High-frequency returns, jumps and the mixture of normals hypothesis

Jeff Fleming *, Bradley S. Paye

Jones Graduate School of Business, Rice University, 77005 Houston, TX, United States

A R T I C L E   I N F O

A B S T R A C T

Previous empirical studies find both evidence of jumps in asset prices and that returns standardized by 'realized volatility' are approximately standard normal. These findings appear to be contradictory. Using a sample of high-frequency returns for 20 heavily traded US stocks, we show how microstructure noise distorts the standard deviation and kurtosis of returns normalized using realized variance. When returns are standardized using a recently developed realized kernel estimator, the resulting series is clearly platykurtotic and the standard normal distribution is soundly rejected. Moreover, daily returns standardized using realized bipower variation, an estimator for integrated variance that is robust to the presence of jumps, are more consistent with the standard normal distribution. These results suggest that there is no empirical contradiction: jumps should be included in stock price models.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Asset prices are frequently modeled as evolving according to a continuous diffusive process. This class of models is both tractable and flexible. By superimposing multiple stochastic volatility diffusive processes, such models can accommodate many of the observed features of asset returns, such as time variation in both conditional volatility and conditional kurtosis.[1] Nevertheless, the assumption that prices follow a continuous sample path is restrictive. It is certainly plausible that prices exhibit occasional jumps, possibly related to the release of firm-specific or macroeconomic news. Such reasoning suggests that asset prices may be more appropriately modeled as jump diffusions. However, this entails greater costs in that jump diffusion models are more difficult to estimate and may include additional sources of risk that are difficult to price empirically using cross-sectional data.

Given this tradeoff, assessing the empirical evidence regarding the existence and relevance of jumps in asset prices is an important area of current research. Andersen et al. (2002) and Chernov et al. (2003) find evidence of jumps based on parametric estimation of jump diffusion models. More recently, Andersen et al. (2007a), Huang and Tauchen (2005), Jiang and Oomen (2007), Tauchen and Zhou (2011) and Barndorff-Nielsen and Shephard (2006) provide evidence of jumps using nonparametric techniques that exploit the information in high-frequency intraday asset returns.

A number of other studies observe that daily returns are nearly Gaussian when standardized by 'realized volatility' constructed using high-frequency intraday returns. For example, Andersen et al. (2001a) examine the unconditional distribution of daily returns for individual stocks standardized by realized volatility and find that "the close approximation afforded by the normal reference density is striking"[2] (pp. 51–52). As noted by Andersen et al. (2003), a standard normal distribution for standardized returns is consistent with a simple continuous diffusion model for asset prices that omits the possibility of occasional jumps.

The finding that daily returns standardized by realized volatility are approximately Gaussian seems inconsistent with the emerging body of literature that finds evidence of a significant jump component in total asset return volatility. This paper resolves the apparent incongruity. We begin by showing how microstructure noise alters the properties of daily returns standardized using the traditional realized variance estimator. Under the pure diffusion null hypothesis, microstructure noise induces bias in realized variance, causing the standard deviation of standardized returns to be greater or less than 1 depending on the direction of the bias. Microstructure noise also induces upward bias in the kurtosis of standardized returns because, when volatility varies across trading days, standardized returns follow a mixture of (near) zero-mean normals. Using a sample of high-frequency returns for 20 heavily traded US stocks, we find empirical evidence consistent with these predictions.

We then construct standardized returns using an alternative realized kernel estimator of integrated variance that is (nearly)

---

* Corresponding author.
  E-mail address: jfleming@rice.edu (J. Fleming).

[1] See, e.g., Chernov et al. (2003).

---

[2] The analysis in ABDE (2001a) relies primarily on graphical devices such as QQ plots. Standard statistical tests do tend to reject the null hypothesis of a Gaussian distribution. Andersen et al. (2001b) obtain similar findings for exchange rate returns.

conditionally unbiased in the presence of microstructure noise. The returns are clearly platykurtotic and inconsistent with the standard normal. As Andersen et al. (2007b) demonstrate, this is precisely what one would expect if occasional, large jumps occur. Thus, our results provide clear evidence of misspecification in the direction of omitted jumps, consistent with findings in recent studies that directly test for the presence of jumps. These results are robust to alternative sampling frequencies, to the choice of using quotes versus transaction prices to construct estimates of integrated variance, and to splitting the sample into subperiods.

