Testing for Jumps: A Delta-Hedging Perspective*

Jia Li
Princeton University
Department of Economics and Bendheim Center for Finance
Princeton, NJ, 08544
This Version: June 6, 2011

Abstract

We measure asset price jumps by the hedging error they induce on a delta-hedged position of European options. Based on high frequency data, we propose a nonparametric estimator for this measure and a test for its positivity. We further construct a Kolmogorov-type test for the presence of jump hedging errors for a possibly infinite-dimensional family of options based on the worst-case contract in this family. Under regularity conditions, these tests are equivalent to tests for jumps. An empirical application on U.S. stocks in 2008 shows that jumps cause statistically significant and economically sizable hedging errors for short-dated call and put options.

JEL classification: C22
Keywords: Hedging, high frequency data, semimartingale, testing for jumps.

1 Introduction

The recent availability of observations on financial returns at increasingly higher frequencies has prompted the development of methodologies designed to test the specification of suitable models for these data. One of the first issues studied in the literature is testing the presence of jumps. Since the seminal work of Barndorff-Nielsen and Shephard (2004), there have been several tests in the literature. However, these tests are typically based on purely statistical measures of jumps and thus lack a clear economic interpretation. Even though the literature has documented that jumps are present in financial data, it is not clear how the economic relevance of jumps is reflected in the

---

*I am grateful to Yacine Aït-Sahalia, Jean Jacod, Ulrich Müller and Mark Watson for their guidance and support. I also thank Paolo Colla, Thomas Eisenbach, Jakub Jurek, Weicheng Lian, Jose Scheinkman, Martin Schmalz, Hyun Shin, Wei Xiong and Dacheng Xiu for helpful discussions. All errors are mine. Email: jiali@princeton.edu. Webpage: http://www.princeton.edu/~jiali.
existing testing procedures. The present paper intends to fill this gap. We measure jumps by the hedging error they induce for European options and design tests for jumps based on this measure.

To illustrate the idea, let us consider the classical problem of Black and Scholes (1973). That is, hedging a short position of a European call option with a long position of the underlying stock, where the hedge ratio (i.e. the number of shares of stock required for hedging one option contract) is the option delta of the Black-Scholes formula. If the stock price follows a geometric Brownian motion, it is well known that the value of the delta-hedging portfolio perfectly replicates the model price of the option. What happens if the stock price jumps? At the jump time, both the model price of the option, which is a nonlinear function of the stock price, and the value of the hedging portfolio, which is a linear function of the stock price, jump. Because the option and its hedge respond to the stock price jump differently, the jump necessarily induces a discrepancy between their values, and thus leads to a hedging error, henceforth a “jump error”. We note that the jump error is not specific to the call option or the Black-Scholes formula. Instead, it is a generic consequence of the nonlinearity of the payoff of a derivative security, coupled with the discrete nature of jumps. Qualitatively, it is clear that jumps cause hedging errors. The remaining econometric questions are: how large is the jump error; is it statistically significant?

First of all, we cannot answer this question by simply looking at the actual hedging error in a trading book. The reason is that the observed total hedging error not only includes the jump error, but also hedging errors coming from other sources such as discrete implementation, transaction costs, and volatility misspecification. The answer to this question would nevertheless be immediate if we could observe the continuous-time sample path of the stock price. In this case, we could observe the jump time, the pre-jump stock price and the jump size, which are sufficient to determine the hedging error contributed by each jump (see Section 2). However, such information is unavailable in reality because we can only observe data sampled at discrete times. The econometric challenge is to use such data to estimate the jump error of an option and then make statistical inference. We propose estimators and tests for such purposes. Our method is nonparametric in the sense that we only require the stock price to be an Itô semimartingale without parametrizing the stochastic volatility or the jump component in any way. Since the jump error is identified in the continuous-time limit, the method relies on the infill asymptotics with the sampling interval going to zero. High frequency data are necessary for implementation.

Our general econometric result is to jointly test the positivity of jump errors of a (possibly infinite dimensional) family of options written on the same asset, e.g. the cross section of call and put options with various strikes. This multivariate setting is motivated by our general interest in the whole market instead of a particular contract. To this end, we introduce the notion of jump error profile, which is defined as the collection of jump errors for a family of options. We propose a consistent estimator for the jump error profile. We also propose a Kolmogorov-type test
to examine whether the profile is zero. The idea behind this test is to look at the jump error of the worst-case contract in the family and determine its statistical significance. A failure of rejection would suggest that, uniformly across this family of contracts, jumps are unlikely to be important for hedging. In the effort of constructing this test, we derive an empirical process-type asymptotic theory for Itô semimartingale models. This result seems to be unconventional for the theory of empirical processes because our Itô semimartingale model is non-ergodic, non-stationary and does not impose weak dependence (independence, mixing, etc.).

We now discuss the related literature and some further results. Our test includes the bipower test of Barndorff-Nielsen and Shephard (2004) as a special case corresponding to a quadratic contract, or equivalently, an equally weighted portfolio of the complete cross section of call options. By embedding the bipower test in our framework, we shed new light on the economic interpretation of this test from a hedging perspective. Another inspiration to our work is the paper by Jiang and Oomen (2008). Jiang and Oomen measure jumps by the swap variance and propose a test based on this quantity. The statistical properties of their test are quite different from ours. The economic interpretability of their test seems to be narrower than what we pursue here since it is restricted to the log contract, which is a fictitious theoretical device and rarely, if ever, traded in the market.\(^1\) In contrast to these papers, we consider arbitrary European options in an infinite dimensional setting, allowing many possibilities for constructing new tests for jumps. Other tests for jumps based on high-frequency data have been proposed by Lee and Mykland (2008), Lee and Hannig (2010) (based on detecting large returns), and Huang and Tauchen (2005), Andersen, Bollerslev, and Diebold (2007), Aït-Sahalia and Jacod (2009), Podolskij and Ziggel (2010), Aït-Sahalia, Jacod, and Li (2011) (based on power variations of jumps). The tests of Podolskij and Ziggel (2010) and Aït-Sahalia, Jacod, and Li (2011) are robust to microstructure noise, which is not considered here. Robustifying the method against microstructure noise is crucial for applications, but it seems to deserve a separate paper.

Our paper is also related to the literature on hedging derivative securities. In the study of discrete hedging in diffusion models, Bertsimas, Kogan, and Lo (2000) introduce the notion of temporal granularity as a measure of the magnitude of the hedging error due to discretization. This result is further extended by Hayashi and Mykland (2005) to continuous Itô semimartingale models. It turns out that the t-statistic of our test associated with a single contract is proportional to the estimated ratio of the jump error to the ex-post temporal granularity. Hence, our test is essentially a comparison between the jump hedging error and the discretization hedging error. A rejection of the null hypothesis that the jump error is zero suggests that the jump error is sufficiently large relative to the discretization error.

The paper is organized as follows. We discuss delta-hedging and characterize the jump error

\(^1\)The statistical validity of Jiang and Oomen’s test does not depend on the log contract being traded.
in Section 2. In Section 3, we present our estimator and show its consistency. In Section 4, we present our test based on a single contract. The testing framework is extended to a functional setting in Section 5, in which we propose a new test based on the worst-case jump error. Section 6 shows simulation results. We present an empirical application in Section 7. Section 8 concludes. Proofs are collected in Section 9.

2 Delta-hedging and hedging errors

In this section, we characterize jump-induced delta-hedging errors of European options. To formalize ideas, let the stock price process $X_t$ be a semimartingale with the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t,$$

where $b_t$ is the instantaneous drift, $\sigma_t^2$ is the instantaneous variance (hence $v_t = \sigma_t/X_t$ is the instantaneous volatility), and $J_t$ is a purely discontinuous process. Assuming $X$ to be a semimartingale is necessary to exclude arbitrage opportunities (see Harrison and Pliska (1981), Delbaen and Schachermayer (1994)). We discuss regularity conditions on the price process in Section 3.

We suppose that an investor delta-hedge an option by dynamically trading the underlying stock. To implement this hedging strategy, it is necessary to specify an option pricing function, which we denote by $V(t, X_t)$. The hedge ratio is then given by the partial derivative $\frac{\partial V}{\partial X}(t, X_t)$, where $X_{t-} = \lim_{s \to t} X_s$ is the pre-jump price. The hedging problem is fully characterized by the pricing function $V(\cdot)$ as it specifies both the term of the option contract and the hedging strategy.

In this paper, we do not wish to enter a discussion about which pricing model should be used for hedging, as there is a large literature on this topic (see e.g. Singleton (2006)). We simply take $V(\cdot)$ as an input and consider the corresponding hedging error. We note that $V(\cdot)$ is not assumed to be correctly specified, if there ever exists a “correctly specified” pricing function.

We denote the size of a jump at time $t$ by $\Delta X_t = X_t - X_{t-}$. For simplicity, we suppose that the bond holding in the hedging portfolio earns a constant risk-free rate $r_f$. In a frictionless market, the hedging error of this strategy from time $t_0$ to time $t$, denoted by $\Pi_t$, is given by

$$\Pi_t = V(t, X_t) - V(t_0, X_{t_0}) - \int_{t_0}^t V_x(s, X_{s-}) dX_s - \int_{t_0}^t r_f(V(s, X_s) - V_x(s, X_s) X_s) ds.$$

To simplify notation, we normalize $t_0 = 0$ below.

---

2 In practice, an investor may also gamma- or vega-hedge an option with the underlying and the other option. In this case, we can transform the gamma- or vega-hedging problem into a delta-hedging problem for a portfolio of options.
By Itô’s formula, we can decompose the instantaneous hedging error as

\[
d\Pi_t = \varphi(t, X_{t-}, \Delta X_t) + \rho(t, X_t, \sigma_t) dt,
\]

where \(\varphi(t, X_{t-}, \Delta X_t) = V(t, X_{t-} + \Delta X_t) - V(t, X_{t-}) - V_x(t, X_{t-}) \Delta X_t,\)

\[\rho(t, X_t, \sigma_t) = V_t(t, X_t) + \frac{1}{2} V_{xx}(t, X_t) \sigma_t^2 - r_f(V(t, X_t) - V_x(t, X_t) X_t),\]

which should be understood in integral form as

\[
\Pi_t = \sum_{s \leq t : \Delta X_s \neq 0} \varphi(s, X_{s-}, \Delta X_s) + \int_0^t \rho(s, X_s, \sigma_s) ds.
\]

The sum in (4) involves countably many nonzero terms because there are countably many jumps and \(\varphi(s, X_{s-}, \Delta X_s) \neq 0\) only if \(\Delta X_s \neq 0\).

The first component on the right side of (4) is the hedging error induced by jumps, i.e. the jump error, and the second component is contributed by the stochastic volatility, henceforth the volatility error. Before elaborating these interpretations, we make a few remarks.

First, \(\Pi_t\) is the hedging error under continuous rebalancing in a frictionless market. In practice, discrete implementation and transaction costs lead to additional hedging errors. Such hedging errors should not be attributed to jumps. Therefore, they should be isolated from the definition of the jump error as considered here.

Second, the notation \(V(t, X_t)\) indicates that \(V(\cdot)\) only depends on time and the spot price. In general, a pricing function may also depend on parameters governing the risk-neutral law of \(X\) and/or unobservable state variables such as the stochastic volatility. These unobservable quantities need to be estimated in order to determine \(V(\cdot)\). By suppressing the dependence of \(V(\cdot)\) on these quantities in our notation, we have implicitly assumed that the estimates of unobservables are not updated during \([t_0, t]\). This assumption is not restrictive since we are mainly interested in the hedging error contributed by intraday jumps accumulated within a short period, say one day or one week.

Third, if we could observe the sample path in continuous-time, we would have complete knowledge about jumps and the volatility. Quantifying each component would be a straightforward calculation. In practice, however, the data are only sampled at discrete times. The main focus of our paper is on the estimation and inference for the jump error based on discretely sampled data.

Fourth, the jump error is defined through a delta-hedging strategy. Excluding other hedging strategies, such as the quadratic hedging (see Schweizer (2001)), is restrictive. Eliminating the delta risk is necessary in order to obtain a central limit theorem and conducting formal statistical tests.

We now interpret the jump error intuitively via the instantaneous representation (1). If a jump occurs at time \(t\) with size \(\Delta X_t\), the model price of the option jumps by \(V(t, X_{t-} + \Delta X_t) - V(t, X_{t-})\)
and the value of the delta-hedging portfolio jumps by $V_x(t, X_{t^-}) \Delta X_t$. The term $\varphi(t, X_{t^-}, \Delta X_t)$, defined as their difference, quantifies the hedging error induced by the jump at the instant $t$. Mathematically, $\varphi(\cdot)$ is the difference between $V(t, X_t)$ and its first-order Taylor approximation around $X_{t^-}$. We can thus think of $\varphi(\cdot)$ as a local measure of the nonlinearity of the pricing function.

We note that jumps induce hedging errors not because the pricing function is possibly misspecified, but rather because the value of the option is nonlinear in $X_t$. Indeed, even if the pricing function $V(\cdot)$ is correctly specified, $\varphi(t, X_{t^-}, \Delta X_t)$ is generically nonzero when $\Delta X_t \neq 0$; on the other hand, when $\Delta X_t = 0$, we have $\varphi(t, X_{t^-}, \Delta X_t) = 0$ regardless whether $V(\cdot)$ is misspecified or not.

To interpret the volatility error, we first consider the basic case in which the pricing function $V(\cdot)$ is derived from the Black-Scholes-Merton model with the local volatility function $(t, x) \mapsto v_{BS}(t, x)$. It is well known that $V(\cdot)$ should solve the Black-Scholes valuation equation:

$$V_t(t, x) + \frac{1}{2} V_{xx}(t, x) x^2 v_{BS}^2(t, x) - r f (V(t, x) - V_x(t, x) x) = 0,$$

subject to a suitable boundary condition determined by the terminal payoff. Combining (3) and (5), we have

$$\rho(t, X_t, \sigma_t) = \frac{1}{2} V_{xx}(t, X_t) \left( \sigma_t^2 - v_{BS}^2(t, X_t) X_t^2 \right).$$

In words, $\rho(\cdot)$ is the difference between the true instantaneous diffusive variance and its model counterpart, scaled by half of the option gamma (i.e. the second partial derivative of option price with respect to the stock price). In our econometric framework, we do not require the pricing function $V(\cdot)$ obey the Black-Scholes equation for any function $v_{BS}(\cdot)$. In this general case, $\rho(t, X_t, \sigma_t)$ measures the degree to which the pricing function violates the Black-Scholes equation at the instant $t$.

Our general econometric framework is based on a family of pricing functions $\mathbb{V} = \{V(\cdot, \cdot; \theta) : \theta \in \Theta\}$. The index $\theta$ may include the configuration of a contract, such as the strike price and the maturity of a call option. It may also include characteristics of the pricing model, such as the volatility parameter in the classical Black-Scholes formula or, more generally, any finite-dimensional parameter which governs the risk-neutral law of the stock price process. We set for each $\theta \in \Theta$,

$$B_T(\theta) = \sum_{s \leq T} \varphi(s, X_{s^-}, \Delta X_s; \theta).$$

We refer to the (random) function $\theta \mapsto B_T(\theta)$ as the jump error profile.
3 Estimation of hedging errors

3.1 Assumptions on the stock price process

We collect some regularity conditions on the stock price process. Our assumptions are weak enough to incorporate most continuous-time models studied in the empirical asset pricing literature. In particular, we allow for stochastic volatility and jumps and impose essentially no restriction on their dynamics. Since the mathematical presentation may appear somewhat technical, readers interested in applications may skip this subsection during the first reading.

We assume that the stock price \( X = (X_t)_{t \geq 0} \) is a one-dimensional process taking values in an open set \( D \subset \mathbb{R} \), and is sampled at regularly spaced discrete times \( i \Delta_n, i \geq 0 \), over a fixed time interval \([0, T]\), with a time lag \( \Delta_n \) which asymptotically goes to 0. The basic assumption is that \( X \) is an Itô semimartingale on a filtered space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})\), which means that it can be written as

\[
X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (\delta 1_{|\delta| \leq 1}) * (\mu - \nu)_t + (\delta 1_{|\delta| > 1}) * \mu_t,
\]

where \( W \) is a Brownian motion, \( \mu \) is a Poisson random measure on \( \mathbb{R}_+ \times E \) and its compensator is \( \nu(dt, dz) = dt \otimes \lambda(dz) \) where \((E, \mathcal{E})\) is an auxiliary space and \( \lambda \) is a \( \sigma \)-finite measure (all these are defined on the filtered space above and we refer for example to Jacod and Shiryaev (2003) for all unexplained terms).

In particular, when \( X \) is continuous, it has the form

\[
X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s. \tag{7}
\]

We further assume for some \( r \in [0, 2] \):

Assumption (H-\( r \)): 1) the process \( (b_t) \) is optional and locally bounded;

2) the processes \( (\sigma_t) \) is càdlàg and adapted;

3) the function \( \delta \) is predictable, and there is a bounded function \( \gamma \) in \( L^r(E, \mathcal{E}, \lambda) \) such that the process \( \sup_{z \in E} (|\delta(\omega, t, z)| \wedge 1)/\gamma(z) \) is locally bounded;

4) we have almost surely \( \int_0^t \sigma_s^2 ds > 0 \) for all \( t > 0 \). \hfill \Box

Assumption (K): We have Assumption (H-2) and \( \sigma_t \) is also an Itô semimartingale which can be written as

\[
\sigma_t = \sigma_0 + \int_0^t \bar{b}_s ds + \int_0^t \bar{\sigma}_s dW_s + M_t + \sum_{s \leq t} \Delta \sigma_s 1_{|\Delta \sigma_s| > \nu},
\]

where \( M \) is a local martingale orthogonal to \( W \) and with bounded jumps and \( \langle M, M \rangle_t = \int_0^t a_s ds \), and the compensator of \( \sum_{s \leq t} 1_{|\Delta \sigma_s| > \nu} \) is \( \int_0^t a'_s ds \), and where \( \bar{b}_t, a_t, a'_t \) are optional locally bounded processes, and \( \bar{\sigma}_t \) is optional and càdlàg, as well as \( b_t \). \hfill \Box
Assumption (H-r) imposes very mild measurability and sample-path regularity. The parameter \( r \) governs the concentration of small jumps. The larger \( r \), the weaker this assumption. When \( r = 2 \), we essentially put no restriction on the concentration of small jumps since jumps of any semimartingale are square-summable. When \( r = 1 \) (resp. \( r = 0 \)), the jump process has finite variation (resp. finite activity). Assumption (K) can be considered as a smoothness condition for \( (\sigma_t)_{t \geq 0} \) in a stochastic sense. However, this assumption does not require sample paths of the volatility process to be smooth. We allow the drift of volatility and the volatility of volatility to be stochastic in a nonparametric manner. We also allow jumps in volatility, which can have finite or infinite activity, or even infinite variation. Although it is fairly unrestrictive, Assumption (K) does exclude the process \( (\sigma_t) \) being a fractional Brownian motion.

