Christoffersen, Kris Jacobs, Xisong Jin and Hugues Langlois: Dynamic Diversification in Corporate Credit

2013-47: Peter Christoffersen, Mathieu Fournier and Kris Jacobs: The Factor Structure in Equity Options

2013-48: Peter Christoffersen, Ruslan Goyenko, Kris Jacobs and Mehdi Karoui: Illiquidity Premia in the Equity Options Market

2013-49: Peter Christoffersen, Vihang R. Errunza, Kris Jacobs and Xisong Jin: Correlation Dynamics and International Diversification Benefits

2013-50: Georgios Effraimidis and Christian M. Dahl: Nonparametric Estimation of Cumulative Incidence Functions for Competing Risks Data with Missing Cause of Failure

2013-51: Mehmet Caner and Anders Bredahl Kock: Oracle Inequalities for Convex Loss Functions with Non-Linear Targets

2013-52: Torben G. Andersen, Oleg Bondarenko, Viktor Todorov and George Tauchen: The Fine Structure of Equity-Index Option Dynamics