We also explore the extent to which jumps can account for the thin-tailed behavior by standardizing returns using realized bipower variation, an estimator of integrated variance robust to the presence of jumps. The kurtosis increases toward 3 for most stocks, and thus the standard normal provides a much better fit. These results complement more comprehensive studies of this question by Andersen et al. (2007b) and ABFN (2009). Although these studies, unlike ours, rely on the standard realized variance estimator, the basic finding that jumps are an important component of stock prices is similar. Andersen et al. (2007b) and ABFN (2009) also provide evidence that leverage effects are important.

Overall, our results provide guidance regarding the specification of continuous time models for asset prices, namely, that jumps are important and should not be ignored in modeling stock prices. The results also demonstrate the value of recently developed high-frequency variance estimators that are consistent in the presence of market microstructure noise.[3] While such estimators are theoretically appealing, evidence continues to accumulate regarding their improvements in practice relative to the standard realized variance estimator. Our analysis provides a striking example of a situation where using a realized kernel estimator debunks a 'stylized fact', i.e., that daily stock returns standardized by integrated variance are approximately Gaussian.

The paper is organized as follows. Section 2 reviews the relevant theory and discusses the effects of microstructure noise and discretization on the statistical properties of standardized returns. Section 3 describes our data. Section 4 presents empirical results regarding the distributional properties of standardized returns and examines the impact of accounting for jumps on the standardized returns. Section 5 concludes.

2. Theory

2.1. The mixture of normals hypothesis

Consider a log stock price process p_t evolving continuously in time. In a frictionless market with no arbitrage opportunities, p_t must obey a semimartingale process on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$, as detailed by Back (1991). Such a process may be decomposed as

$$p_t = A_t + M_t,$$

where A_t is a (local) finite variation process and M_t is a (local) martingale.

A common specification assumes that stock prices follow a jump diffusion process:

$$p_t = \int_0^t \mu(s)\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}W(s) + \sum_{j=1}^{N(t)} \kappa_j \tag{1}$$

where $\mu(s)$ is a local finite variation process, $\sigma(s)$ is a cadlag volatility process, $\mathrm{d}W(s)$ represents increments to a Brownian motion $W(s)$, N is a counting process, and κ_j are the associated nonzero jump increments. The process in (1) may be characterized as a

Brownian semimartingale with finite jump activity. Note that this process is quite general. A time-varying drift is acceptable as long as the finite variation condition is satisfied. Stochastic volatility models are also permitted, including multi-factor volatility processes as studied by Chernov et al. (2003), and the innovations to volatility and stock prices may be correlated to accommodate the so-called leverage effect. The primary restriction regarding the jump component is that an infinite number of jumps in a 'small' period is forbidden.

A leading special case of the jump diffusion process is the Brownian semimartingale without leverage effects:

$$p_t = \int_0^t \mu(s)\mathrm{d}s + \int_0^t \sigma(s)\,\mathrm{d}W(s); \quad \mu_s, \sigma_s \text{ jointly } \perp\!\!\!\perp \text{ of } W(s). \tag{2}$$

The 'pure diffusion' model in (2) exhibits continuous sample paths and additionally rules out leverage effects since the mean and volatility processes $\mu(s)$ and $\sigma(s)$ are jointly independent of the innovation process $W(s)$. Although (2) offers a convenient representation of stock prices, it rules out features that may be important empirically. In addition to potential leverage effects, it is possible that stock prices exhibit jumps resulting from the sudden arrival of firm-specific or macroeconomic news.

In order to investigate the validity of these restrictions, let n index trading days and n^O and n^C represent the opening and closing of trading on day n. For the jump diffusion process in (1), the increment in the *quadratic variation* process $[p]_t$ on trading day n is

$$[p]_{n^C} - [p]_{n^O} = \int_{n^O}^{n^C} \sigma^2(s)\mathrm{d}s + \sum_{j=N(n^O)+1}^{N(n^C)} \kappa_j^2, \tag{3}$$

where the first term in (3) captures the *integrated variance* on day n, denoted as IV_n, and the second term captures any jump component of variance on that day. ABDL (2003) note that, under the pure diffusion model (2), and assuming that the mean process $\mu(s)$ is zero,

$$\frac{r_n}{\sqrt{IV_n}} \sim N(0, 1), \tag{4}$$

where $r_n = p_{n^C} - p_{n^O}$ is the daily return.[4] As ABDL (2003) point out, the distributional result in (4) conditions on *ex post* sample path realizations of $\sigma(s)$. If ruling out jumps and leverage effects is empirically valid, then daily returns standardized by the square root of the integrated variance will follow a standard normal distribution.