### 3.2 Assumptions on the pricing function

We collect some regularity conditions on the family \( \mathbb{V} \) of pricing functions.

**Assumption (V):** Let \((\Theta, d)\) be a compact metric space in \( \mathbb{R}^q \), \( q \geq 1 \). For any bounded subset \( \mathcal{K} \subset \mathcal{D} \), there exists some finite constant \( C > 0 \), such that the family \( \mathbb{V} \) satisfies the following conditions.

1) For each \( \theta \in \Theta \), \( V(t, x; \theta) \) is differentiable in \( t \) and fourth order differentiable in \( x \). The partial derivatives are continuous in \((t, x)\).

2) For \( h \in \{V_{xx}, V_{xxx}, V_{xxxx}\} \), \( \sup_{(t,x,\theta) \in [0,T] \times \mathcal{K} \times \Theta} |h(t, x; \theta)| < \infty \).

3) For \( h \in \{V_{xx}, V_{xxx}\} \), \( \sup_{(t,x) \in [0,T] \times \mathcal{K}} |h(t, x; \theta') - h(t, x; \theta)| \leq C d(\theta', \theta), \theta, \theta' \in \Theta \).

4) For some \( \kappa > 0 \), \( N(\varepsilon, \Theta, d) \leq C \varepsilon^{-\kappa} \) for any \( \varepsilon \in (0, 1] \), where \( N(\varepsilon, \Theta, d) \) is the minimum number of \( \varepsilon \)-balls needed for covering \( \Theta \).

5) We have almost surely \( \int_0^T V_{xx}(s, X_s; \theta)^2 ds > 0 \) for all \( \theta \in \Theta \). \( \square \)

Assumptions (V1-3) are satisfied when the pricing function \( V(\cdot) \) is sufficiently smooth in its state variables and the parameter \( \theta \). In many cases, we can take the metric \( d(\cdot, \cdot) \) to be Euclidean and Assumption (V4) automatically holds for \( \kappa = q \). In general, Assumption (V4) restricts the complexity of the set \( \Theta \). Assumption (V5) requires that the pricing function has nonzero curvature over the horizon. We need this assumption to avoid degenerate central limit theorems.

A basic example which verifies Assumption (V) is the following. Let \( \mathbb{V} \) be the collection of Black-Scholes formulas of a family of European call options indexed by their strike prices. In this case, the index \( \theta \) is the strike price. More generally, \( \theta \) may include the volatility parameter in the Black-Scholes model and the maturity of the contract. In this general case, \( \mathbb{V} \) represents a cross section of call options hedged under various volatility parameters.
3.3 Consistent estimators of hedging errors

We propose an estimator for the jump error profile $B_T(\cdot)$. For any process $Y$, we denote its increment over the interval $((i - 1)\Delta_n, i\Delta_n]$, $i \geq 1$, by

$$\Delta^n_i Y = Y_i\Delta_n - Y_{(i-1)\Delta_n}.$$  

(The notation $\Delta^n_i Y$ is not meant to denote the $n$th difference, but rather highlights that $\Delta^n_i Y$ forms a triangular array.) The estimator for the jump error $B_T(\theta)$ is given by

$$B_T^n(\theta) = A^n_T(\theta) - \tilde{A}_T^n(\theta),$$

where

$$A_T^n(\theta) = \sum_{i=1}^{[T/\Delta_n]} \varphi((i-1)\Delta_n, X_{(i-1)\Delta_n}; \Delta^n_i X; \theta)$$

$$\tilde{A}_T^n(\theta) = \frac{1}{2m_2} \sum_{i=1}^{[T/\Delta_n]-1} V^n_{xx,i-1}(\theta) |\Delta^n_i X| |\Delta^n_{i+1} X|$$

$$V^n_{xx,i}(\theta) = V_{xx}(i\Delta_n, X_{i\Delta_n}; \theta), \ i \geq 0,$$

and for $w > 0$, $m_w$ is the $w$th absolute moment of a standard normal variable, that is, $m_w = \mathbb{E}[|U|^w], U \sim N(0, 1)$. Following the convention in the literature, we refer to $B_T^n(\theta)$ as the realized jump error and the functional estimator $B_T^n(\cdot)$ as the realized jump error profile.

**Theorem 1** Suppose that Assumptions (H-2) and (V) hold. We have uniformly in $\theta \in \Theta$,

$$A_T^n(\theta) \overset{p}{\rightarrow} \frac{1}{2} \int_0^T V_{xx}(s, X_s; \theta) \sigma_s^2 ds + B_T(\theta), \quad (8)$$

$$\tilde{A}_T^n(\theta) \overset{p}{\rightarrow} \frac{1}{2} \int_0^T V_{xx}(s, X_s; \theta) \sigma_s^2 ds, \quad (9)$$

and consequently $B_T^n(\theta) \overset{p}{\rightarrow} B_T(\theta)$.

Theorem 1 shows the consistency of the realized jump error and the uniform consistency of the realized jump error profile. The theorem also clarifies the role of $A_T^n(\cdot)$ and $\tilde{A}_T^n(\cdot)$ in our construction. The estimator $A_T^n(\cdot)$ is an intuitive sample analog to $B_T(\cdot)$. However, it carries a bias $\frac{1}{2} \int_0^T V_{xx}(s, X_s; \cdot) \sigma_s'^2 ds$ in the first-order asymptotics as shown in (8). This bias is corrected by $\tilde{A}_T^n(\cdot)$. The correction term $\tilde{A}_T^n(\cdot)$ generalizes the bipower estimator originally proposed by Barndorff-Nielsen and Shephard (2004) with random weights.

We also have a consistent estimator of the volatility error as a by-product of Theorem 1. For brevity, we only present the result in the case when $V(\cdot)$ satisfies the Black-Scholes equation. We need this result in our empirical application in which we use the volatility error as a yardstick to gauge the relative magnitude of the jump error.
Corollary 1 Suppose that the same conditions as in Theorem 1 hold. If the function \( V(\cdot, \cdot) \) satisfies the Black-Scholes equation for some local volatility function \((t, x) \mapsto v_{BS}(t, x)\), then the realized volatility error defined by

\[
D^n_T = A^n_T - \frac{\Delta n}{2} \sum_{i=1}^{[T/\Delta n]} V^n_{xx,i-1} v_{BS}^2 ((i - 1) \Delta n, X_{(i-1)\Delta n}) X^2_{(i-1)\Delta n}.
\]

converges in probability to the volatility error

\[
D_T \equiv \int_0^T \rho(s, X_s, \sigma_s) \, ds = \frac{1}{2} \int_0^T V_{xx}(s, X_s) \left( \sigma^2_s - v_{BS}^2(s, X_s) X^2_s \right) \, ds.
\]

3.4 A sensitivity measure of the jump error

It is helpful to understand how sensitive the jump error \( B_T(\theta) \) is with respect to changes in \( \theta \). When \( \theta \) is one- or two-dimensional, we can visualize the dependence of \( B_T(\theta) \) on \( \theta \) by plotting the functional estimator \( B^n_T(\theta) \) versus \( \theta \). However, such presentation is infeasible when \( \theta \) is of high dimension. Even in the low dimensional case, it is still useful to characterize the sensitivity with a few numbers in addition to presenting a curve or a surface. For this purpose, we propose a local sensitivity measure which is simply defined as the gradient of \( B_T(\theta) \) with respect to \( \theta \), i.e., \( \partial B_T(\theta) / \partial \theta \). This quantity is unobservable but can be consistently estimated.

Corollary 2 Suppose that Assumption (H-2) holds and the family of functions \( \{\partial V(\cdot, \cdot, \theta) / \partial \theta : \theta \in \Theta\} \) satisfies Assumption (V). The variables \( \partial B^n_T(\theta) / \partial \theta \) converges in probability uniformly in \( \theta \in \Theta \) to \( \partial B_T(\theta) / \partial \theta \).

4 Univariate tests

We construct tests in order to determine the statistical significance of jump errors. In this section, we consider a univariate problem by fixing one option contract which is delta-hedged according to a fixed pricing function \((t, x) \mapsto V(t, x)\), so the family \( V \) in Assumption (V) is a singleton. We thus suppress the index \( \theta \) for notational simplicity. We present this special case separately to help develop intuition and relate our approach to the literature.

4.1 Equivalence to tests for jumps

The testing problem can be formally stated as follows: for a given realization \( \omega \in \Omega \),

one-sided test \( \left\{ \begin{array}{ll}
H_0 : & \omega \notin \Omega^b_T, \\
H_a : & \omega \in \Omega^b_T,
\end{array} \right. \) \( \Omega^b_T = \{B_T > 0\}, \)
two-sided test \[ \begin{cases} H_0 : & \omega \notin \Omega^b_T, \\ H_a : & \omega \in \Omega^b_T, \end{cases} \quad \Omega^b_T = \{ B_T \neq 0 \}. \]

The set \( \Omega^{b+}_T \) is the event in which the jump error is positive, representing an economic loss for the option seller. The one-sided test is designed to detect the down-side risk. On the other hand, the sample path of the stock price falls in the event \( \Omega^b_T \) as long as the jump error is nonzero, so the jump error may be either a loss or a gain in economic terms. The risk captured by \( \Omega^b_T \) is two-sided.

We now show that under regularity conditions, testing for the presence of jump errors is equivalent to testing for jumps. The latter can be formally stated as

\[ \begin{cases} H_0 : & \omega \notin \Omega^j_T, \\ H_a : & \omega \in \Omega^j_T, \end{cases} \]

where \( \Omega^j_T = \{ \omega : t \mapsto X_t(\omega) \text{ is discontinuous on } [0, T] \} \). It is clear that \( \Omega^{b+}_T \subseteq \Omega^{b}_T \subseteq \Omega^j_T \). The reverse inclusion is given below.

**Theorem 2** Suppose that for each \( t \in [0, T] \), the function \( x \mapsto V(t, x) \) is twice continuously differentiable on \( D \).

(a) If \( V_{xx}(\cdot, \cdot) > 0 \) on \( [0, T] \times D \), then \( \Omega^{b+}_T = \Omega^{b}_T = \Omega^j_T \).

(b) Assume that (i) the jumps of \( X \) have finite activity with successive arrival times denoted by \( T_q, q \in \mathbb{N} \); (ii) the conditional distribution of \( \Delta X_{T_q} \) given \( \mathcal{F}_{T_q-} \) is absolutely continuous for each \( q \in \mathbb{N} \); and (iii) for each \( t \), the function \( x \mapsto V(t, x) \) is non-affine on any open interval in \( D \). Then \( \Omega^b_T = \Omega^j_T \) almost surely.

Theorem 2(a) shows that if the pricing function has positive gamma, then jumps necessarily induce a positive jump error, and thus a loss to the hedger who shorts the option. As the jump error can never be negative in this case, it is natural to conduct the one-sided test. Basic examples of this sort include the classical Black-Scholes model (Black and Scholes (1973)) and Merton’s jump-diffusion model (Merton (1976)) for vanilla call and put options.

Without the single-signed gamma, it is possible that the loss induced by one jump is cancelled by the gain from the other. As a result, the presence of jumps does not necessarily imply nonzero jump error. However, Theorem 2(b) shows that such a possibility is negligible. The conditions imposed here are unlikely to be the weakest, but seem to be general enough for most applications. Condition (i) only requires that there are finitely many jumps over a finite time period. This condition ensures that jumps can be ordered in time so that condition (ii) is a valid statement. Condition (ii) is satisfied if the distribution of the jump size, conditional on the pre-jump information, has a density. Condition (iii) is generically true for derivatives with nonlinear payoffs.
The significance of Theorem 2 is that under the null hypotheses, there are no jumps. For testing purposes, it is enough to characterize the sampling variability of the realized jump errors under the null, as we discuss below.

4.2 Second order asymptotic properties

In this subsection, we describe the convergence in law of the realized jump error when the underlying process is continuous. For a sequence of random variables $\xi_n$, we write $\xi_n \xrightarrow{L} (s) MN (0, \Sigma)$ if $\xi_n$ converges stably in law to a mixture normal distribution $MN (0, \Sigma)$ which, conditional on the information set ($\sigma$-algebra) $F$, is centered Gaussian with variance $\Sigma$. Stable convergence is a slightly stronger notion than the usual weak convergence. Its key property which is useful for us is that if there is a sequence of estimators $\hat{\xi}_n$ which consistently estimate $\xi$, then $\hat{\xi}_n = \xi + o_p(\xi)$ converges weakly to a standard normal distribution. This result can not be obtained from the usual weak convergence because $\xi$ is typically random in our setting. For a detailed discussion about the stable convergence, see Jacod and Shiryaev (2003).

**Theorem 3** Suppose that Assumptions (K) and (V) hold and $X$ is continuous. We have

$$
\Delta_n^{-1/2} B_T \xrightarrow{L} (s) MN (0, \Sigma_T),
$$

where

$$
\Sigma_T = \tilde{k} \int_0^T V_{xx} (s, X_s)^2 \sigma_s^4 ds, \quad \tilde{k} = \frac{\pi^2 + 4\pi - 20}{16}.
$$

In order to implement Theorem 3 for statistical inference, we need to consistently estimate the asymptotic variance $\Sigma_T$. We propose two estimators:

$$
\hat{\Sigma} (1)_T^n = \frac{\tilde{k}}{m_{4/3}^{1/3} \Delta_n} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor -2} (V_{xx,i,i-1}^n)^2 \prod_{j=1}^3 \left| \Delta_{n,j-1}^n X \right|^{4/3},
\hat{\Sigma} (2)_T^n = \frac{\tilde{k}}{3 \Delta_n} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} (V_{xx,i,i-1}^n)^2 \left( \Delta_{n}^X \right)^4 1 \{ |\Delta_{n}^X| \leq u_n \},
$$

where the truncation level $u_n$ is given by

$$
u_n = \alpha \Delta_n^{\varpi}, \quad \alpha > 0, \quad \varpi \in (0, 1/2).
$$

These estimators generalize those proposed by Barndorff-Nielsen and Shephard (2006), Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) and Jacod (2008) by accommodating random weights. We refer to $\hat{\Sigma} (1)_T^n$ as the multipower variance estimator, and $\hat{\Sigma} (2)_T^n$ as the truncation variance estimator. Both estimators are consistent, regardless of whether there are jumps or not.

**Theorem 4** Suppose that Assumptions (H-r) and (V) hold for some $r \in [0, 2)$. Let $\hat{\Sigma}_T^n = \hat{\Sigma} (j)_T^n$ for $j = 1$ or 2. When $j = 2$, we further assume that $r < 2$ and $\varpi \in [\frac{1}{4-r}, \frac{1}{2})$. Then $\hat{\Sigma}_T^n = \Sigma_T + o_p(1).$
4.3 The test: size and power

The t-statistic of our test is given by

\[ S_n^T = \frac{\Delta_n^{-1/2} B_n^T}{\sqrt{\Sigma_n^T}}, \]

where \( \Sigma_n^T = \tilde{\Sigma} (j)_T \) for \( j = 1 \) or \( 2 \).

When the price of the underlying is continuous, \( S_n^T \) converges in distribution to a standard normal variable (Theorems 3 and 4). When there are jumps, \( S_n^T \) diverges in probability to \( +\infty \) in the restriction to \( \Omega^b_T \), and \( |S_n^T| \) diverges in probability to \( +\infty \) in the restriction to \( \Omega^b_T \) (Theorems 1 and 4). Size and power properties of tests based on \( S_n^T \) readily follow.

**Corollary 3** For \( q \in (0, 1) \), let \( z_q \) be the \( 1-q \) quantile of \( N(0,1) \). Suppose that Assumption \((K)\) and the same conditions as in Theorem 4 hold.

(a) Under the same conditions as in Theorem 2(a), the asymptotic rejection probability of the one-sided critical region \( C_{T+}^n = \{ S_n^T \geq z_q \} \) is \( q \) under the null hypothesis \( B_T = 0 \), and one under the alternative hypothesis \( B_T > 0 \).

(b) Under the same conditions as in Theorem 2(b), the asymptotic rejection probability of the two-sided critical region \( C_T^n = \{ |S_n^T| \geq z_q/2 \} \) is \( q \) under the null hypothesis \( B_T = 0 \), and one under the alternative hypothesis \( B_T \neq 0 \).

We can interpret the test as a comparison between the jump error and the discretization error, i.e., the hedging error resulted from the discrete implementation of the dynamic hedging strategy. Bertsimas, Kogan, and Lo (2000) and Hayashi and Mykland (2005) show that when \( X \) is continuous, the discretization error is approximately normally distributed with mean zero and standard deviation \( \Delta^{1/2} \sqrt{\Sigma_T / 2k} \), where \( \Sigma_T \) and \( \tilde{k} \) are the same as in Theorem 3 and \( \Delta_{dis} \) is the duration between portfolio rebalancing. Following Bertsimas, Kogan, and Lo (2000), we refer to the quantity \( \sqrt{\Sigma_T / 2k} = \sqrt{\frac{1}{k} \int_0^T V_{xx}(s, X_s)^2 \sigma_s^4 ds} \) as the ex-post temporal granularity.\(^3\) Rewriting the t-statistic as

\[ S_n^T = \frac{1}{\sqrt{2k}} \left( \frac{\Delta_{dis}}{\Delta_n} \right)^{1/2} \frac{B_n^T}{\Delta_{dis}^{1/2} \sqrt{\Sigma_n^T / 2k}}, \]  \( (12) \)

we see that it is proportional to the ratio of the realized jump error to the discretization error measured by the estimate of the ex-post temporal granularity. Up to a multiplicative constant \( 1/\sqrt{2k} \), the scaling factor is determined by the relative magnitude of the “business time scale” \( \Delta_{dis} \) with respect to the statistical time scale \( \Delta_n \). Equation (12) clarifies the economic intuition

\(^3\)Bertsimas, Kogan, and Lo (2000) define the temporal granularity for a derivative contract with pricing function \( (t, x) \rightarrow V(t, x) \) as \( \frac{1}{2} E_0 \left[ \int_0^T V_{xx}(s, X_s)^2 \sigma_s^4 ds \right] \), where \( E_0 \) is the conditional expectation operator given the information at time 0.
behind our test. In practice, a trader can only hedge at discrete times and always suffers from the discretization error, regardless of whether there are jumps or not. Given that the discretization error is inevitable, the trader may use it as a yardstick to judge the relative importance of the jump error, which is implicit in the test. As a by-product of our analysis, we derive consistent estimators for the ex-post temporal granularity, which might also be useful for other applications.

We now discuss the relation between our tests and the bipower test of Barndorff-Nielsen and Shephard (2004). The bipower test is a special case of our one-sided test with $V(t, x) = x^2$. Ignoring the dependence of the option price on the time variable, this pricing function corresponds to a quadratic contract. In particular, the jump error is reduced to the jump quadratic variation $\sum_{s \leq T} (\Delta X_s)^2$. This quantity has received a fair amount of attention in the recent literature of financial econometrics, see e.g., Andersen, Bollerslev, and Diebold (2007), Andersen, Bollerslev, and Dobrev (2007) and Andersen, Bollerslev, and Huang (2010). We can also interpret the jump quadratic variation and the bipower test based on an argument of static hedging. As shown by Leland (1980), one can replicate the final payoff of a quadratic contract with an equally weighted portfolio of the complete cross section of call options, that is, $X_T^2 = 2 \int_0^\infty \max \{X_T - \theta, 0\} d\theta$. Therefore, the jump error of the quadratic contract is also the total jump error of this portfolio of call options. The bipower test can thus be considered as a special way of testing the joint hypothesis that jump errors of the cross section of call options are identically zero by assigning equal weight to each contract.

5 The worst-case jump error and the SUP test

In this section, we extend the univariate test to a multivariate setting by developing a functional sampling theory for the realized jump error profile. We begin with a motivating example. Imagine a trader hedges a few call options which are deeply in the money. Since these options behave quite similar as the underlying stock, they are relatively easy to delta hedge. Consequently, the trader is likely to observe a small realized jump error for her portfolio and conclude that the jump error is insignificant. But we may ask: is this because the trader is lucky enough to manage a trading book which happens to be “jump-neutral” or is it because jumps are not important for hedging at all?

We address these questions by investigating the entire jump error profile in order to get a global view for a family of contracts, instead of any specific portfolio of them. To add some concreteness, we note that the family of vanilla call options is of fundamental interest for two reasons. First, vanilla call and put options are the backbone of the derivative market. Because of the put-call parity, the jump error of a call option is identical to that of a put option with the same configuration. By studying call options, we automatically take care of both types of contracts.
Second, by the result of Leland (1980), the payoff of many options—of which the quadratic contract
is an example—can be statically replicated by holding the underlying stock and a portfolio of call
options. The study of jump errors of these options can thus be reduced to the study of the jump
error profile for vanilla calls. We consider call options in the empirical analysis, although our
econometric framework is not restricted to this special case.

In order to examine whether jumps are costly for hedging in a global manner, we consider a
pessimistic measure defined by
\[ \mathcal{B}_T = \sup_{\theta \in \Theta} B_T (\theta), \]
i.e., the worst-case jump error for the family of interest. For call options, the worst-case contract is often “near the money”, while the exact moneyness depends on the realization of the stock price process, especially when there are multiple
jumps occurring at various price levels. An analysis of the worst case scenario tells us what would
have happened if we were hedging a portfolio which is most sensitive to the jump risk. We also
consider the worst-case absolute jump error defined by
\[ \mathcal{B}_T^a = \sup_{\theta \in \Theta} |B_T (\theta)|. \]

We now discuss the econometric results. We characterize the sampling variability of the realized
jump error profile by extending Theorem 3 to a functional setting. Again, we only consider the
stable convergence in law in the absence of jumps for the testing purpose considered here.

**Theorem 5** Suppose that Assumptions (K) and (V) hold and \( X \) is continuous. The processes
\( (\Delta_n^{-1/2} \mathcal{B}_T^n (\theta) )_{n \geq 1} \) converge stably in law to a process \( U (\theta) \) defined on an extension of the
probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) which, conditionally on \( \mathcal{F} \), is a centered Gaussian process with covariance
function
\[ \Sigma_T (\theta, \theta') = k \int_0^T V_{xx} (s, X_s; \theta) V_{xx} (s, X_s; \theta') \sigma_s^4 ds, \quad \theta, \theta' \in \Theta. \]

The asymptotic properties of \( \mathcal{B}_T^n \) and \( \mathcal{B}_T^a \) follow directly from Theorems 1 and 5:

**Corollary 4** (a) Under the same conditions as in Theorem 1, \( \mathcal{B}_T^n = \mathcal{B}_T + o_p (1) \) and \( \mathcal{B}_T^a = \mathcal{B}_T^a + o_p (1) \).

(b) Under the same setting as in Theorem 5, \( \Delta_n^{-1/2} \mathcal{B}_T^n \overset{\mathcal{L}}{\rightarrow} G_T = \sup_{\theta \in \Theta} U (\theta) \) and
\( \Delta_n^{-1/2} \mathcal{B}_T^a \overset{\mathcal{L}}{\rightarrow} G_T' = \sup_{\theta \in \Theta} |U (\theta)|. \)

Corollary 4(a) shows that \( \mathcal{B}_T^n \) and \( \mathcal{B}_T^a \) are consistent estimators of \( \mathcal{B}_T \) and \( \mathcal{B}_T^a \) respectively.
Corollary 4(b) characterizes the sampling variability of the realized worst-case jump errors under
the null hypothesis. To implement Corollary 4 as a test, we need to estimate quantiles of \( G_T \) and
\( G_T' \).
$G'_T$. For $q \in (0, 1)$, let $\kappa_q$ and $\kappa'_q$ be the $1 - q$ quantiles of $G_T$ and $G'_T$ respectively. We first extend the estimators in (10) by setting $\hat{\Sigma}_T^n (\cdot, \cdot) = \hat{\Sigma}_T^n (j; \cdot, \cdot)$, where for $j = 1$ or 2,

\[
\hat{\Sigma}_T^n (1; \theta, \theta') = \frac{\bar{k}}{n^{4/3} \Delta_n} \sum_{i=1}^{[T/\Delta_n]} V_{xx,i-1}^n(\theta) V_{xx,i-1}^n(\theta') \prod_{j=1}^3 |\Delta_{i+j-1}^n X|^4/3,
\]

\[
\hat{\Sigma}_T^n (2; \theta, \theta') = \frac{\bar{k}}{3 \Delta_n} \sum_{i=1}^{[T/\Delta_n]} V_{xx,i-1}^n(\theta) V_{xx,i-1}^n(\theta') (\Delta_{i}^n X)^4 1\{|\Delta_{i}^n X| \leq u_n\}.
\]

Similar as in the univariate setting, we label these two estimators as multipower- and truncation-based respectively. Let $U_n (\theta)_{\theta \in \Theta}$ be a centered mixture Gaussian process with conditional covariance function $\Sigma_T^n (\cdot, \cdot)$. We set $\hat{\kappa}_q^n$ and $\hat{\kappa}'_q^n$ to be the $1 - q$ quantiles of $\sup_{\theta \in \Theta} U_n (\theta)$ and $\sup_{\theta \in \Theta} |U_n (\theta)|$ respectively. These quantiles, which can be obtained via simulation, serve as consistent estimators of $\kappa_q$ and $\kappa'_q$. Below, we only state results for the test based on the realized worst-case jump error $\bar{B}_T^n$. We refer to this test as the SUP test. A test based on $\bar{B}_T^n$ can be constructed similarly.

**Theorem 6** Suppose that Assumption (K) and the same conditions as in Theorems 4 hold. Then $\hat{\Sigma}_T^n (\theta, \theta')$ converges in probability to $\Sigma_T (\theta, \theta')$ uniformly in $\theta, \theta' \in \Theta$. Moreover, for each $q \in (0, 0.5]$, $\hat{\kappa}_q^n$ (resp. $\hat{\kappa}'_q^n$) converges in probability to $\kappa_q$ (resp. $\kappa'_q$). If we further assume that $V_{xx} (\cdot, \cdot; \theta) > 0$ for some $\theta \in \Theta$, then the asymptotic rejection probability of the critical region $\{\bar{B}_T^n \geq \hat{\kappa}_q^n\}$ is $q$ under the null hypothesis $\bar{B}_T = 0$, and one under the alternative hypothesis $\bar{B}_T > 0$.

### 6 Simulation results

#### 6.1 The setting

We now examine the validity of the asymptotic theory above in a simulation setting designed to approximate the constraints faced in a typical real life application. The data generating process is given by

\[
X_t = X_t^c + J_t,
\]

\[
dX_t^c/X_t^c = v_t dW_t, \quad v_t = c_t^{1/2},
\]

\[
c_t = \kappa (\beta - c_t) dt + \gamma c_t^{1/2} dB_t, \quad \mathbb{E} [dW_t dB_t] = \rho dt.
\]

Here, $X_t^c$ is the continuous part of the price and $J_t$ is a pure jump process to be specified below. The drift part is omitted because it plays little role in the high-frequency setting. Parameters governing the dynamics of the stochastic volatility process are calibrated according to the estimates

\footnote{For any real random variable $\xi$, the $\alpha$-quantile of $\xi$ is defined as $\inf\{x : \text{Prob}(\xi \leq x) \geq \alpha\}$.}
in Aït-Sahalia and Kimmel (2007): \( \beta^{1/2} = 0.4, \gamma = 0.5, \kappa = 5, \rho = -0.5 \) in annualized terms. The continuous-time process is simulated at 5-second interval under the Euler scheme. We then resample the process at every minute or every 5 minutes. We set the horizon \( T \) to be 5 trading days, with each day consisting of 6.5 hours. There are 5,000 simulations in each experiment. We consider one-sided univariate tests based on weekly vanilla call options, henceforth the VC tests, with strike prices spanning the interval \([20, 30]\). For each call option, we set the pricing function \( V(\cdot) \) to be the Black-Scholes formula with zero risk-free rate and dividend yield. The volatility parameter in the Black-Scholes formula is taken to be \( \beta^{1/2} \). Because the pricing function is non-smooth at the strike price at maturity, we smooth away this singularity by adding half a trading day to the maturity of the option. We also conduct the SUP test based on the whole family of call options. For comparison, we conduct the bipower test as a special case of the univariate test corresponding to \( V(t, x) = x^2 \).

In the power analysis, we consider simple jump processes which are engineered to highlight the features of our tests. We consider sample paths with one jump occurring at fixed time \( \tau = 1, 2.5 \) or 4 days in order to examine how the power of the test depends on the timing of jumps. We set the jump size to be \( mX_{\tau-}\sqrt{3} \times 5 \) minutes, \( m = \pm 2.5, \pm 5, \pm 7.5 \), so the jump size is \( m \) times the average standard deviation of the diffusive increment over 5 minutes. To simplify the discussion, we fix the prejump price \( X_{\tau-} = 25 \) so a call option with strike 25 is at the money when the jump occurs. This design help illustrates how the power depends on the moneyness of the option. We refer to VC tests with strikes 22.5, 25 and 27.5 as in-the-money (ITM), at-the-money (ATM) and out-of-the-money (OTM) tests, respectively. In the size analysis, we set the jump process \( J_t \) to be identically zero and \( X_0 = 25 \).

We find that tests based on the truncation variance estimator control size better than those based on the multipower variance estimator. To save space, we only present results based on the truncation variance estimator. We set the truncation level \( u_n = \bar{\alpha} \times 25 \times \beta^{1/2} \Delta_n^{0.49} \) with \( \bar{\alpha} = 5 \). Taking \( \bar{\alpha} = 4 \) or 6 yields similar results which are omitted for brevity.

### 6.2 Size

Table 1 shows the rejection rates of 5% and 10% level tests in the absence of jumps. At the 1-minute sampling frequency, all tests control size well, suggesting that the asymptotic theory is valid. In the 5-minute case, univariate tests (bipower and VC) still have good size control. However, the SUP test tends to over-reject: its rejection rate is 6.5% at the 5% level and 13.1% at the 10% level. We further investigate the sizes of all VC tests with strikes spanning \([20, 30]\). Figure 1 shows the results for 5% level tests. In most cases, the finite-sample rejection rate differs from the nominal level by less than one percentage point. Overall, we find that the tests control size well, as predicted by the asymptotic theory.
6.3 Power

Tables 2 and 3 show the rejection rates of tests under the presence of jumps for 1-minute and 5-minute data, respectively. For brevity, the nominal level is fixed at 5%. An immediate observation is that the rejection rate is higher when data are sampled more frequently. Below, we discuss the results in more details.

Jump size. Not surprisingly, the larger the jump size, the higher the power of each test. For data sampled at 5 minutes, the finite-sample power is very close to the nominal level when the jump size is \( m = 2.5 \) times the standard deviation of the diffusive increment over 5 minutes. In this case, the jump is too small to be detected because it has little effect on hedging. On the other hand, when jumps are large ( \( m = 7.5 \) ) and data are sampled at every minute, the bipower, the SUP and the ATM tests have almost perfect power.

Jump time and moneyness. In Figure 2, we plot the rejection rate versus the jump time for various jump magnitudes (rows) and sampling frequencies (columns). The rejection rate of the bipower test (dashed line) does not depend the jump time because it corresponds to the time-homogeneous function \( V(t, x) = x^2 \). On the contrary, when the jump occurs near maturity, the SUP test (square) and the ATM test (circle) reject more often, but the ITM test (X) and the OTM test (triangle) reject less. The intuition is that as the time-to-maturity approaches zero, the pricing function become “more nonlinear (resp. linear)” around (resp. far from) the strike price. Therefore, a jump around (resp. far from) the strike price has a larger effect if it occurs later (resp. earlier) during the life of the option. In most cases, the ATM test is more powerful than the ITM and the OTM tests, reflecting the fact that jumps induce large hedging errors for call options when they occur at the money. In Figure 3, we further plot the rejection rates of all VC tests versus their strikes. The power curve forms a peak around the prejump price, i.e. 25, when the jump occurs late in the sample (dot-dashed and dashed lines); however, the curve is relatively flat across strikes when the jump occurs early (solid line). This pattern suggests that the power of VC tests is more sensitive to the strike price if the jump time is closer to maturity.

Bipower, SUP and ATM. The ATM test is always more powerful than the SUP test. Indeed, the ATM test directly exploits the fact that the strike price of the worst-case contract is close to the pre-jump price. The pre-jump price is known in our experiment by design but is unknown in practice because jumps are not directly observable. On the other hand, the SUP test has to “find” the worst-case contract. Therefore, the ATM test can be roughly thought of as an infeasible version of the SUP test. As the former exploits extra information, it is not surprising it has higher power than the latter. Except for a few cases, the ATM is also more powerful than the bipower test. The comparison between the SUP test and the bipower test depends on the jump time: the former is more powerful than the latter when the jump is closer to maturity, and vice versa.
Positive and negative jumps. Rejection rates of the bipower, the SUP and the ATM tests do not seem to depend on the sign of the jump. On the other hand, in most cases, the OTM test tends to reject more often for positive jumps than for negative jumps; and the opposite is true for the ITM test. To illustrate this asymmetry graphically, we plot in Figure 4 the rejection rates of VC tests versus their strike prices for positive and negative jumps with various magnitudes. When the strike is low, positive jumps (solid lines) lead to lower rejection rates than negative jumps (dashed lines), and vice versa. Intuitively, when the option is in the money, a positive jump makes the pricing function “more linear”, while a negative jump makes the pricing function “more nonlinear”. It is thus easier for an ITM test to pick up negative jumps than positive ones.

To summarize, we illustrate how the power of various tests depends on features of the jump process, such as the magnitude, the direction and the timing of jumps. The power of a test is high if the corresponding option is difficult to hedge under the presence of jumps. The purpose of this analysis is not to conduct a horse race among various tests. Instead, we attempt to provide some economic insight on why one test is more powerful than the other given certain features of the jump process. Through the lens of hedging, our view is that jumps “deserve” to be detected only if they have important economic consequences. This idea might be unconventional from a statistical point of view, but seems natural in an econometric context.

7 Empirical analysis

7.1 The setting

Our sample consists of the 30 component stocks of Dow Jones Industrial Average (DJIA) in 2008; the data source is the TAQ database. Because the composition of the DJIA changes over time, we use the 30 stocks that are the components of the index as of September, 16th, 2010. We use filters to eliminate clear data errors (price set to zero, etc.) as is standard in the empirical market microstructure literature. For each trading day in 2008, we collect all transactions from 9:30am until 4:00pm, and compute the volume-weighted average of transaction prices at each time stamp for each one of these stocks. Following the common practice in the literature, we sparsely sample the intraday data at every 5 minutes to reduce the effect of microstructure noise. We perform the tests for intraday returns. We also conduct the same analysis including overnight returns within each week. In so doing, we consider intraweek stock price jumps in U.S. trading time under the assumption that the hedging portfolio is locked up after the U.S. stock market closes. Each week and stock is treated on its own.

The empirical setting is similar as in the simulation. We consider vanilla call options with
moneyness spanning \([0.85,1.15]\). We take the pricing function \(V(\cdot)\) to be the classical Black-Scholes formula with constant volatility which is updated on a weekly basis. The interest rate and the dividend yield are set to be zero. Without any attention for advocating the Black-Scholes model as a hedging device, we choose this simple strategy only for the purpose of illustrating our method. Nevertheless, we note that this simple setup might convey a broader message. Indeed, as documented by Bakshi, Cao, and Chen (1997) and Chernov and Ghysels (2000), the Black-Scholes strategy has similar hedging performance as models with stochastic volatility and jumps. We conjecture that using more sophisticated pricing models may lead to similar results as ours. An extensive comparison is beyond the scope of this paper.

The key tuning parameter in the Black-Scholes hedging strategy is the volatility parameter, which we denote by \(v_{BS}\). For each stock-week, we calibrate \(v_{BS}\) to be the ex-post average volatility \(\sqrt{\frac{1}{T} \int_0^T v_s^2 ds}\) for the week, where we recall that \(v_t = \sigma_t / X_t\) is the volatility. We approximate this quantity by the truncated pre-averaging estimator, which, denoted by \(\hat{v}\), varies across stocks and weeks. We are thus considering the benchmark case in which we are “correct”, up to estimation errors, about the volatility on average. We perform a sensitivity analysis with other choices of \(v_{BS}\) later.

We implement the tests with the truncation variance estimator. The truncation level is set to be \(u_n = 5 \hat{X} \hat{v} \Delta_n^{0.49}\), where \(\hat{X}\) is the weekly average stock price. Similarly as in the simulation, we adding half a trading day (3.25 hours) to the time-to-maturity in order to smooth away the singularity of the pricing function at the strike at maturity.

### 7.2 An illustrative example

To help visualize concepts, we start by illustrating the econometric framework via an example based on the stock of Microsoft Corporation (NASDAQ: MSFT). In the top panel of Figure 5, we plot the time series of the stock price for two weeks: week-A beginning August 11th (left) and week-B beginning November 10th (right). A visual inspection suggests that week-A is likely to have a continuous sample path, even though overnight returns (circles) are included, and week-B is likely to have jumps: one candidate is the overnight drop at the market’s open on Friday, and

---

\(^5\) We define the moneyness of a call option as the ratio of the strike price to the weekly opening stock price. For computational purposes, we discretize the interval \([0.85,1.15]\) into grids with mesh size 0.01. Preliminary results (not presented here) suggest that finer mesh size (0.0025) has little effect on the results.