2.2. The properties of returns standardized using realized variance

Assessing the adequacy of the pure diffusion specification via (4) requires an estimator for the latent quantity IV_n. Suppose that a set of $m + 1$ intraday prices are available at equally spaced intervals of length $\Delta = 1/m$. These price observations give rise to m intradaily returns $r_n^i = p_n^i - p_n^{i-1}$ for $i = 1, \ldots, m$, where $p_n^0 = p_{n^O}$ and $p_n^m = p_{n^C}$.

Econometric measurement of the increments in quadratic variation is often based on the *realized variance* (RV). The realized variance estimator on day n is obtained by summing the squared returns for each subperiod,

$$RV_n^m = \sum_{i=1}^{m} (r_n^i)^2. \tag{5}$$

---

[3] See, for example, Ait-Sahalia et al. (2005), Zhang et al. (2005), Zhang (2006), and Barndorff-Nielsen et al. (2008a).

---

[4] Ignoring time variation in the conditional mean should be innocuous in practice because this variation is empirically an order of magnitude smaller than return variation at short horizons.

The number of intraday returns m appears as a superscript in the notation RV_n^m to emphasize the dependence of the estimator on the time grid of intraday returns. ABDL (2001b) and Barndorff-Nielsen and Shephard (2002) formalized realized variance econometrically, focusing on the fact that, as $\Delta \to 0$, RV_n^m converges in probability to the daily increment in the quadratic variation of the price process p_t.

When stock prices follow a pure diffusion process (i.e., no jumps), the decomposition in (3) implies that the quadratic variation and integrated variance are identical. By using intraday returns sampled at sufficiently high frequencies, we theoretically obtain an error-free measure of the integrated variance. A feasible test of the mixture of normals hypothesis can therefore be based on $r_n/\sqrt{RV_n^m}$. In practice, however, microstructure noise and discretization effects distort the statistical properties of returns standardized using the realized variance estimator.

2.2.1. Microstructure noise

Assume that the true latent log price process is $p_t^* = p_t + u_t$, where p_t is the observed log price process and u_t represents a noise process capturing market microstructure effects. In the presence of noise, the realized variance estimator RV_n^m is biased and no longer consistent for the increment in the quadratic variation of the true, latent price process. When the price and noise processes are independent, the realized variance estimator is biased upward. When the price and noise processes are dependent, it is possible that the bias is negative. Hansen and Lunde (2006) provide evidence of such negative bias at very high sampling frequencies when observed prices are estimated using bid–ask quotes.

Now consider the consequences for the time series of standardized returns $r_n/\sqrt{RV_n^m}$.[5] When RV_n^m is upward (downward) biased, daily returns tend to be scaled by an overly large (small) number, and the standard deviation of the standardized returns will be biased downward (upward) from 1. Microstructure noise also affects the kurtosis of the standardized returns. This is particularly easy to see when the quadratic variation of the noise process $[u]$ is well-defined. In this case, assuming that the noise and price processes are independent, Hansen and Lunde (2006) show that

$$RV_n^m - IV_n \xrightarrow{p} [u]_n \equiv [u]_{n^C} - [u]_{n^O}.$$

Consequently, as $m \to \infty$ and under the null hypothesis that prices follow (2), the standardized returns $r_n/\sqrt{RV_n^m}$ converge in distribution:

$$\frac{r_n}{\sqrt{RV_n^m}} \xrightarrow{d} N\left(0, \frac{IV_n}{IV_n + [u]_n}\right), \tag{6}$$

so the time series of returns (asymptotically) follows a mixture of normals with mixing weights governed by the time series behavior of IV_n and $[u]_n$. It can be shown that a mixture of zero-mean normals, such as (6), is always leptokurtotic (see, e.g., Alexander and Narayanan (2001)). In the limit, returns standardized using the realized variance estimator will be *artificially fat-tailed due to the effects of market microstructure noise*.

More often in the literature, microstructure noise is assumed to be such that $RV_n^m \to \infty$ as $m \to \infty$, i.e., realized variance 'explodes' in the limit (see, e.g., Bandi and Russell (2006, 2008) and Zhang et al. (2005)). In an unpublished appendix (available from the authors), we provide simulation results illustrating that standardized returns are also artificially fat-tailed at high sampling frequencies for this type of noise.