\(^6\) The estimate is computed based on data sampled at every 5 seconds. This estimator is robust against both microstructure noise and jumps. See Lemma 6 of Aït-Sahalia, Jacod, and Li (2011) for its definition and theoretical justification.

\(^7\) Among all 53 weeks in 2008, there are two weeks in which Fridays are national holidays (Good Friday and Independence Day). For these weeks, we consider a weekly option with maturity being 4.5 days. The last day of 2008 is Wednesday. The maturity of the option in this week is set to be 3.5 days.
another is the large drop near the market’s close on the same day.

For each week, we plot the realized jump error profile with overnight returns included (solid line) or excluded (dashed line) versus the option moneyness in the bottom panel of Figure 5. Clearly, the profiles of week-B are more pronounced than those of week-A. In addition, overnight returns in week-B have an evident effect on the jump error, but not so much in week-A. These observations are consistent with what we observe in the time series. We also plot the 5% critical value of each VC test (dotted line). In week-A, the realized jump error is less than the critical value for every call option. In week-B, we reject the null hypothesis of zero intraday jump error when the moneyness is, roughly speaking, between 0.9 and 0.94. To see the intuition, we note that the large drop at the end of the sample occurs when the stock price is around $20.50 which is approximately 0.94 times the weekly opening price $21.85. Options with moneyness between 0.9 and 0.94 is “near the money” when this large jump occurs and thus leads to a large jump error. When we include overnight returns into the analysis, VC tests reject the null hypothesis for moneyness between 0.89 and 1.13, suggesting that overnight jumps have a significant effect on the whole cross section of call options.

The SUP test eliminates the ambiguity in the contract-by-contract exercise by concentrating on the worst-case scenario. In the bottom panel of Figure 5, we plot the uniform acceptance region (shaded area) of the SUP test, where the upper bound of the acceptance region is \( \bar{\kappa}_{0.05} \) (Theorem 6). Not surprisingly, we reject the null hypothesis that the worst-case jump error is zero for week-B at 5% nominal level no matter whether overnight returns are included or not. We do not reject the same null hypothesis for week-A as the acceptance region covers the realized jump error profiles.

### 7.3 Cross-sectional analysis

We start by performing tests at 5% nominal level for each week and stock. Figure 6 plots the empirical rejection rate of VC tests versus the option moneyness (solid line). We first discuss the case when overnight returns are excluded (left panel). When the moneyness is near 1, VC tests reject the null hypothesis at about 41% of the time. The rejection rate decreases as the moneyness deviates from 1: the rejection rates at moneyness = 0.9 and 1.1 are 26% and 28% respectively. The pattern is similar when we include overnight returns in the calculation (right panel). Because overnight returns often appear to be jumps to the naked eye, it is not surprising that the rejection rate increases after overnight returns are included. What is remarkable is the magnitude of this increase—the rejection rate approximately doubles—suggesting that overnight jumps make a major contribution to the jump error.

We then perform the SUP test for these call options. We also conduct the bipower test for comparison. Table 4 compares the rejection rates of the SUP test, the bipower test and the VC test associated with the ATM option, which, with some abuse of notation, will be referred to as the
The rejection rates of these tests turn out to be quite similar: when overnight returns are excluded, the rejection rate ranges from 41% to 46%; when overnight returns are included, the rejection rate ranges from 86% to 90%. After all, each of these three tests can be thought of a test based on the jump error profile of the cross section of call options, but corresponding to different weights on the options.

In view of the large amount of attention that has been given to the bipower test in empirical studies, we find the similarity between results of the ATM test, the SUP test, and the bipower test comforting. Through a hedging perspective, our empirical finding sheds some new light on the economic implication of the existing results based on the bipower test. This said, it should also be kept in mind when and how these tests may give different results, as we have discussed in the simulation study. What is more, by examining the jump error profile as an entity, our approach provides a more complete picture of the economic relevance of jumps; it is thus largely complementary to the existing tests for jumps.

We have documented that jumps induce statistically significant hedging errors. We now discuss whether the jump errors are economically sizable. Below, our purpose is to summarize the point estimates of the jump error for all stock-weeks in our sample, instead of making formal statistical inferences. To gauge the economic significance of the jump error, we compute the realized volatility error (Corollary 1) as a yardstick. Because the volatility error is two-sided but the jump error is one-sided for call options, we focus on the upper percentiles of these hedging errors in order to make a meaningful comparison. Table 5 shows the results. In Panel A, the estimated hedging errors are normalized with respect to the model price of the ATM call option, which is computed based on the Black-Scholes model with the same configuration as that of the hedging strategy. In Panel B, we express the hedging error in terms of implied volatility by normalizing it with respect to the option vega (i.e. the partial derivative of the option price with respect to the volatility). We shall focus our discussion on Panel B, noting that Panel A delivers a similar message. We make two observations. First, the realized intraday jump error (the second row) has a similar magnitude as the realized volatility error in the tail. Second, when overnight jumps are included (the third row), the realized jump error increases significantly over all percentiles: in terms of implied volatility, the median increases from 0.9% to 2.9%, and the 99th percentile increases from 9.1% to 45.5%.

To have a rough sense about the magnitude of these estimates, we use option data to compute the premium of short-dated ATM options, which is defined as the difference between the implied volatility of a 30-day call or put option and the realized volatility of the underlying stock. We find

---

8 Here, the ATM test is defined differently than that in the simulation study. The moneyness is defined relative to the opening price of the stock, while in the simulation study, it is defined relative to the stock price at the jump time.

9 The data source is the Option Metrics database. For each trading day, we extract the implied volatility smile from the volatility surface file for 30-day call and put options. The implied volatility of an ATM option is obtained
that for an average DJIA stock on an average day in 2008, the premium is approximately 1.9%.
Compared with this premium, the realized jump errors seem quite large, especially in the tail.

These results suggest that hedging errors contributed by intraday jumps are economically sizable and have similar magnitude as those contributed by stochastic volatility. Yet the most significant source of hedging error is the overnight jumps. We thus argue that, compared with intraday jumps and stochastic volatility, the periodic closure of the stock market might be a greater source of market incompleteness, at least in the simple hedging context considered here.\footnote{This finding points out a limitation of prior works on empirical option pricing (see Pan (2002), Eraker (2004), Broadie, Chernov, and Johannes (2007)). These studies are based on daily returns and make no distinction between intraday and overnight movements in stock prices. Of course, such a distinction would be redundant if the hedger only rebalance her portfolio at daily or lower frequencies. However, improvements in trading technologies and the falling transaction cost make the assumption of infrequent rebalancing increasingly unreasonable. When intraday rebalancing is feasible, the hedger could reduce the risk underlying intraday price movements by dynamic hedging, but can not do so during the market closure. Pooling intraday and overnight returns obscures the risk factor carried by each component. Modelling intraday and overnight price dynamics and risk premia separately seems to deserve some attention in future research on empirical option pricing.}

7.4 Robustness check

So far, we have only considered the Black-Scholes hedging strategy with the a posteriori “correct” weekly average volatility $\tilde{v}$. To check the robustness of our results, we set $v_{BS} = \eta \tilde{v}$, $\eta = 0.75$ or 1.25, and then repeat the tests. Note that when $\eta = 1$, we are back to the setting in Section 7.3. To make a meaningful comparison across different values of $v_{BS}$, we adjust the definition of moneyness by setting:

$$\text{Adjusted Moneyness} = \frac{\text{Moneyness} - 1}{\eta} + 1.$$  

Figure 6 plots the rejection rate of VC tests versus the adjusted moneyness. As $v_{BS}$ increases, the curve of rejection rate levels up by roughly the same amount across moneyness. When $v_{BS}$ varies by a factor of 0.25, the rejection rate of the ATM test changes by 1 percentage point when overnight returns are excluded (Panel A) and by 2 percentage points when overnight returns are included (Panel B). Since we have considered a fairly large perturbation on $v_{BS}$, the VC tests seem to be quite robust relative to the choice of this tuning parameter. For the same perturbation, we find that the rejection rate of the SUP test varies by less than 1 percentage point.

We further investigate the degree to which the realized jump error depends on $v_{BS}$ by computing the sensitivity measure $\partial B^n_T (v_{BS}) / \partial v_{BS}$ for ATM options as described in Corollary 2. We evaluate this measure at $v_{BS} = \tilde{v}$ and express it in terms of implied volatility. Table 6 summarizes the results. In Column 1, we report the sample median of the absolute value of the sensitivity measure (MAS). We find that the realized jump error is quite robust to local changes in $v_{BS}$: if $v_{BS}$ changes by linearly interpolating the volatility smile.
by 100 basis points (bp), the realized jump error only changes by 2 bps (resp. 5 bps) in terms of implied volatility on the median when overnight returns are excluded (resp. included). In order to set a yardstick for these estimates, we report the sample standard deviation (SD) of the realized jump errors in column 2 and shows the ratio of MAS to SD in column 3. Again, we find that the realized jump error is robust to changes in $v_{BS}$: a change of 100 bps in $v_{BS}$ only leads to a 0.01 standard deviation change of the realized jump error, regardless whether overnight returns are included or not.

8 Concluding remarks

Derivative hedging seems to be a proper domain for a discussion of the jump risk. We measure stock price jumps by the hedging error they induce and propose a general framework for nonparametrically testing the presence of jumps based on this measure. In a hedging context, we shed new light on the economic relevance of jump tests. We demonstrate how to apply the framework in practice and document the economic relevance of jumps detected in high frequency data.

In principle, our measure relies on the derivative security and the form of the hedging strategy, so our conclusion regarding the economic importance of jumps might depend on these choices. However, we consider this flexibility as an advantage instead of a drawback. After all, hedging is the key element of derivative trading and specifying what and how to hedge is the minimal requirement for discussing hedging in the first place. By considering various derivative securities and delta-hedging specifications, we can get a better understanding about how the jump risk manifests in various contexts. Our functional framework is general enough for studying such multi-dimensional or even infinite-dimensional problems.

References


9 Proofs

9.1 Proof of Theorem 1

As often in this kind of problem, it is convenient to strengthen Assumptions (H-r) and (K) as follows:

Assumption (SH-r): We have Assumption (H-r), \( sup_{z\in E} |\delta(\omega, t, z)| \leq \gamma(z) \), and the processes \( b_t, \sigma_t, \) and \( X \) itself are bounded.

\[ \square \]

Assumption (SK): We have Assumptions (K) and (SH-2), and further the processes \( \tilde{b}_t, a_t, a'_t, \tilde{\sigma}_t \) are bounded.

We first partially extend Theorem 8.4.1 and Theorem 9.2.1 in Jacod and Protter (2010) in order to accommodate random weights. These results are used repeatedly in our proofs. For \( a \geq 0 \) and \( k \geq 1 \), we denote by \( \rho^k_\alpha \) the joint law of \( k \) independent centered Gaussian variables with variance \( a^2 \). For any measurable function \( f \) on \( \mathbb{R}^k \), we write \( \rho^k_\alpha(f) = \int f(x_1, \ldots, x_k) \rho^k_\alpha(dx_1, \ldots, dx_k) \), provided that the integral is well-defined. We write \( \rho^k_\alpha \) as \( \rho_\alpha \) for notational simplicity. Below, we denote by \( K \) a generic constant which varies from line to line.

Lemma 1 Suppose that Assumption (SH) holds. Let \( Y \) a bounded càdlàg adapted process and \( F \) be a continuous function on \( D^k, k \geq 1 \), satisfying

\[ |F(x_1, \ldots, x_k)| \leq \prod_{j=1}^k \psi(|x_j|) \left( 1 + |x_j|^2 \right), \]

where \( \psi \) is a continuous function on \([0, \infty)\) which goes to zero at infinity. We have

\[
\Delta_n \sum_{i=1}^{[t/\Delta_n]-k+1} Y_{(i-1)\Delta_n} F \left( \Delta_n^i X/\sqrt{\Delta_n}, \ldots, \Delta_n^{i-k+1} X/\sqrt{\Delta_n} \right) \to \int_0^t Y_s \rho^{k}_\alpha(F) \, ds. \tag{13}
\]

Proof. We set

\[
\begin{align*}
\beta^n_{i,j} &= \sigma_{(i-1)\Delta_n} \Delta_n^{i+j-1} W/\sqrt{\Delta_n}, \\
\zeta^n_i &= F \left( \beta^n_{i,1}, \ldots, \beta^n_{i,k} \right), \\
\chi^n_i &= \Delta_n \left[ F \left( \Delta_n^i X/\sqrt{\Delta_n}, \ldots, \Delta_n^{i+k-1} X/\sqrt{\Delta_n} \right) - \zeta^n_i \right], \\
H^n_t &= \mathbb{E} \left( \sum_{i=1}^{[t/\Delta_n]-k+1} |\chi^n_i| \right), \\
H'^n_t &= \mathbb{E} \left( \sum_{i=1}^{[t/\Delta_n]-k+1} |Y_{(i-1)\Delta_n} \chi^n_i| \right).
\end{align*}
\]

Since \( Y \) is bounded, \( H'^n_t \leq KH^n_t \). It is shown in the proof of Theorem 8.4.1 of Jacod and Protter (2010) that \( H'^n_t \to 0 \). We thus have \( H'^n_t \to 0 \). Therefore, it is enough to show (13) with the left side replaced by \( \Delta_n \sum_{i=1}^{[t/\Delta_n]-k+1} Y_{(i-1)\Delta_n} \zeta^n_i \).

We set \( \zeta(1)_i^n = \mathbb{E} \left[ \zeta^n_i |F_{(i-1)\Delta_n} \right] \) and \( \zeta(2)_i^n = \zeta^n_i - \zeta(1)_i^n \). Note that \( \zeta(1)_i^n = \rho^k_{\sigma_{(i-1)\Delta_n}}(F) \). We use the Riemann approximation to derive

\[
\Delta_n \sum_{i=1}^{[t/\Delta_n]-k+1} Y_{(i-1)\Delta_n} \zeta(1)_i^n \to \int_0^t Y_s \rho^k_{\sigma_s}(F) \, ds. \tag{14}
\]
Since \( \mathbb{E} \left[ (2)_n \mid F_{(i-1)\Delta_n} \right] = 0 \) by construction, and \( (2)_n \) is \( F_{(i+k-1)\Delta_n} \)-measurable, \( (2)_i^n \) and \( (2)_{i+j}^n \) are uncorrelated whenever \( |j| \geq k \). Noting that \( \mathbb{E}[(2)_i^n] \leq K \), we use the Cauchy-Schwarz inequality to derive

\[
\mathbb{E} \left[ \Delta_n \sum_{i=1}^{[t/\Delta_n] - k + 1} Y_{(i-1)\Delta_n} (2)_i^n \right]^2 \leq K \Delta_n \to 0. \tag{15}
\]

We finish the proof of (13) by combining (14) and (15).

**Lemma 2** Suppose that Assumption (H-r) holds for some \( r \in [0, 2] \), and let \( u_n \) satisfy (11). Let \( Y \) a bounded càdlàg adapted process and \( F \) be a continuous function on \( \mathbb{R}^k \) which satisfies for some \( p \geq 2 \),

\[
|F(x_1, \ldots, x_k)| \leq K \prod_{j=1}^k (1 + \|x_j\|^p).
\]

The variables \( \xi_i^n \) defined by

\[
\xi_i^n = \Delta_n \sum_{i=1}^{[t/\Delta_n] - k + 1} Y_{(i-1)\Delta_n} F \left( \Delta_i^n X/\sqrt{\Delta_n}, \ldots, \Delta_{i+k-1}^n X/\sqrt{\Delta_n} \right) \prod_{j=1}^k 1_{\{|\Delta_{n,j-1}^n X| \leq u_n\}}
\]

converge in probability to \( \int_0^t Y_s \rho_s^k (F) \, ds \) when one of the following conditions holds: (a) \( X \) is continuous; (b) \( p = 2 \); (c)

\[
p > 2, r < 2, \omega \geq \frac{p - 2}{2(p - r)}.
\]

**Proof.** By localization, we suppose that Assumption (SH-r) holds. Let \( \psi \) be a \( C^\infty \) function on \( \mathbb{R}_+ \), satisfying \( 1_{[1,\infty)} \leq \psi(x) \leq 1_{[1/2,\infty)}(x) \), \( x \in \mathbb{R}_+ \). For each \( m \in \mathbb{N} \), we set \( \psi^'_m(x) = 1 - \psi(x/m) \), \( x \in \mathbb{R}_+ \), and \( F_m(x_1, \ldots, x_k) = F(x_1, \ldots, x_k) \prod_{j=1}^k \psi^'_m(x_j) \). We set

\[
\gamma_i^n = \Delta_n F \left( \Delta_i^n X/\sqrt{\Delta_n}, \ldots, \Delta_{i+k-1}^n X/\sqrt{\Delta_n} \right) \prod_{j=1}^k 1_{\{|\Delta_{n,j-1}^n X| \leq u_n\}},
\]

\[
\gamma(m)_i^n = \Delta_n F_m \left( \Delta_i^n X/\sqrt{\Delta_n}, \ldots, \Delta_{i+k-1}^n X/\sqrt{\Delta_n} \right),
\]

\[
H(m)_i^n = \mathbb{E} \left( \sum_{i=1}^{[t/\Delta_n] - k + 1} |\gamma_i^n - \gamma(m)_i^n| \right),
\]

\[
\tilde{H}(m)_i^n = \mathbb{E} \left( \sum_{i=1}^{[t/\Delta_n] - k + 1} |Y_{(i-1)\Delta_n} |\gamma_i^n - \gamma(m)_i^n| \right).
\]

The following convergence is implicit in the proof of Theorem 9.2.1 in Jacod and Protter (2010):

\[
\lim_{m \to \infty} \limsup_{n \to \infty} H(m)_i^n = 0.
\]

Since \( \tilde{H}(m)_i^n \leq KH(m)_i^n \), we have

\[
\lim_{m \to \infty} \limsup_{n \to \infty} \tilde{H}(m)_i^n = 0. \tag{16}
\]

Let \( \gamma(m)_i^n = \sum_{j=i}^{[t/\Delta_n] - k + 1} Y_{(i-1)\Delta_n} \gamma(m)_i^n \). Since \( F_m \) is continuous and supported on a compact set, we use Lemma 1 to derive

\[
m \geq 1 \Rightarrow \gamma(m)_i^n \to^p \int_0^t Y_s \rho_s^k (F_m) \, ds. \tag{17}
\]

29
Since \( F_m \to F \), the dominated convergence theorem yields
\[
\int_0^t Y_s \rho_x^k (F_m) \, ds \to \int_0^t Y_s \rho_x^k (F) \, ds, \quad \text{as } m \to \infty.
\] (18)
Note that \( \mathbb{E}[(m)_t^p - (\xi)_t^p] \leq H (m)_t^p \). With an appeal to Markov’s inequality, we finish the proof by combining (16), (17) and (18).

Proof of Theorem 1. It suffices to show (i) \( A_T^{m} (\cdot), \tilde{A}_T^{m} (\cdot) \) and \( B_T^{m} (\cdot) \) are stochastically equicontinuous and (ii) the convergences in probability hold for each \( \theta \in \Theta \). By localization, we suppose that Assumption (SH-2) holds without loss of generality.

Rewriting \( \varphi (t, x, z) \) in integral form as \( \int_0^z dy_1 \int_0^{y_1} V_{xx} (t, x + y_2) \, dy_2 \), we have for \( \theta, \theta' \in \Theta, \)
\[
A_T^{m} (\theta') - A_T^{m} (\theta) = \sum_{i=1}^{[T/\Delta_n]} \int_0^{\Delta_n} dy_1 \int_0^{y_1} (V_{xx} ((i - 1) \Delta_n, X_{(i-1)\Delta_n-} + y_2; \theta') - V_{xx} ((i - 1) \Delta_n, X_{(i-1)\Delta_n-} + y_2; \theta)) \, dy_2.
\] (19)
Combined with the Lipschitz condition in Assumption (V), (19) yields \( |A_T^{m} (\theta') - A_T^{m} (\theta)| \leq K d (\theta', \theta) \sum_{i=1}^{[T/\Delta_n]} (\Delta_n^p X)^2 \). Since the realized variance \( \sum_{i=1}^{[T/\Delta_n]} (\Delta_n^p X)^2 = O_p (1) \), \( A_T^{m} (\cdot) \) is stochastically equicontinuous by Theorem 21.10 of Davidson (1994). Similarly, \( |\tilde{A}_T^{m} (\theta') - \tilde{A}_T^{m} (\theta)| \leq K d (\theta', \theta) \sum_{i=1}^{[T/\Delta_n]} (\Delta_n^p X) |\Delta_n^p X - (\Delta_n^p X)| = O_p (1) \), \( A_T^{m} (\cdot) \) is stochastically equicontinuous by the same reasoning. The stochastic equicontinuity of \( B_T^{m} (\cdot) \) readily follows.

We now fix some \( \theta \in \Theta \). For notational simplicity, we suppress the dependence on \( \theta \). We prove (8) by verifying the conditions of Theorem 7.3.6 of Jacod and Protter (2010). We define a real-valued function \( F \) on \( (\Omega \times [0, T] \times \mathbb{R}, \mathcal{F} \otimes \mathcal{B} ([0, T]) \otimes \mathcal{R}) \) by setting
\[
F (\omega, t, z) = \varphi (t, X_{t-} (\omega), z), \quad \omega \in \Omega, \ t \in [0, T], \ z \in \mathbb{R}.
\]
For any sequences \( t_n \) and \( z_n \) satisfying \( t_n < t \), \( t_n \to t \) and \( z_n \to z \), we have \( F (\omega, t_n, z_n) \to F (\omega, t, z) \) for every \( \omega \in \Omega \), because \( \varphi \) is continuous and \( X_- \) is left-continuous. Let \( \alpha_t (\omega) = \frac{1}{2} V_{xx} (t, X_{t-} (\omega)) \). The process \( \alpha \) is measurable with càglàd paths. The Taylor expansion yields, for some \( \tilde{z} \) between 0 and \( z \),
\[
\varphi (t, X_{t-} (\omega), z) - \alpha_t (\omega) z^2 = \frac{1}{2} V_{xxx} (t, X_{t-} (\omega) + \tilde{z}) (\tilde{z})^2.
\]
For \( |z| \leq 1 \), we then use Assumption (V) to derive
\[
|F (\omega, t, z) - \alpha_t (\omega) z^2| \leq K |z|^3.
\]
With the above setting, we readily verify the conditions of Theorem 7.3.6 of Jacod and Protter (2010) and obtain
\[
\sum_{i=1}^{[T/\Delta_n]} F (\cdot, (i - 1) \Delta_n, \Delta_n^p X) \overset{p}{\to} \frac{1}{2} \int_0^T V_{xx} (s, X_s) \sigma_s^2 \, ds + B_T.
\] (20)
Note that \( \sum_{i=1}^{[T/\Delta_n]} F (\cdot, (i - 1) \Delta_n, \Delta_n^p X) = \sum_{i=1}^{[T/\Delta_n]} \varphi ((i - 1) \Delta_n, X_{(i-1)\Delta_n-}, \Delta_n^p X) \). Since \( X \) has no fixed time of discontinuity, on almost every path, we have \( X_{(i-1)\Delta_n} = X_{(i-1)\Delta_n-} \) for all \( 1 \leq i \leq [T/\Delta_n] \) and \( n \in \mathbb{N} \). Therefore, almost surely, \( A_T^{m} = \sum_{i=1}^{[T/\Delta_n]} F (\cdot, (i - 1) \Delta_n, \Delta_n^p X) \). Combining this result with (20), we have (8).

Applying Lemma 1 with \( F (x_1, x_2) = \frac{1}{2m} |x_1| |x_2| \) and \( Y_t = V_{xx} (t, X_t) \), we have (9). The convergence of \( B_T^{m} \) readily follows. \( \square \)
9.2 Proof of Theorem 2

Proof. (a) The claim is an easy consequence of the Taylor expansion.

(b) For each \((r, t, x) \in \mathbb{R} \times [0, T] \times D\), we define a set \(Z(r, t, x) = \{z \in \mathbb{R} : z \neq 0, \varphi(t, x, z) + r = 0\}\). Under condition (iii), this set is countable. By construction, for each \(k \geq 1\), \(\sum_{q \leq k} \varphi(T_q, X_{T_q -}, \Delta X_{T_q}) = 0\) is equivalent to \(\Delta X_{T_k} \in Z(\sum_{q \leq k} \varphi(T_q, X_{T_q -}, \Delta X_{T_q}), T_k, X_{T_k -})\). Since the \(Z(\cdot)\) is countable and only replies on the information in \(\mathcal{F}_{T_k -}\) (noting that the stopping time \(T_k\) is \(\mathcal{F}_{T_k -}\)-measurable), condition (ii) implies that for each \(k \in \mathbb{N}\), \(\sum_{q \leq k} \varphi(T_q, X_{T_q -}, \Delta X_{T_q}) = 0\) almost surely. Since there are only finitely many jumps in finite horizon (condition (i)), we have \(\Omega_T^n = \Omega_T^b\) almost surely. \(\square\)

9.3 Proof of Theorems 3 and 5

For \(\theta \in \Theta\), let

\[
A_T^n (\theta) = \frac{1}{2} \sum_{i=1}^{[T/\Delta_n]} V_{XXX} ((i - 1) \Delta_n, X_{(i-1)\Delta_n}; \theta) (\Delta_n^3 X)^2,
\]

\[
B_T^n (\theta) = A_T^n (\theta) - \bar{A}_T^n (\theta).
\]

We first show that when \(X\) is continuous, \(A_T^n (\cdot)\) and \(B_T^n (\cdot)\) are accurate approximations for \(A_T^\infty (\cdot)\) and \(B_T^\infty (\cdot)\) respectively in a functional sense.

**Lemma 3** Suppose that Assumptions (SH-2) and (V) hold. When \(X\) is continuous, \(\Delta_n^{-1/2} (A_T^n (\theta) - A_T^\infty (\theta)) = o_p (1)\) uniformly in \(\theta \in \Theta\). Consequently, \(\Delta_n^{-1/2} (B_T^n (\theta) - B_T^\infty (\theta)) = o_p (1)\) uniformly in \(\theta \in \Theta\).

Proof. We set

\[
\psi(t, x, z; \theta) = V(t, x + z; \theta) - V(t, x; \theta) - \sum_{j=1}^{3} \frac{1}{j!} \frac{\partial^j V(t, x; \theta)}{\partial x^j} z^j.
\]

We have the following decomposition

\[
\Delta_n^{-1/2} (A_T^n (\theta) - A_T^\infty (\theta)) = \sum_{j=1}^{2} \zeta(j; \theta)_T^n, \text{ where}
\]

\[
\zeta(1; \theta)_T^n = \frac{1}{6} \Delta_n \sum_{i=1}^{[T/\Delta_n]} V_{XXX} ((i - 1) \Delta_n, X_{(i-1)\Delta_n}; \theta) \left(\frac{\Delta_n^3 X}{\Delta_n^2}\right)^3
\]

\[
\zeta(2; \theta)_T^n = \Delta_n^{-1/2} \sum_{i=1}^{[T/\Delta_n]} \psi ((i - 1) \Delta_n, X_{(i-1)\Delta_n}; \Delta_n^3 X; \theta).
\]

Since \(X\) is bounded under Assumption (SH-2), the mean-value theorem and Assumption (V) imply \(\sup_{\theta \in \Theta} |\zeta(2; \theta)_T^n| \leq K \Delta_n^{-1/2} \sum_{i=1}^{[T/\Delta_n]} |\Delta_n^3 X|^4\). Since \(X\) is continuous, we use the Burkholder-Davis-Gundy inequality to derive \(\mathbb{E}[|\Delta_n^3 X|^4] \leq K \Delta_n^3\). Thus \(\mathbb{E}[\sup_{\theta \in \Theta} |\zeta(2; \theta)_T^n|] \leq K \Delta_n^{1/2}\) and \(\zeta(2; \theta)_T^n = o_p (1)\) uniformly in \(\theta \in \Theta\).
It remains to show that $\zeta (1; \theta)_T^n \equiv o_p(1)$ uniformly in $\theta \in \Theta$. Since $X$ is bounded, Assumption (V3) yields that for $\theta, \theta' \in \Theta$,

$$|\zeta (1; \theta')_T^n - \zeta (1; \theta)_T^n| \leq K d(\theta, \theta') B_n,$$

where $B_n = \Delta_n \sum_{i=1}^{[T/\Delta_n]} \left( \frac{\Delta_n X}{\Delta_n^{1/2}} \right)^3$.

By the Burkholder-Davis-Gundy inequality, we derive $B_n = o_p(1)$. With an appeal to Theorem 21.10 of Davidson (1994), we see that $(\zeta (1; \theta)_T^n : n \geq 1, \theta \in \Theta)$ is stochastically equicontinuous. Therefore, it is enough to show that $\zeta (1; \theta)_T^n = o_p(1)$ for each $\theta \in \Theta$. We hence fix some $\theta$ and suppress the dependence on $\theta$ in our notation below. Consider a real-valued function on $[0, T] \times [0, T] \cap F \otimes B ([0, T]) \otimes \mathcal{R}$ defined by

$$G(\omega, t, z) = \frac{1}{6} V_{xxx}(t, X_t(\omega)) z^3, \quad \omega \in \Omega, \; t \in [0, T], \; z \in \mathbb{R}.$$

Because $X$ and $V_{xxx}$ are continuous, the mapping $(t, z) \mapsto G(\omega, t, z)$ is continuous for every $\omega$. Moreover, $\sup_{t \in [0, T], \omega \in \Omega} V_{xxx}(t, X_t(\omega)) < \infty$ because of Assumptions (SH-2) and (V2). We have verified the conditions of Theorem 7.2.2 of Jacod and Protter (2010). Taking $\tau(n, i) = (i - 1) \Delta_n$ in that theorem and noting that $G(\omega, t, z)$ is odd in $z$, we obtain

$$\zeta (1)_T^n \overset{p}{\to} \int_0^t ds \int_\mathbb{R} G(\cdot, s, z) \rho_{\sigma_s} (dz) = 0.$$ 

This finishes the proof. \hfill \Box

We now study the functional stable convergence in law of $(B^n_T(\theta))_{\theta \in \Theta}$. To simplify notations, we set $F(x_1, x_2) = |x_1|^2 - \frac{1}{m_1} |x_1| |x_2|$ and $g(s, x; \theta) = \frac{1}{2} V_{xx}(s, x; \theta)$. Denoting $g^n_i(\theta) = g(i \Delta_n, X_i \Delta_n; \theta)$, we define

$$\bar{V}^n_T(\theta) = \Delta_n^{-1/2} (\bar{V}^n_T(\theta) - V_T^0(\theta)), \quad \text{where}
$$

$$V^0_T(\theta) = \Delta_n \sum_{i=1}^{[T/\Delta_n]-1} g^n_{i-1}(\theta) F \left( \Delta_n X_i / \Delta_n^{1/2}, \Delta_n X_{i+1} / \Delta_n^{1/2} \right),$$

$$V_T^0(\theta) = \int_0^T g(s, X_s; \theta) \rho_{\sigma_s}^2 (F) \, ds.$$

Note that $B^n_T(\theta) = V^n_T(\theta)$ so it is enough to consider the convergence of the process $(\bar{V}^n_T(\theta); \theta \in \Theta)$. Our approach replies on the following decomposition

$$\bar{V}^n_T(\theta) = \bar{V}^n_T(\theta) + \Delta_n^{-1/2} (V^n_T(\theta) - V^n_T(\theta)) - \Delta_n^{-1/2} (V_T^0(\theta) - V_T^0(\theta)),$$

where

$$\bar{V}^n_T(\theta) = \Delta_n \sum_{i=1}^{[T/\Delta_n]-1} g^n_i(\theta) F \left( \sigma_{i-1}^{(n)} \Delta_n, \sigma_{i+1}^{(n)} W / \Delta_n^{1/2} \right),$$

$$V_T^0(\theta) = \Delta_n \sum_{i=1}^{[T/\Delta_n]-1} g^n_{i-1}(\theta) \rho_{\sigma_{i-1}^{(n)}}^2 (F).$$

The last two terms on the right side of (21) are negligible as shown in the following
Proposition 1 Suppose that Assumptions (SK) and (V) hold and $X$ is continuous. We have
\[
\Delta_n^{-1/2} (V^n_T (\theta) - V_T^n (\theta)) = o_p (1) \\
\Delta_n^{-1/2} (V'_T (\theta) - V'_T (\theta)) = o_p (1)
\]
uniformly in $\theta \in \Theta$.

The proof of Proposition 1 requires two lemmas.

Lemma 4 Let $Y$ be an Itô semimartingale taking values in some bounded subset $E \subset \mathbb{R}^m$, $m \geq 1$, with the form
\[
Y_t = Y_0 + \int_0^t b^Y_s \, ds + \int_0^t \sigma^Y_s \, dW_s + M^Y_t,
\]
where $M^Y$ is a local martingale orthogonal to $W$ with bounded jumps, and $\langle M^Y, M^Y \rangle_t = \int_0^t \sigma^Y_s \, ds$, and where $b^Y_t, \sigma^Y_t$ are optional bounded processes. Let $(\Theta, d)$ be a totally bounded subspace of $\mathbb{R}^q, q \geq 1$ and $\{ f (\cdot; \theta) : \theta \in \Theta \}$ be a family of real-valued continuous functions on $[0, T] \times \mathcal{E}$. Suppose that for some constant $C > 0$ and some increasing continuous function $\Psi$ on $\mathbb{R}_+$ with $\Psi (0) = 0$, the following conditions hold:
(a) $\sup_{(s, y, \theta) \in [0, T] \times \mathcal{E}} (|f (s, y, \theta)| + \| \nabla_y f (s, y, \theta) \|) < \infty$;
(b) $\sup_{(s, y) \in [0, T] \times \mathcal{E}} \| \nabla_y f (s, y, \theta') - \nabla_y f (s, y, \theta) \| \leq C d (\theta, \theta')$, for all $\theta, \theta' \in \Theta$;
(c) $\sup_{\theta \in \Theta} | f (s', y'; \theta) - f (s, y; \theta) - \nabla_y f (s, y, \theta) (y' - y) | \leq C | s' - s | + \Psi (\| y' - y \|) \| y' - y \|$, for all $s, s' \in [0, T], y, y' \in \mathcal{E}$.

Then,
\[
\sup_{\theta \in \Theta} \Delta_n^{-1/2} \left| \int_0^T f (s, Y_s) \, ds - \Delta_n \sum_{i=1}^{[T/\Delta_n]} f ((i - 1) \Delta_n, Y_{(i-1)\Delta_n}) \right| = o_p (1).
\]

Proof. The local martingale $M^Y$ can be decomposed as
\[
M^Y_t = \int_0^t \tilde{\sigma}^Y_s \, d\tilde{W}_s + \delta' * (\mu - \nu),
\]
where $\delta'$ is a predictable function, and $\tilde{W}$ is a Wiener process orthogonal to $W$. Because $a^Y$ is bounded, $\sigma^Y$ and $\delta'$ can be chosen such that
\[
\| \tilde{\sigma}^Y_t (\omega) \| \leq K, \| \delta' (\omega, t, z) \| \leq K, \int_{\mathcal{E}} \| \delta' (\omega, t, z) \|^2 \lambda (dz) \leq K.
\]
For any $s < t$ in $[0, T]$, we have the following estimate
\[
E \left[ \| Y_t - Y_s \|^2 \right] \leq K E \left[ \left( \int_s^t b^Y_u \, du \right)^2 \right] + KE \left[ \left( \int_s^t \sigma^Y_u dW_u + M^Y_t - M^Y_s \right)^2 \right] \leq K | t - s |,
\]
where the second inequality follows from the Burkholder-Davis-Gundy inequality and the boundedness of the coefficients of $Y$. 
Let \( I(n,i) = ((i-1)\Delta_n, i\Delta_n] \). We have the following decomposition

\[
\Delta_n^{-1/2} \left( \int_0^t f(s, Y_s) ds - \Delta_n \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} f \left( (i-1) \Delta_n, Y_{(i-1)\Delta_n} \right) \right)
\]

\[
= \bar{\eta}_n(\theta) + \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \eta_{i}^n(\theta) + \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \eta_{i}^n(\theta).
\]

where

\[
\bar{\eta}_n(\theta) = \Delta_n^{-1/2} \int_{\lfloor t/\Delta_n \rfloor \Delta_n}^t f(u, Y_u; \theta) \, du,
\]

\[
\eta_i^n(\theta) = \Delta_n^{-1/2} \int_{I(n,i)} f(u, Y_u; \theta) - f \left( (i-1) \Delta_n, Y_{(i-1)\Delta_n}; \theta \right)
- \nabla_y f \left( (i-1) \Delta_n, Y_{(i-1)\Delta_n}; \theta \right) (Y_u - Y_{(i-1)\Delta_n}) \, du,
\]

\[
\eta_i^n(\theta) = \Delta_n^{-1/2} \nabla_y f \left( (i-1) \Delta_n, Y_{(i-1)\Delta_n}; \theta \right) \int_{I(n,i)} (Y_u - Y_{(i-1)\Delta_n}) \, du.
\]

By condition (a), we have \( \sup_{\theta \in \Theta} |\bar{\eta}_n(\theta)| \leq \Delta_n^{-1/2} \int_{\lfloor t/\Delta_n \rfloor \Delta_n}^t \sup_{\theta \in \Theta} |f(u, Y_u; \theta)| \, du \leq K\Delta_n^{1/2} \). Hence, \( \sup_{\theta \in \Theta} |\bar{\eta}_n(\theta)| = o_p(1) \).