2.2.2. Discretization effects

Discretization error associated with the realized variance RV_n^m for finite m also impacts the statistical properties of standardized returns. Peters and de Vilder (2006) study discretization effects for the special case of a pure diffusion process where returns are sampled over intervals of equal length. Assuming that the instantaneous volatility remains constant over each interval, they show that the density function for returns standardized using the realized variance estimator RV_n^m is given by

$$f_{\frac{r_n}{\sqrt{RV_n^m}}}(r) = \frac{\Gamma(m/2)}{\sqrt{\pi m}\Gamma((m-1)/2)}\left(1 - \frac{r^2}{m}\right)^{\frac{m-3}{2}}$$
$$\times I\left(-\sqrt{m} \le r \le \sqrt{m}\right), \tag{7}$$

where $I(\cdot)$ represents the indicator function.

The mean, variance and skewness under (7) accord with the corresponding moments of the limiting standard normal, while the kurtosis is given by $3m/(m + 2)$. The distribution of $r_n/\sqrt{RV_n^m}$ is therefore *thin-tailed* relative to the normal due to discretization effects. Note that when $m = 1$ (i.e., realized variance equals the daily squared return), the kurtosis of standardized returns is 1. As the sampling frequency increases, the density in (7) converges to the standard normal density. Intraday variation in volatility can exacerbate the effects of discretization. In fact, for any discrete sampling frequency m, a pure diffusion with cadlag volatility paths may be constructed so that the kurtosis of standardized returns is close to 1, rather than the kurtosis of $3m/(m + 2)$ under constant intraday volatility. An expanded discussion of this point is provided in the unpublished appendix.

The contrasting effects of discretization and microstructure noise with respect to sampling frequency lead to an empirical prediction for the 'signature plot' of the kurtosis of returns standardized using the realized variance estimator. The kurtosis will be significantly less than 3 at lower sampling frequencies, such as hourly, where discretization effects are strong and microstructure effects are minimal. The kurtosis will then increase with sampling frequency, as discretization effects weaken and microstructure effects become stronger, and ultimately exceed 3 at very high sampling frequencies, where microstructure effects are the most pronounced.

2.3. Consistent estimation of realized variance in the presence of noise

Several recent papers propose alternative high-frequency variance estimators that remain consistent in the presence of microstructure noise.[6] The fact that these estimators are consistent in the presence of noise permits sampling at very high frequencies, which affords more accurate estimation of the integrated variance while minimizing distortions due to discretization. Furthermore, these estimators incorporate a 'bias correction' leading to (nearly) conditionally unbiased estimation of integrated variance. For these reasons, constructing standardized returns using noise-robust estimators of integrated variance should permit more reliable inference regarding the extent to which scaled returns conform to the standard normal.

In the subsequent empirical analysis we utilize the *realized kernel* estimator proposed by Barndorff-Nielsen et al. (2008b). This estimator takes the form

$$RV(\text{BN-HLS})_n^m = \sum_{h=-H}^{H} k(h/(H+1))\gamma_h, \tag{8}$$

where $k(\cdot)$ is a kernel weighting function and $\gamma_h = \sum_{j=|h|+1}^{m} r_n^j r_n^{j-h}$.

---

[5] Since microstructure noise has only a negligible effect on the daily return, we simply assume that $r_n = r_n^*$, i.e., that the true daily return is observed.

[6] See, for example, Ait-Sahalia et al. (2005), Barndorff-Nielsen et al. (2008a), Hansen and Lunde (2003, 2006), Zhang (2006) and Zhang et al. (2005).

The parameter H is a bandwidth parameter that determines the maximum lag length captured by the covariance terms. We follow BN-HLS (2008b, 2009) by adopting the Parzen kernel for $k(\cdot)$. This kernel choice satisfies certain smoothness conditions and guarantees a non-negative variance estimate. For $H \propto m^\eta$, BN-HLS (2008b) show that $\eta \in \left(\frac{1}{2}, 1\right)$ delivers a consistent estimator when the noise process is covariance stationary. We set the bandwidth parameter H to $m^{3/5}$, which BN-HLS (2008b) establish as the optimal rate under MSE loss.[7]

3. Data

Our data consist of intradaily transaction prices and bid–ask quotes for the 20 stocks in the Major Market Index (MMI).[8] We obtain these data from the Trade and Quote (TAQ) database distributed by the New York Stock Exchange (NYSE). The majority of our analysis is based on intradaily returns constructed using bid–ask midpoints ('quotes'). For comparative purposes, we sometimes also consider intradaily returns constructed using transaction prices ('trades'). The sample period is January 4, 1993 to December 31, 2003 (2770 trading days).[9]

We implement a data cleaning procedure to filter the raw price data prior to constructing the intradaily returns. The importance of cleaning high-frequency intraday data prior to volatility estimation is emphasized by a number of studies (e.g., Dacorogna et al. (2001, Chapter 4) and Hansen and Lunde (2006)). Our data cleaning procedure is similar, although not identical, to that proposed by BN-HLS (2009). A detailed discussion is provided in the unpublished appendix.