By condition (c), we have,

\[
\sup_{\theta \in \Theta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\eta_i^n(\theta)|
\leq \Delta_n^{-1/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \int_{I(n,i)} \left[ C|u - (i-1) \Delta_n| + \Psi \left( \|Y_u - Y_{(i-1)\Delta_n}\| \right) \|Y_u - Y_{(i-1)\Delta_n}\| \right] \, du
\leq K\Delta_n^{1/2} + \Delta_n^{-1/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \int_{I(n,i)} \Psi \left( \|Y_u - Y_{(i-1)\Delta_n}\| \right) \|Y_u - Y_{(i-1)\Delta_n}\| \, du
\leq K\Delta_n^{1/2} + \Delta_n^{-1/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \int_{I(n,i)} \Psi (\varepsilon) \|Y_u - Y_{(i-1)\Delta_n}\| \, du
+ K\Delta_n^{-1/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \int_{I(n,i)} \frac{\|Y_u - Y_{(i-1)\Delta_n}\|^2}{\varepsilon} \, du.
\]

Combining (25) and (27), we have

\[
\mathbb{E} \left[ \sup_{\theta \in \Theta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\eta_i^n(\theta)| \right] \leq K\Delta_n^{1/2} + K\Psi (\varepsilon) + \frac{K\Delta_n^{1/2}}{\varepsilon}.
\]

By first letting \( \Delta_n \rightarrow \infty \) and then sending \( \varepsilon \rightarrow 0 \), we see that the left side of (28) goes to zero. Hence, \( \sup_{\theta \in \Theta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} |\eta_i^n(\theta)| = o_p(1) \).
Let \( Q_n (\theta) = \sum_{i=1}^{[t/\Delta_n]} \eta_i^n (\theta) \). For any \( \theta, \theta' \in \Theta \), we use the Cauchy-Schwarz inequality and condition (b) to derive

\[
|Q_n (\theta') - Q_n (\theta)| \\
\leq \Delta_n^{-1/2} \sum_{i=1}^{[t/\Delta_n]} \left\| \nabla_y f\left((i-1)\Delta_n, Y_{(i-1)\Delta_n}, \theta'\right) - \nabla_y f\left((i-1)\Delta_n, Y_{(i-1)\Delta_n}, \theta\right) \right\| \\
\times \int_{I(n,i)} \left\| Y_u - Y_{(i-1)\Delta_n} \right\| du \\
\leq K d (\theta, \theta') B_n,
\]

where

\[
B_n = \Delta_n^{-1/2} \sum_{i=1}^{[t/\Delta_n]} \int_{I(n,i)} \left\| Y_u - Y_{(i-1)\Delta_n} \right\| du.
\]

We use (25) and Jensen’s inequality to derive \( B_n = O_p (1) \). By Theorem 21.10 of Davidson (1994), we see that \( (Q_n (\theta) : \theta \in \Theta, n \geq 1) \) is stochastically equicontinuous.

Now, fix any \( \theta \in \Theta \). We shall show \( Q_n (\theta) = o_p (1) \). We suppress the dependence on \( \theta \). For example, we write \( Q_n \) and \( \eta_i^n \) instead of \( Q_n (\theta) \) and \( \eta_i^n (\theta) \). Let \( \eta (1)_i^n = \mathbb{E}[\eta_i^n | F_{(i-1)\Delta_n}] \) and \( \eta (2)_i^n = \eta_i^n - \eta (1)_i^n \). Setting \( Q (j)_n = \sum_{i=1}^{[t/\Delta_n]} \eta (j)_i^n \), \( j = 1, 2 \), we have \( Q_n = Q (1)_n + Q (2)_n \). By condition (a),

\[
|\eta (1)_i^n| = \left| \Delta_n^{-1/2} \nabla_y f\left((i-1)\Delta_n, Y_{(i-1)\Delta_n}\right) \int_{I(n,i)} \left( \int_u^{\Delta_n} b^y du \right) \right| du \\
\leq K \Delta_n^{3/2}.
\]

Thus \( |Q (1)_n| \leq K \Delta_n^{3/2} \to 0 \). By condition (b), the Cauchy-Schwarz inequality and (25), we have \( \mathbb{E}[\eta_i^n] \leq K \Delta_n^2 \). By construction, \( \eta (2)_i^n \) forms a martingale difference sequence, so

\[
\mathbb{E}\left[ Q (2)_n^2 \right] = \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}\left[ (\eta (2)_i^n)^2 \right] \leq K \Delta_n.
\]

Hence, \( Q (2)_n = o_p (1) \). We thus have \( Q_n (\theta) = o_p (1) \) for any \( \theta \in \Theta \). Since \( Q_n (\cdot) \) is stochastically equicontinuous, we have \( Q_n (\theta) = o_p (1) \) uniformly in \( \theta \in \Theta \).

We have shown that each of the three components on the right side of (26) is \( o_p (1) \) uniformly on \( \Theta \). The claim (23) readily follows. \( \square \)

**Lemma 5** Suppose that Assumptions (SK) holds and \( X \) is continuous. Let \((\Theta, d)\) be a totally bounded subspace of \( \mathbb{R}^k \) and \( \{ (Z_t (\theta))_{t \geq 0} : \theta \in \Theta \} \) be a family of bounded adapted processes. Suppose that there exists a finite constant \( C > 0 \) such that for any \( \theta, \theta' \in \Theta \), and \( t \in [0, T] \), \( |Z_t (\theta') - Z_t (\theta)| \leq C d (\theta, \theta') \). The variables

\[
\xi_n (\theta) = \Delta_n^{1/2} \sum_{i=1}^{[t/\Delta_n]-1} Z_{(i-1)\Delta_n} (\theta) \left[ F \left( \Delta_n^1 X/\Delta_n^{1/2}, \Delta_n^{n+1} X/\Delta_n^{1/2} \right) - F (\beta_i^1, \beta_i^2) \right]
\]

converge to zero in probability uniformly in \( \theta \in \Theta \).
Proof. It is enough to show two results: (1) \( \xi_n (\cdot) \) is stochastically equicontinuous; (2) \( \xi_n (\theta) = o_p (1) \) for any \( \theta \in \Theta \).

Step 1) In this step, we prove that \( \xi_n (\cdot) \) is stochastically equicontinuous. We start with some estimates. Let \( I (n, i) = ((i - 1)\Delta_n, i\Delta_n] \). Note that
\[
\mathbb{E} \left[ (\Delta_i^n X - \sigma_{(i-1)\Delta_n} \Delta_i^n W)^2 \right] \\
\leq K \mathbb{E} \left[ \left( \int_{I(n,i)} b_s ds \right)^2 \right] + K \mathbb{E} \left[ \left( \int_{I(n,i)} (\sigma_s - \sigma_{(i-1)\Delta_n}) dW_s \right)^2 \right] \\
\leq K \Delta_i^2 + K \mathbb{E} \left[ \int_{I(n,i)} (\sigma_s - \sigma_{(i-1)\Delta_n})^2 ds \right] \\
\leq K \Delta_i^2,
\]
where the first inequality follows from (7), the second inequality follows from \( |b_s| \leq K \) and the Burkholder-Davis-Gundy inequality and the last inequality follows from (25). It is obvious that \( \mathbb{E}[(\Delta_i^n X)^2] + \mathbb{E}[|\sigma_{(i-1)\Delta_n} \Delta_i^n W|^2] \leq K \Delta_i^2 \). Combining these estimates with the Cauchy-Schwarz inequality, we derive
\[
\mathbb{E} \left[ (\Delta_i^n X)^2 - (\sigma_{(i-1)\Delta_n} \Delta_i^n W)^2 \right] \\
\leq \left\{ \mathbb{E} \left[ (\Delta_i^n X - \sigma_{(i-1)\Delta_n} \Delta_i^n W)^2 \right] \right\}^{1/2} \left\{ \mathbb{E} \left[ (\Delta_i^n X + \sigma_{(i-1)\Delta_n} \Delta_i^n W)^2 \right] \right\}^{1/2} \\
\leq K \Delta_i^{3/2},
\]
and similarly
\[
\mathbb{E} \left[ |\Delta_i^n X| |\Delta_{i+1}^n X| - |\sigma_{(i-1)\Delta_n} \Delta_i^n W||\sigma_{(i-1)\Delta_n} \Delta_{i+1}^n W| \right] \\
\leq \mathbb{E} \left[ |\Delta_i^n X| |\Delta_{i+1}^n X - \sigma_{(i-1)\Delta_n} \Delta_{i+1}^n W| \right] \\
+ \mathbb{E} \left[ |\Delta_i^n X - \sigma_{(i-1)\Delta_n} \Delta_i^n W| |\sigma_{(i-1)\Delta_n} \Delta_{i+1}^n W| \right] \\
\leq K \Delta_i^{3/2}.
\]
Setting \( F_i^n = F(\Delta_i^n X/\Delta_i^{1/2}, \Delta_{i+1}^n X/\Delta_i^{1/2}) - F(\beta_{i,1}^n, \beta_{i,2}^n) \), we combine (31) and (32) to derive
\[
\mathbb{E}[|F_i^n|] \leq K \Delta_i^{1/2}.
\] (33)

For any \( \theta, \theta' \in \Theta \), we use the Lipschitz condition on \( Z_t (\theta) \) to derive \( |\xi_n (\theta') - \xi_n (\theta)| \leq K d (\theta, \theta') B_n \), where \( B_n = \Delta_i^{1/2} \sum_{i=1}^{[t/\Delta_n]} |F_i^n| \). Following (33), we have \( B_n = O_p (1) \). By Theorem 21.10 of Davidson (1994), we conclude that \( \{ \xi_n (\theta) : n \geq 1, \theta \in \Theta \} \) is stochastically equicontinuous.

Step 2) Fix any \( \theta \in \Theta \). We shall show that \( \xi_n (\theta) = o_p (1) \). We note that our function \( F \) satisfies the conditions of Lemma 11.2.7 in Jacod and Protter (2010). Our proof is a straightforward adaptation of the proof of Jacod and Protter (2010). Because the full proof is quite lengthy, we only emphasize the modification here. For the sake of notational simplicity, we suppress the dependence on \( \theta \). We adopt the
same notation as in Jacod and Protter (2010):

\[
\begin{align*}
X^n_i & = \left( \Delta^n_i X/\Delta^{1/2}_n, \Delta^n_{i+1} X/\Delta^{1/2}_n \right) \\
\beta^n_i & = (\beta^n_{i,1}, \beta^n_{i,2}) \\
\chi^n_i & = \Delta^{1/2}_n \left( F \left( X^n_i \right) - F \left( \beta^n_i \right) \right) \\
\chi^n_{i,j} & = \mathbb{E} \left[ \chi^n_i | \mathcal{F}_{(i-1)\Delta_n} \right], \\
\chi^n_{i,j} & = \chi^n_i - \chi^n_j.
\end{align*}
\]

It is enough to show that

\[
\begin{align*}
\sum_{i=1}^{[t/\Delta_n]-k+1} Z_{(i-1)\Delta_n} \chi^n_{i,j} & = o_p(1) \\
\sum_{i=1}^{[t/\Delta_n]-k+1} Z_{(i-1)\Delta_n} \chi^n_{i,j} & = o_p(1).
\end{align*}
\]

By construction, \( Z_{(i-1)\Delta_n} \chi^n_{i,j} \) and \( Z_{(j-1)\Delta_n} \chi^n_{i,j} \) are uncorrelated whenever \( |i-j| \geq 2 \). Since \( Z \) is bounded, we use the Cauchy-Schwarz inequality to derive

\[
\mathbb{E} \left[ \left( \sum_{i=1}^{[t/\Delta_n]-1} Z_{(i-1)\Delta_n} \chi^n_{i,j} \right)^2 \right] \leq K \mathbb{E} \left( \sum_{i=1}^{[t/\Delta_n]-1} (\chi^n_{i,j})^2 \right).
\]

It is shown in Jacod and Protter (2010) that the right side of (36) goes to zero, see (11.2.43) there. We thus have (34).

It remains to show (35). Setting

\[
\begin{align*}
Y^n_{i,0} & = (\beta^n_{i,1}, \beta^n_{i,2}) , \quad Y^n_{i,1} = \left( \Delta^n_i X/\Delta^{1/2}_n, \beta^n_{i,2} \right) , \quad Y^n_{i,2} = \left( \Delta^n_i X/\Delta^{1/2}_n, \Delta^n_{i+1} X/\Delta^{1/2}_n \right) \\
\chi^n_{i,j} & = \Delta^{1/2}_n \left[ F \left( Y^n_{i,j} \right) - F \left( Y^n_{i,j-1} \right) | \mathcal{F}_{(i-1)\Delta_n} \right], \quad j = 1, 2,
\end{align*}
\]

we have \( \chi^n_{i,j} = \sum_{j=1}^k \chi^n_{i,j} \). Hence, it is enough to show that for each \( j = 1, 2 \),

\[
\sum_{i=1}^{[t/\Delta_n]-1} Z_{(i-1)\Delta_n} \chi^n_{i,j} = o_p(1).
\]

Defining \( \delta'(r)_i^r, 1 \leq r \leq 7 \), in the same way as in (11.2.49) in Jacod and Protter (2010), we have \( \chi^n_{i,j} = \sum_{r=1}^7 \mathbb{E} \left[ \delta'(r)_i^r - \mathcal{F}_{(i-1)\Delta_n} \right] \). By (5.3.48) and (11.2.52) in Jacod and Protter (2010), we have

\[
\begin{align*}
\begin{cases}
\quad r = 4, 5, 6, 7 \Rightarrow \sum_{i=1}^{[t/\Delta_n]-k+1} \mathbb{E} \left[ \left| \delta'(r)_i^r \right| \right] & \rightarrow 0 \\
\quad r = 1, 2, 3 \Rightarrow \mathbb{E} \left[ \delta'(r)_i^r | \mathcal{F}_{(i-1)\Delta_n} \right] & = 0.
\end{cases}
\end{align*}
\]

Since \( Z \) is bounded, (37) readily follows from (38). This finishes the proof.

**Proof of Proposition 1.** We put \( Z_\theta(s) = g(s, X_s; \theta) \). Since \( g(s, x; \theta) = \frac{1}{2} V_{xx}(s, x; \theta) \), we readily see that \( Z \) satisfies the conditions in Lemma 5 under Assumptions (SK) and (V). Lemma 5 yields \( \Delta_n^{-1/2} (V^n_1 (\theta) - V^n_t (\theta)) = o_p(1) \) uniformly in \( \theta \).
We put \( Y = (X, \sigma) \). Under Assumption (SK), \( Y \) satisfies the conditions in Lemma 4. We also put \( f(s, (x, \sigma); \theta) = g(s, x; \theta) \rho_x^2(F) \). Noting that \( \rho_x^2(F) = c\sigma^4 \) for some constant \( c \), we verify that \( f \) satisfies the conditions in Lemma 4 under Assumptions (SK) and (V). We use Lemma 4 to derive \( \Delta_n^{-1/2} (V_1^\ell (\theta) - V_1^\ell (\theta)) = o_p(1) \) uniformly in \( \theta \in \Theta \). \( \square \)

We now study the stable convergence in law of the process \( (\tilde{V}_T^n (\theta) : \theta \in \Theta) \). We first consider the convergence of the marginals and then prove the stochastic equicontinuity of this process.

**Proposition 2** Suppose that Assumptions (SK) and (V) hold and \( X \) is continuous. Let \( d \in \mathbb{N} \) and \( \theta_1, \ldots, \theta_d \in \Theta \). The \( d \)-dimensional variables \( \left( \tilde{V}_T^n (\theta_1) \right)_{1 \leq 1 \leq d} \) converge stably in law to a \( d \)-dimensional variable \( (U (\theta_1))_{1 \leq 1 \leq d} \) defined on an extension of the space \((\Omega, \mathcal{F}, \mathbb{P})\), which conditionally on \( \mathcal{F} \), is centered Gaussian with covariance defined by

\[
\text{Cov} (U (\theta_1), U (\theta_1)) = C_1 \int_0^T g(s, X_s; \theta_1) g(s, X_s; \theta_1') \sigma_s^4 ds,
\]

where \( C_1 = \frac{\pi^2}{4} + \pi - 5 \).

**Proof.** For any \( a > 0 \), we set \( F_1 (a; x_1) = \mathbb{E} [F(x_1, aU)] \), where \( U \) is an \( \mathcal{N} (0, 1) \) variable. Explicitly, we have \( F_1 (a; x_1) = |x_1|^2 - \frac{a}{m_1} |x_1| \). By some straightforward manipulation, we decompose

\[
\tilde{V}_T^n (\theta) = R_T^n (\theta) + \sum_{i=2}^{[T/\Delta_n] - 1} \zeta (\theta)_i^n, \quad \text{where}
\]

\[
\zeta (\theta)_i^n = \Delta_n^{1/2} g_{i-1}^{n, 2} (\theta) \left[ F \left( \beta_i^{n, 1, 1}; \beta_i^{n, 1, 2} \right) - F_1 \left( \sigma_{(i-1)\Delta_n}; \beta_i^{n, 1, 1} \right) \right]
+ \Delta_n^{1/2} g_i^{n, 1} (\theta) F_1 \left( \sigma_{(i-1)\Delta_n}; \beta_i^{n, 1, 1} \right),
\]

\[
R_T^n (\theta) = \Delta_n^{1/2} g_{i-1}^{n, 2} (\theta) \left( F \left( \beta_i^{n, 1, 1}; \beta_i^{n, 1, 2} \right) - F_1 \left( \sigma_{(i-1)\Delta_n}; \beta_i^{n, 1, 1} \right) \right)
+ \Delta_n^{1/2} g_i^{n, 1} (\theta) F_1 \left( \sigma_{0}; \beta_i^{n, 1, 1} \right).
\]

It is easily seen that \( R_T^n (\theta) = o_p(1) \) for each \( \theta \in \Theta \). We note that \( (\zeta (\theta)_i^n, 1 \leq i \leq d)_{i \geq 1} \) is a triangular array of martingale differences by construction. By Theorem IX.7.28 of Jacod and Shiryaev (2003), in order to prove the claim of this proposition, it is sufficient to check the following conditions: for \( \theta, \theta' \in \Theta \), and any bounded martingale \( N \) orthogonal to \( W \),

\[
\sum_{i=2}^{[T/\Delta_n] - 1} \mathbb{E} \left[ \zeta (\theta)_i^n \zeta (\theta')_i^n | \mathcal{F}_{(i-1)\Delta_n} \right] \overset{p}{\to} C_1 \int_0^T g(s, X_s; \theta) g(s, X_s; \theta') \sigma_s^4 ds, \tag{39}
\]

\[
\sum_{i=2}^{[T/\Delta_n] - 1} \mathbb{E} \left[ |\zeta (\theta)_i^n|^4 | \mathcal{F}_{(i-1)\Delta_n} \right] \overset{p}{\to} 0, \tag{40}
\]

\[
\sum_{i=2}^{[T/\Delta_n] - 1} \mathbb{E} \left[ \Delta_i^{n, 2} W \zeta (\theta)_i^n | \mathcal{F}_{(i-1)\Delta_n} \right] \overset{p}{\to} 0, \tag{41}
\]

\[
\sum_{i=2}^{[T/\Delta_n] - 1} \mathbb{E} \left[ \Delta_i^{n, 2} N \zeta (\theta)_i^n | \mathcal{F}_{(i-1)\Delta_n} \right] \overset{p}{\to} 0. \tag{42}
\]

38
Denoting $Z^n_i = \Delta^n_i W / \Delta^{1/2}_n$, we can rewrite
\[
\zeta^n_i (\theta) = \Delta^{1/2}_n g_{i-2}^n (\theta) \sigma^n_{(i-2)\Delta_n} \left[ \frac{1}{m_1} |Z^n_{i-1}| - \frac{1}{m_1} |Z^n_{i-1}| |Z^n_i| \right] + \Delta^{1/2}_n g_{i-1}^n (\theta) \sigma^n_{(i-1)\Delta_n} \left[ |Z^n_i|^2 - \frac{1}{m_1} |Z^n_i| \right].
\]
Since $Z^n_i$ is $\mathcal{N}(0, 1)$ conditionally on $\mathcal{F}_{(i-1)\Delta_n}$, direct calculation gives
\[
\mathbb{E} \left[ \zeta^n_i (\theta) \zeta^n_{i^*} (\theta^*) | \mathcal{F}_{(i-1)\Delta_n} \right] = \sum_{j=1}^{4} \xi^n_{i,j}, \text{ where}
\]
\[
\xi^n_{1,j} = \Delta_n c_1 g_{i-2}^n (\theta) g_{i-2}^n (\theta^*) \sigma^n_{(i-2)\Delta_n} |Z^n_{i-1}|^2 / m_1,
\]
\[
\xi^n_{2,j} = \Delta_n c_2 g_{i-1}^n (\theta) g_{i-2}^n (\theta^*) \sigma^n_{(i-1)\Delta_n} \sigma^n_{(i-2)\Delta_n} |Z^n_{i-1}| / m_1,
\]
\[
\xi^n_{3,j} = \Delta_n c_3 g_{i-1}^n (\theta) g_{i-2}^n (\theta^*) \sigma^n_{(i-2)\Delta_n} \sigma^n_{(i-1)\Delta_n} |Z^n_{i-1}| / m_1,
\]
\[
\xi^n_{4,j} = \Delta_n c_4 g_{i-1}^n (\theta) g_{i-1}^n (\theta^*) \sigma^n_{(i-1)\Delta_n},
\]
and $c_1 = \frac{1}{m_1} - \frac{1}{m_1}$, $c_2 = c_3 = \frac{1}{m_1} - \frac{m_3}{m_1}$, $c_4 = m_4 + \frac{1}{m_1} - \frac{2m_3}{m_1}$.

Moreover, we have
\[
\mathbb{E} \left[ (g_{i-1}^n (\theta) \sigma^n_{(i-1)\Delta_n} - g_{i-2}^n (\theta) \sigma^n_{(i-2)\Delta_n})^2 \right] 
\leq K \mathbb{E} \left[ |\sigma_{(i-1)\Delta_n} - \sigma_{(i-2)\Delta_n}|^2 + \Delta_n^2 + |X_{(i-1)\Delta_n} - X_{(i-2)\Delta_n}|^2 \right] 
\leq K \Delta_n,
\]
where the first inequality follows from Assumptions (SK) and (V) and the second inequality is a standard estimate for Itô semimartingales with bounded coefficients, see (25). We set $\xi^n_{2,i} = \Delta_n c_2 g_{i-2}^n (\theta) g_{i-2}^n (\theta^*) \sigma^n_{(i-2)\Delta_n} |Z^n_{i-1}| / m_1$. By the Cauchy-Schwarz inequality and (44), we have
\[
\mathbb{E} \left| \xi^n_{2,i} - \xi^n_{1,i} \right| \leq K \Delta_n^{3/2}.
\]
Hence, $\sum_{i=2}^{[T/\Delta_n]} \xi^n_{2,i} = \sum_{i=2}^{[T/\Delta_n]} \xi^n_{2,i} + o_p(1)$. Using Lemma 1 with $Y_i = c_2 m_1 g(t, X_i; \theta) g(t, X_i; \theta^*) \sigma^n_t$ and taking $X$ in that lemma to be the Wiener process $W$, we derive $\sum_{i=2}^{[T/\Delta_n]} \xi^n_{2,i} = c_2 \int_0^T g(s, X_s; \theta) g(s, X_s; \theta^*) \sigma^4_s ds + o_p(1)$. We thus have
\[
\sum_{i=2}^{[T/\Delta_n]} \xi^n_{2,i} \xrightarrow{p} c_2 \int_0^T g(s, X_s; \theta) g(s, X_s; \theta^*) \sigma^4_s ds.
\]
Using the same argument, we can show for all $j = 1, 2, 3, 4$,
\[
\sum_{i=2}^{[T/\Delta_n]} \xi^n_{j,i} \xrightarrow{p} c_j \int_0^T g(s, X_s; \theta) g(s, X_s; \theta^*) \sigma^4_s ds.
\]
Because $C_4 = \sum_{j=1}^4 c_j$, we combine (43) and (45) to derive (39).
It is obvious that $\mathbb{E}[(\zeta(\theta))^n | \mathcal{F}_{(i-1)\Delta_n}] \leq K \Delta_n^2$, which implies (40). Noting that $\zeta(\theta)^n$ is an even function in $Z^i_n = \Delta^n W^i | \mathcal{D}_n^{1/2}$, we have $\mathbb{E}\Delta^n \zeta(g)^n | \mathcal{F}_{(i-1)\Delta_n} = 0$ and thus (41). Let $N$ be a bounded martingale orthogonal to $W$. We note that conditionally on $\mathcal{F}_{(i-1)\Delta_n}$, $\zeta^n$ is a martingale increment adapted to the filtration generated by the Wiener process $W_t - W_{(i-1)\Delta_n}$, $t \geq (i-1)\Delta_n$. Using the martingale representation theorem and the orthogonality between $N$ and $W$, we have $\mathbb{E}[\Delta^n N \zeta(g)^n | \mathcal{F}_{(i-1)\Delta_n}] = 0$ and thus (42). This finishes the proof.

We now show that the sequence of processes $(\mathbb{V}^n_T(\theta)_{\theta \in \Theta})_{n \geq 1}$ is stochastically equicontinuous for the metric $d$.

**Proposition 3** Suppose that Assumptions (SK) and (V) hold. Then $(\mathbb{V}^n_T(\theta)_{\theta \in \Theta})_{n \geq 1}$ is stochastically equicontinuous for each $d$. Below, we fix $j = 1$ or 2.

Let $G^n_i = \mathcal{F}_{(i+1)\Delta_n}$. Since $\mathbb{E}[F^n_i | \mathcal{F}_{(i-1)\Delta_n}] = 0$ and $F^n_i \in \mathcal{F}_{(i+1)\Delta_n}$, $(g^n_i(\theta) F^n_i, G^n_i)_{i \in I_n(j)}$ form an array of martingale differences. Hence,

$\mathbb{E}\left[(\xi^n_1(\theta') - \xi^n_1(\theta))^q\right] = \Delta^{2/q} \mathbb{E}\left[\left(\sum_{i \in I_n(j)} (g^n_{i-1}(\theta') - g^n_{i-1}(\theta)) F^n_i\right)^q\right]$

$\leq K \Delta^{2/q} \mathbb{E}\left[\left(\sum_{i \in I_n(j)} (g^n_{i-1}(\theta') - g^n_{i-1}(\theta))^2 (F^n_i)^2\right)^{q/2}\right]$,

$\leq K \Delta^{q/2} d(\theta, \theta') q \mathbb{E}\left[\left(\sum_{i \in I_n(j)} (F^n_i)^2\right)^{q/2}\right]$

$\leq K d(\theta, \theta') q$,

where the first inequality follows from the Burkholder-Davis-Gundy inequality, the second inequality follows from Assumption (V3), and the last inequality follows from Hölder’s inequality. Therefore, $\|\xi^n_1(\theta') - \xi^n_1(\theta)\|_q \leq K d(\theta, \theta')$. By Theorem 2.2.4 of van der Vaart and Wellner (1996), we have for any $\eta, \delta > 0$,

$\|\sup_{d(\theta, \theta') \leq \delta} \xi^n_1(\theta') - \xi^n_1(\theta)\|_q \leq K \int_0^\eta D(\varepsilon, \Theta, d)^{1/q} d\varepsilon + \delta D(\eta, \Theta, d)^{2/q}$,

where for any $\varepsilon > 0$, the packing number $D(\varepsilon, \Theta, d)$ is the maximum number of points in $\Theta$ such that each pair of points have distance strictly greater than $\varepsilon$. 

40
By Assumption (V4) and the relation between covering and packing numbers, the packing number
\( D(\varepsilon, \Theta, d) \) is bounded above by \( C\varepsilon^{-\kappa} \) for some finite constant \( C \). Since \( q > \kappa \),
\[
\int_0^1 D(\varepsilon, \Theta, d)^{1/q} d\varepsilon \leq K \int_0^1 \varepsilon^{-\kappa/q} d\varepsilon < \infty.
\]
Therefore, by the dominated convergence theorem, \( \lim_{\eta \to 0} \int_0^\eta D(\varepsilon, \Theta, d)^{1/q} d\varepsilon = 0 \). By first taking \( \eta \) small and then choosing \( \delta \) small, the right side of (46) can be made arbitrarily small. Therefore, \( \lim_{\delta \to 0} ||\sup_{d(\theta, \theta') \leq \delta} (\xi^j_n(\theta') - \xi^j_n(\theta))||_q = 0 \). The stochastic equicontinuity of \((\xi^j_n(\theta_{\inTheta})_{n \geq 1} \) readily follows.

**Proof of Theorems 3 and 5.** By localization, we assume that Assumption (SK) holds. By Lemma 3 and Proposition 1, \( B^n_T(\cdot) \) has the same asymptotic distribution as \( \tilde{V}^n_T(\cdot) \). Proposition 2 shows the stable convergence in law of \( \tilde{V}^n_T(\cdot) \) on finite dimensions and Proposition 3 establishes its stochastic equicontinuity. Combining these two results, we see that \( \tilde{V}^n_T(\cdot) \) converges stably in law of the limiting process described in Theorem 5. Theorem 3 follows as a corollary.

### 9.4 Proof of Theorems 4 and 6

We only need to prove Theorem 6 which includes Theorem 4 as a special case. We first prove a lemma regarding the weak convergence of Gaussian processes. For any centered Gaussian process \( Z(\theta)_{\inTheta} \), its standard deviation semimetric is defined as \( \sigma(\theta, \theta') = \{E[(Z(\theta) - Z(\theta'))^2]\}^{1/2} \), for \( \theta, \theta' \in \Theta \).

**Lemma 6** Let \((\Theta, d)\) be a semimetric space. Let \( Z_n(\theta)_{\inTheta} \) be a separable centered Gaussian process defined on \((\Omega_n, \mathcal{F}_n, \mathbb{P}_n)\) with (deterministic) standard deviation semimetric \( \sigma_n(\theta, \theta') \) and \( Z(\theta)_{\inTheta} \) be a centered Gaussian process defined on \((\Omega, \mathcal{F}, \mathbb{P})\). Suppose (a) the covariance function of \( Z_n(\cdot) \) converges to the covariance function of \( Z(\cdot) \) pointwisely; (b) for some constant \( C > 0 \), \( \sigma_n(\theta, \theta') \leq C d(\theta, \theta') \) for all \( \theta, \theta' \in \Theta \); (c) for some \( \eta > 0 \), \( \int_0^\eta \sqrt{\log N(\varepsilon, \Theta, d)} d\varepsilon < \infty \). Then \( Z_n(\cdot) \) converges weakly to \( Z(\cdot) \).

**Proof.** By condition (a), \( Z_n(\cdot) \) converges weakly to \( Z(\cdot) \) on finite dimensions. Condition (c) implies that \((\Theta, d)\) is totally bounded. It remains to show that \( Z_n(\cdot) \) is stochastically equicontinuous for the metric \( d \). Since \( Z_n(\cdot) \) is a Gaussian process, it is sub-Gaussian for the semimetric \( \sigma_n(\cdot, \cdot) \). By Corollary 2.2.8 of van der Vaart and Wellner (1996), we have for some universal constant \( K \),
\[
\mathbb{E}_n\left( \sup_{\sigma_n(s,t) \leq \delta} |Z_n(s) - Z_n(t)| \right) \leq K \int_0^\delta \sqrt{\log N(\varepsilon/2, \Theta, \sigma_n)} d\varepsilon.
\]
By condition (b), we have \( \sup_{d(s,t) \leq \delta/C} |Z_n(s) - Z_n(t)| \leq \sup_{\sigma_n(s,t) \leq \delta} |Z_n(s) - Z_n(t)| \) and \( \log N(\varepsilon/2, \Theta, \sigma_n) \leq \log N(\varepsilon/2C, \Theta, d) \). Combining these estimates and using a change of variable, we derive that
\[
\mathbb{E}_n\left( \sup_{d(s,t) \leq \delta} |Z_n(s) - Z_n(t)| \right) \leq 2CK \int_0^{\delta/2} \sqrt{\log N(\varepsilon, \Theta, d)} d\varepsilon.
\]
By condition (c), the right side of (47) goes to zero as \( \delta \to 0 \). This convergence, combined with Markov’s inequality, yields the stochastic equicontinuity of \( Z_n(\cdot) \). \( \square \)
Proof of Theorem 6. Step 1) In this step, we show that $\hat{\Sigma}_n^T (1; \theta, \theta') = \Sigma_T (\theta, \theta') + o_p(1)$ uniformly in $\theta, \theta' \in \Theta$. By localization, we can suppose that Assumption (SH-2) holds. In particular, $X$ only takes values in a bounded set $K \subset D$. Let $\theta, \theta' \in \Theta$. Setting $Y_t = V_{xx} (t, X_t; \theta) V_{xx} (t, X_t; \theta')$, and $F (x_1, x_2, x_3) = \prod_{j=1}^3 |x_j|^{4/3}$, we verify the conditions in Lemma 1 and derive $\hat{\Sigma}_n^T (1; \theta, \theta') = \Sigma_T (\theta, \theta') + o_p(1)$ pointwise. We equip the product space $\Theta^2$ with the product metric $d_2((\theta_1, \theta_2), (\theta'_1, \theta'_2)) = \sum_{j=1}^2 d(\theta_j, \theta'_j)$. Since $(\Theta, d)$ is totally bounded, $(\Theta^2, d_2)$ is also totally bounded. It remains to establish that $\hat{\Sigma}_n^T (1; \cdot, \cdot)$ is stochastically equicontinuous for $d_2$. Under Assumption (V), we have

$$\left| \hat{\Sigma}_n^T (1; \theta_1, \theta_2) - \hat{\Sigma}_n^T (1; \theta'_1, \theta'_2) \right| \leq K d_2 \left( (\theta_1, \theta_2), (\theta'_1, \theta'_2) \right) B_n,$$

where $B_n = \frac{1}{\Delta_n} \sum_{i=1}^{[T/\Delta_n]-2} \prod_{j=1}^3 |\Delta_n^{t+j-1}|^{4/3}$. Applying Lemma 1 again but with $Y_t \equiv 1$, we have $B_n = O_p(1)$. The stochastic equicontinuity of $\hat{\Sigma}_n^T (1; \cdot, \cdot)$ follows from Theorem 21.10 of Davidson (1994).

Step 2) In this step, we show that $\hat{\Sigma}_n^T (2; \theta, \theta') = \Sigma_T (\theta, \theta') + o_p(1)$ uniformly in $\theta, \theta' \in \Theta$, so we have Assumption (H-r), $r < 2$ and $\bar{w} \in \left( \frac{1}{2}, \frac{1}{2} \right)$. By localization, we can suppose that Assumption (SH-r) holds. Setting $Y_t = V_{xx} (t, X_t; \theta) V_{xx} (t, X_t; \theta')$ and $F (x) = |x|^4$, Lemma 2 yields $\hat{\Sigma}_n^T (2; \theta, \theta') = \Sigma_T (\theta, \theta') + o_p(1)$ pointwise. By using a similar argument as in the previous step, we can show that $\hat{\Sigma}_n^T (2; \cdot, \cdot)$ is stochastically equicontinuous for the product metric $d_2$. The uniform convergence of $\hat{\Sigma}_n^T (2; \cdot, \cdot)$ towards $\Sigma_T (\cdot, \cdot)$ readily follows.

Step 3) In this step, we show that almost surely, the $F$-conditional distribution of $G_T$ and $G_T'$ are absolutely continuous. We first consider $G_T'$. For $\theta \in \Theta$, denote by $h_\theta$ the coordinate projection mapping, that is, $h_\theta : x \mapsto x(\theta); C(\Theta) \mapsto \mathbb{R}$ where $C(\Theta)$ is the collection of continuous functions on $\Theta$. Let $H = \{ h = h_\theta \text{ or } -h_\theta : \theta \in \Theta \}$. Consider the functional $f (x) = \sup_{h \in H} h (x), x \in C(\Theta)$. Then $G_T'$ can be represented as $f (U)$ and almost surely, $G_T' = h (U)$ for some $h \in H$. By Assumption (V), almost surely, $h (U)$ has strict positive variance for every $h \in H$. Using Theorem 11.4 of Davydov, Lifshits, and Smorodina (1998), we derive that the $F$-conditional distribution of $G_T'$ is absolutely continuous on $\mathbb{R}$. For $G_T$, we use the same proof but set $H = \{ h_\theta : \theta \in \Theta \}$ instead.