On most days, trading begins at 9:30 am (EST) and ends at 4:00 pm, a period of 390 min. We divide the trading period into m intervals to construct intradaily returns with a sampling frequency of $390/m$ min. We consider a range of values for m, from $m = 780$ (i.e., 30 s intervals) to $m = 6.5$ (i.e., 60 min intervals).[10] For a given value of m, we treat the first quote in the cleaned database as the starting price for the interval in which the observation occurs. We then estimate the price at the end of this interval, and at the end of each successive interval throughout the day, using the *previous-tick* approach (see, for example, Hansen and Lunde (2006)). The $390/m$ min intradaily returns are then computed as the first differences of the logs of these prices. For our comparative analysis using trades, we estimate the price at the end of each interval using *linear interpolation* (see Andersen and Bollerslev (1997)) to reduce the effect of bid–ask bounce at high sampling frequencies.

IBM is the most actively quoted stock in the sample, with an average of 3099 quotes per day, and EK is the least, with an average of 1406 quotes per day. Sears has the highest volatility of daily returns, 32.8% annualized, and XOM has the lowest, 21.4% annualized. These basic statistics are qualitatively similar for returns based on trades (instead of quotes); however, differences arise in the serial correlation of the intradaily returns. Using quotes,
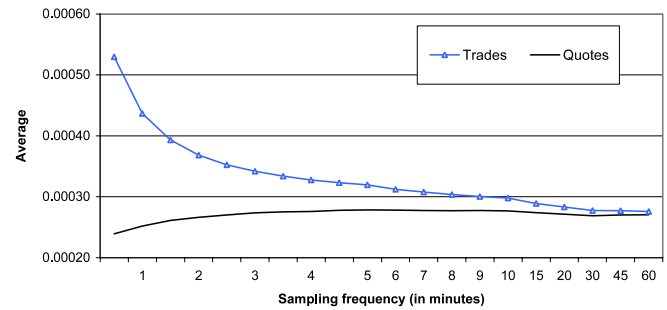

Fig. 1. Volatility signature plots. The figure plots the mean daily realized variance constructed using transaction prices (trades) and using bid–ask midpoints (quotes). For each stock, we construct the realized variances using sampling frequencies ranging from 30 s to 60 min. The figure plots the mean realized variance for each sampling frequency, averaged across stocks. The sample period is January 4, 1993, to December 31, 2003 (2770 observations for most stocks).

the first-order serial correlation for returns at very high sampling frequencies is typically positive (for 30 sec returns, the average value for the 20 stocks is 0.06), but close to zero for returns at lower frequencies (e.g., 5 min returns). In contrast, using trades, the first-order serial correlation at very high sampling frequencies is negative (−0.15 for 30 sec returns) and the magnitude decays more slowly with sampling frequency (−0.07 for 5 min returns). These findings are consistent with previous studies such as Hansen and Lunde (2006).

4. Empirical results

4.1. Are standardized returns approximately standard normal?

4.1.1. Daily returns standardized using realized variance

Fig. 1 plots the average realized variance, averaged across days and across stocks, for sampling frequencies ranging from $m = 780$ to $m = 6.5$. The figure is similar to the volatility signature plot proposed by ABDL (2000).[11] ABDL (2000) recommend selecting the finest sampling frequency at which microstructure effects do not appear to bias the estimates. For realized variances constructed using quotes, the estimates are downward biased at the highest sampling frequencies, e.g., 30 sec, consistent with findings in Hansen and Lunde (2006). A sampling frequency of about four minutes appears to be optimal. For realized variances constructed using trades, the estimates are substantially upward biased at the five-minute sampling frequency, and the bias sharply increases at the highest frequencies.