Step 4) In this step, we show that almost surely, the mappings $q \mapsto \kappa_q$ and $q \mapsto \kappa_q'$ are continuous on $(0, 1/2]$. We first consider $\kappa_q$. It amounts to showing that almost surely, the $F$-conditional distribution function of $G_T$ is strictly increasing on $[\kappa_{1/2}, \kappa_0]$, recalling that $\kappa_q$ is the $1-q$ quantile of $G_T$ conditional on $\mathcal{F}$. Let $\left( \hat{\omega}, \hat{\mathcal{F}}_n, \hat{\mathbb{P}}_n \right)$ be the extension space on which the limiting process $U (\cdot)$ is defined. Under Assumptions (SH-2) and (V), $\sup_{\theta \in \Theta} \mathbb{E} \left[ U (\theta)^2 \right] = \sup_{\theta \in \Theta} \mathbb{E} \left[ \sigma^2 \right] \leq K$. Because of this estimate and the continuity result established in the previous step, Theorem 11.6 of Davydov, Lifshits, and Smorodina (1998) yields that on almost every path, on the interval $[\kappa_{1/2}, \kappa_0]$, the derivative (possibly one-sided) of the $F$-conditional distribution function of $G_T$ is bounded below by a strictly positive number, which implies that conditional distribution function is strictly increasing on this interval. The claim involving $\kappa_q'$ is proven in exactly the same way. Without loss of generality, below, we assume that $q \mapsto \kappa_q$ and $q \mapsto \kappa_q'$ are continuous on $(0, 1/2]$ on all paths.

Step 5) In this step, we show the second claim of Theorem 6, i.e., $\tilde{\kappa}_n^q = \kappa_q + o_p(1)$ and $\tilde{\kappa}_n^q = \kappa_q + o_p(1)$ for $q \in (0, 1/2]$. We denote by $\left( \hat{\Omega}_n, \hat{\mathcal{F}}_n, \hat{\mathbb{P}}_n \right)$ the extension on which the process $U_n (\cdot)$ is defined. We
represent the extentions of the original probability space explicitly as follows
\[
\tilde{\Omega} = \Omega \times \Omega', \quad \tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \quad \tilde{\mathbb{P}}(d\omega, d\omega') = \mathbb{P}(d\omega) \mathbb{Q}_{\omega}(d\omega')
\]
\[
\tilde{\Omega}_n = \Omega \times \Omega'_n, \quad \tilde{\mathcal{F}}_n = \mathcal{F} \otimes \mathcal{F}'_n, \quad \tilde{\mathbb{P}}_n(d\omega, d\omega') = \mathbb{P}(d\omega) \mathbb{Q}_{n\omega}(d\omega')
\]
for some auxiliary measurable spaces \((\Omega', \mathcal{F}')\) and \((\Omega'_n, \mathcal{F}'_n)\) and transition probabilities \(\mathbb{Q}_{\omega}(d\omega')\) and \(\mathbb{Q}_{n\omega}(d\omega')\).

Denote by \(\sigma_n\) the \(\mathcal{F}\)-conditional standard probability deviation semimetric of \(U_n(\cdot)\). We have
\[
\sigma_n(\theta, \theta') = \begin{cases} 
\left( \frac{k}{m^{3/3} \Delta_n} \sum_{i=1}^{[T/\Delta_n]} -2 (g^n_{i-1}(\theta) - g^n_{i-1}(\theta'))^2 \prod_{j=1}^{3} |\Delta_{i+j-1}X_i|^{4/3} \right)^{1/2} & \text{if } \hat{\Sigma}_T^n(\cdot, \cdot) = \hat{\Sigma}_T^n(1; \cdot, \cdot) \\
\left( \frac{k}{m^n \Delta_n} \sum_{i=1}^{[T/\Delta_n]} (g^n_{i-1}(\theta) - g^n_{i-1}(\theta'))^2 (\Delta^n_i X_i)^4 \prod_{j=1}^{4} \{\Delta^n_j X \leq u_n\} \right)^{1/2} & \text{if } \hat{\Sigma}_T^n(\cdot, \cdot) = \hat{\Sigma}_T^n(2; \cdot, \cdot)
\end{cases}
\]
where we recall the notion \(g^n_i(\theta) = V_{x, x}(i\Delta_n, X, \Delta_n; \theta)\). By Assumption (V), for \(K\) defined in step 1, there exists some \(C > 0\) such that \(|g^n_{i-1}(\theta) - g^n_{i-1}(\theta')| \leq Cd(\theta, \theta')\). Letting
\[
B'_n = \begin{cases} 
C \left( \frac{k}{m^{3/3} \Delta_n} \sum_{i=1}^{[T/\Delta_n]} -2 \prod_{j=1}^{3} |\Delta_{i+j-1}X_i|^{4/3} \right)^{1/2} & \text{if } \hat{\Sigma}_T^n(\cdot, \cdot) = \hat{\Sigma}_T^n(1; \cdot, \cdot) \\
C \left( \frac{k}{m^n \Delta_n} \sum_{i=1}^{[T/\Delta_n]} (\Delta^n_i X_i)^4 \prod_{j=1}^{4} \{\Delta^n_j X \leq u_n\} \right)^{1/2} & \text{if } \hat{\Sigma}_T^n(\cdot, \cdot) = \hat{\Sigma}_T^n(2; \cdot, \cdot)
\end{cases}
\]
we have \(\sigma_n(\theta, \theta') \leq B'_n d(\theta, \theta')\). By Lemmas 1 and 2, we have \(B'_n = O_p(1)\) for \(j = 1, 2\).

Fix any \(m > 0\) and let \(\Omega_{n,m} = \{B'_n \leq m\}\). Let \(N_1\) be any subsequence of \(\mathbb{N}\). Since \(\hat{\Sigma}_n(\cdot, \cdot) = \Sigma_T(\cdot, \cdot) + o_p(1)\) uniformly, there exists a further subsequence \(N_2 \subset N_1\) such that \(\hat{\Sigma}_n(\cdot, \cdot) \Rightarrow \Sigma_T(\cdot, \cdot)\) uniformly on a \(P\)-full event \(\Omega^*\), where \(\Rightarrow\) indicates the convergence is along the subsequence \(N_2\). Now, fix any \(\omega \in \Omega_{n,m} \cap \Omega^*\). For such \(\omega\), we have (i) \(U_n(\cdot)\) and \(U(\cdot)\) are centered Gaussian processes under \(\mathbb{Q}_{n\omega}\) and \(\mathbb{Q}_{\omega}\) respectively; (ii) \(\hat{\Sigma}_n(\cdot, \cdot) \Rightarrow_{\mathbb{Q}_{n\omega}} \Sigma_T(\cdot, \cdot)\) uniformly as deterministic functions; (iii) \(\sigma_n(\theta, \theta') \leq md(\theta, \theta')\). By Assumption (V), we have \(\int_0^1 \sqrt{\log N(\varepsilon, \Theta, \Theta)d\varepsilon < \infty}\). Thus, we have verified the conditions of Lemma 6, which yields the weak convergence of \(U_n(\cdot) \Rightarrow_{\mathbb{Q}_{n\omega}} U(\cdot)\), where \(\Rightarrow_{\mathbb{Q}_{n\omega}}\) indicates the weak convergence along \(N_2\) under transition probabilities \(\mathbb{Q}_{n\omega}\) for each \(r\) at which the mapping \(r \mapsto \Sigma_T(\cdot, \cdot)\) is continuous. By the continuous mapping theorem, \(\hat{\Sigma}_n(\cdot, \cdot) \Rightarrow_{\mathbb{Q}_{n\omega}} \Sigma_T(\cdot, \cdot)\). Combining this convergence with Lemma 21.2 of van der Vaart (1998), we have \(\hat{k}_n(\omega) \Rightarrow_{\mathbb{Q}_{n\omega}} \kappa_T(\omega)\) for each \(r\) at which the mapping \(r \mapsto \kappa_T(\omega)\) is continuous. So \(\hat{k}_n^q(\omega) \Rightarrow_{\mathbb{Q}_{n\omega}} \kappa^q(\omega)\) for all \(q \in (0, 1/2]\). Because \(P(\Omega^*) = 1\), we have \((\hat{k}_n^q - \kappa^q) 1_{\Omega_{n,m}} \Rightarrow_{\mathbb{Q}_{n\omega}} 0\) almost surely. Now, we have shown that for any subsequence of \((\hat{k}_n^q - \kappa^q) 1_{\Omega_{n,m}}\), we can extract a further subsequence which converges to zero almost surely. Therefore, \((\hat{k}_n^q - \kappa^q) 1_{\Omega_{n,m}} = o_p(1)\) for each \(m > 0\). Since \(B'_n = O_p(1)\), by taking \(m\) sufficiently large, \(\Omega_{n,m}^c\) bears arbitrarily small probability uniformly in \(n\). Therefore, \(\hat{k}_n^q = \kappa^q + o_p(1)\) for each \(q \in (0, 1/2]\). Similarly, we have \(\hat{k}_n^q = \kappa^q + o_p(1)\).

Step 6) Combining Corollary 4 with the second claim of Theorem 6, we derive the last claim. □
<table>
<thead>
<tr>
<th></th>
<th>$\Delta = 1$ minute</th>
<th></th>
<th>$\Delta = 5$ minutes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>level = 5%</td>
<td>level = 10%</td>
<td>level = 5%</td>
<td>level = 10%</td>
</tr>
<tr>
<td>Bipower</td>
<td>4.7</td>
<td>9.4</td>
<td>5.3</td>
<td>10.5</td>
</tr>
<tr>
<td>SUP</td>
<td>5.2</td>
<td>10.5</td>
<td>6.5</td>
<td>13.1</td>
</tr>
<tr>
<td>ATM</td>
<td>4.8</td>
<td>9.3</td>
<td>5.6</td>
<td>10.6</td>
</tr>
<tr>
<td>ITM</td>
<td>4.3</td>
<td>9.1</td>
<td>5.1</td>
<td>9.8</td>
</tr>
<tr>
<td>OTM</td>
<td>4.8</td>
<td>8.9</td>
<td>4.6</td>
<td>9.8</td>
</tr>
</tbody>
</table>

Table 1: Finite-sample sizes (%) of tests.
<table>
<thead>
<tr>
<th>Jump Time ($\tau$)</th>
<th>Jump Size ($m$)</th>
<th>Bipower</th>
<th>SUP</th>
<th>ATM</th>
<th>ITM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.5</td>
<td>14.3</td>
<td>11.6</td>
<td>14.7</td>
<td>15.4</td>
<td>15.2</td>
</tr>
<tr>
<td></td>
<td>-2.5</td>
<td>13.9</td>
<td>11.4</td>
<td>14.1</td>
<td>15.3</td>
<td>15.2</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>83.5</td>
<td>74.8</td>
<td>81.5</td>
<td>75.2</td>
<td>71.0</td>
</tr>
<tr>
<td></td>
<td>-5.0</td>
<td>84.3</td>
<td>74.3</td>
<td>82.7</td>
<td>68.4</td>
<td>78.4</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>99.9</td>
<td>99.9</td>
<td>99.9</td>
<td>95.7</td>
<td>92.6</td>
</tr>
<tr>
<td></td>
<td>-7.5</td>
<td>100</td>
<td>100</td>
<td>99.9</td>
<td>88.7</td>
<td>96.1</td>
</tr>
<tr>
<td>2.5</td>
<td>2.5</td>
<td>14.2</td>
<td>13.6</td>
<td>15.7</td>
<td>9.2</td>
<td>11.9</td>
</tr>
<tr>
<td></td>
<td>-2.5</td>
<td>14.0</td>
<td>13.6</td>
<td>15.8</td>
<td>11.2</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>84.3</td>
<td>84.8</td>
<td>87.7</td>
<td>38.1</td>
<td>58.1</td>
</tr>
<tr>
<td></td>
<td>-5.0</td>
<td>84.7</td>
<td>84.4</td>
<td>87.9</td>
<td>49.2</td>
<td>47.4</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>66.9</td>
<td>90.1</td>
</tr>
<tr>
<td></td>
<td>-7.5</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>82.0</td>
<td>80.4</td>
</tr>
<tr>
<td>4</td>
<td>2.5</td>
<td>13.8</td>
<td>18.8</td>
<td>20.1</td>
<td>4.9</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>-2.5</td>
<td>14.2</td>
<td>19.1</td>
<td>20.9</td>
<td>5.4</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>84.7</td>
<td>94.9</td>
<td>95.3</td>
<td>7.3</td>
<td>20.8</td>
</tr>
<tr>
<td></td>
<td>-5.0</td>
<td>84.3</td>
<td>94.7</td>
<td>95.2</td>
<td>15.2</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>11.2</td>
<td>46.2</td>
</tr>
<tr>
<td></td>
<td>-7.5</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>32.5</td>
<td>17.0</td>
</tr>
</tbody>
</table>

Table 2: Monte Carlo rejection rates (%) of tests under the presence of jumps for 1-minute data. VC tests with strikes 22.5, 25 and 27.5 are labelled as ITM, ATM and OTM, respectively.
<table>
<thead>
<tr>
<th>Jump Time ($\tau$)</th>
<th>Jump Size ($m$)</th>
<th>Bipower</th>
<th>SUP</th>
<th>ATM</th>
<th>ITM</th>
<th>OTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.5</td>
<td>6.7</td>
<td>7.6</td>
<td>6.7</td>
<td>5.6</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>-2.5</td>
<td>6.8</td>
<td>7.6</td>
<td>7.0</td>
<td>6.9</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>24.3</td>
<td>20.4</td>
<td>24.1</td>
<td>19.8</td>
<td>23.7</td>
</tr>
<tr>
<td></td>
<td>-5.0</td>
<td>23.2</td>
<td>19.9</td>
<td>23.4</td>
<td>23.4</td>
<td>20.4</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>75.4</td>
<td>67.5</td>
<td>74.5</td>
<td>64.4</td>
<td>63.1</td>
</tr>
<tr>
<td></td>
<td>-7.5</td>
<td>75.6</td>
<td>66.4</td>
<td>74.2</td>
<td>60.9</td>
<td>68.8</td>
</tr>
<tr>
<td>2.5</td>
<td>2.5</td>
<td>6.7</td>
<td>8.1</td>
<td>7.3</td>
<td>5.0</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>-2.5</td>
<td>6.5</td>
<td>7.6</td>
<td>7.1</td>
<td>6.0</td>
<td>5.6</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>24.1</td>
<td>24.8</td>
<td>26.7</td>
<td>10.9</td>
<td>19.1</td>
</tr>
<tr>
<td></td>
<td>-5.0</td>
<td>23.7</td>
<td>24.6</td>
<td>25.9</td>
<td>17.0</td>
<td>12.7</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>75.6</td>
<td>76.0</td>
<td>79.0</td>
<td>29.5</td>
<td>55.3</td>
</tr>
<tr>
<td></td>
<td>-7.5</td>
<td>75.7</td>
<td>76.2</td>
<td>79.2</td>
<td>47.1</td>
<td>37.4</td>
</tr>
<tr>
<td>4</td>
<td>2.5</td>
<td>6.8</td>
<td>8.4</td>
<td>8.2</td>
<td>4.9</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>-2.5</td>
<td>6.7</td>
<td>8.4</td>
<td>8.2</td>
<td>5.6</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>24.1</td>
<td>31.6</td>
<td>31.7</td>
<td>5.5</td>
<td>11.1</td>
</tr>
<tr>
<td></td>
<td>-5.0</td>
<td>24.2</td>
<td>32.5</td>
<td>32.7</td>
<td>9.1</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>75.8</td>
<td>88.8</td>
<td>88.0</td>
<td>7.0</td>
<td>24.3</td>
</tr>
<tr>
<td></td>
<td>-7.5</td>
<td>75.3</td>
<td>88.2</td>
<td>88.4</td>
<td>18.4</td>
<td>8.9</td>
</tr>
</tbody>
</table>

Table 3: Monte Carlo rejection rates (%) of tests under the presence of jumps for 5-minute data. VC tests with strikes 22.5, 25 and 27.5 are labelled as ITM, ATM and OTM, respectively.
Table 4: Rejection rates of the ATM test, the SUP test and the bipower test at 5% nominal level for all 1590 stock-weeks.

<table>
<thead>
<tr>
<th></th>
<th>Overnight returns excluded</th>
<th>Overnight returns included</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>41%</td>
<td>86%</td>
</tr>
<tr>
<td>SUP</td>
<td>46%</td>
<td>89%</td>
</tr>
<tr>
<td>Bipower</td>
<td>45%</td>
<td>90%</td>
</tr>
</tbody>
</table>
Table 5: Empirical percentiles of realized volatility errors and realized jump errors for ATM options. For each stock-week, hedging errors are normalized with respect to the model price of the at-the-money call option (Panel A) and the option vega (Panel B) in percentage terms.
Table 6: Sensitivity analysis of realized hedging errors with respect to $v_{BS}$ for ATM options. We express the realized hedging errors in terms of implied volatility in percentage terms by normalizing them with respect to the option vega. Column 1 shows the sample median of the absolute values of the sensitivity measures (MAS). Column 2 shows the sample standard deviations (SD) of the realized jump error.

<table>
<thead>
<tr>
<th></th>
<th>MAS (%)</th>
<th>SD (%)</th>
<th>MAS/SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realized intraday jump error</td>
<td>0.02</td>
<td>2.06</td>
<td>0.01</td>
</tr>
<tr>
<td>Realized jump error</td>
<td>0.05</td>
<td>8.15</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Figure 1: Finite-sample sizes of 5% level VC tests.
Figure 2: Comparison of rejection rates for various jump times in simulation. Data are sampled at 1 minute (left) and 5 minutes (right). From top to bottom, the jump size is $m \times 25\sqrt{3} \times 5$ minutes with $m = 2.5, 5$ and 7.5. Legend: bipower (dashed line), SUP (square), ATM (circle), ITM (X) and OTM (triangle).
Figure 3: Monte Carlo rejection rates of 5% level VC tests. Data are sampled at every 5 minutes. The jump size parameter $m = 5.$
Figure 4: Monte Carlo rejection rates of 5% level VC tests. Data are sampled at every 5 minutes. The jump time is the middle of Wednesday ($\tau = 2.5$). The jump is positive (solid line) or negative (dashed line). From bottom to top, the jump magnitude is given by $|m| = 2.5$, 5, and 7.5.
Figure 5: Case study for MSFT. We consider two weeks in 2008: week-A beginning August 11th (left) and week-B beginning November 10th (right). In the top panel, we plot the time series of the stock price and highlight the overnight returns (circle). In the bottom panel, we plot the realized jump error profile for weekly call options versus the moneyness when overnight returns are included (solid line) and excluded (dashed line). The critical values of the VC tests (dotted line) and the uniform acceptance region of the SUP test (shaded area) is based on the truncation variance estimator. All tests are implemented at 5% nominal level.
Figure 6: Empirical rejection rate of 5% level VC tests. The volatility parameter in the Black-Scholes formula is set to be $v_{BS} = \eta \hat{v}$, where $\hat{v}$ is the estimated weekly average volatility and $\eta = 0.75$ (dashed line), 1 (solid line) and 1.25 (dash-dotted line). The adjusted moneyness is defined as $\eta^{-1}(K/X_0 - 1) + 1$, where $K$ is the strike price and $X_0$ is the weekly opening stock price.