Fig. 2 shows the signature plots for estimates of the standard deviation, skewness and kurtosis of daily returns standardized by realized variance. Panel A illustrates the contrasting behavior of the standard deviation of scaled returns at high sampling frequencies for realized variances based on quotes versus those based on trades. For realized variances based on quotes, the standard deviation is significantly greater than 1 at the highest sampling frequencies due to downward bias in the realized variance estimator, whereas the standard deviation using realized variances based on trades is significantly less than 1 due to the corresponding upward bias in the realized variance estimator. Panel B shows that the skewness of standardized returns is reasonably flat as a function of the sampling frequency for both quotes and trades. In Panel C, the kurtosis of standardized returns

---

[7] Our main findings are robust to other methods of setting the bandwidth. These robustness results are reported in the unpublished appendix.

[8] These stocks are American Express (AXP), AT&T (T), ChevronTexaco (CVX), Coca-Cola (KO), Disney (DIS), Dow Chemical (DOW), DuPont (DD), Eastman Kodak (EK), Exxon-Mobil (XOM), General Electric (GE), General Motors (GM), International Business Machines (IBM), International Paper (IP), Johnson & Johnson (JNJ), McDonald's (MCD), Merck (MRK), 3M (MMM), Philip Morris (MO), Procter and Gamble (PG), and Sears (S).

[9] Philip Morris has one less day than the other firms because its stock did not open on May 25, 1994, in advance of a board meeting regarding a proposal to split the firm's food and tobacco businesses.

[10] For fractional values of m, the final fractional return is treated as though it were a full $390/m$ min return.

[11] To conserve space, we do not present separate figures for each stock; however, these figures show that the individual stocks generally share the features shown in Fig. 1.
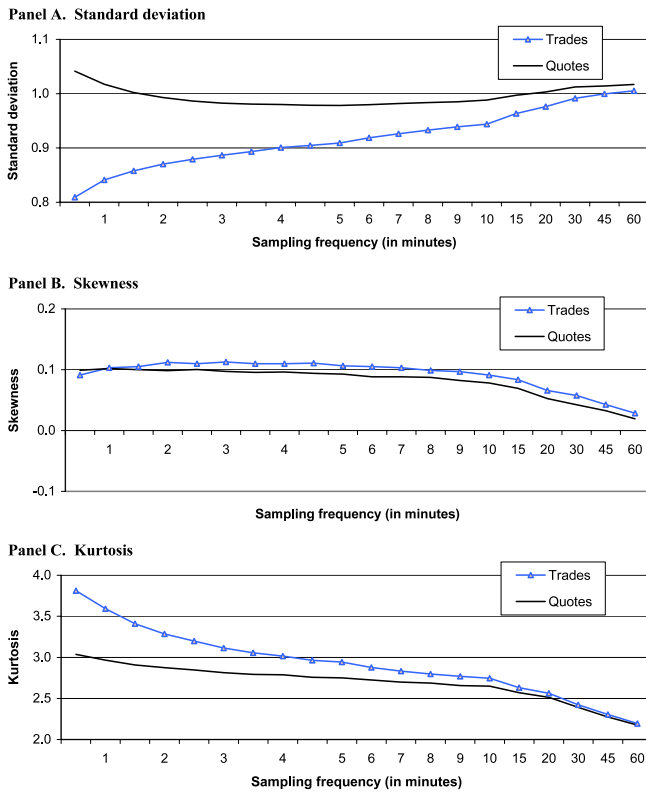
Fig. 2. Signature plots for returns standardized by realized variance. The figure plots the standard deviation, skewness and kurtosis of daily returns standardized by realized variances constructed using transaction prices (trades) and using bid–ask midpoints (quotes). We construct the realized variances using sampling frequencies ranging from 30 s to 60 min. The figure plots the mean standard deviation (Panel A), skewness (Panel B) and kurtosis (Panel C) estimates for the MMI stocks. The sample period is January 4, 1993, to December 31, 2003 (2770 observations for most stocks).

steadily decreases as the sampling frequency decreases, for both quotes and trades. Using trades, the kurtosis signature plot for most stocks is a downward sloping curve which starts at close to 4.0 at the highest sampling frequency, flattens out between the three- and ten-minute sampling frequencies, and decreases below 2.5 beyond the 20 min sampling frequency. The kurtosis signature plot using quotes exhibits a similar pattern, but starting just slightly above 3.0 at the highest sampling frequency.

These empirical findings are consistent with our discussion of the impact of microstructure noise and discretization on the moments of standardized returns in Section 2. Specifically, sampling at the highest frequencies, microstructure noise induces upward bias in realized variances constructed using trades and downward bias in realized variances constructed using quotes. These biases, in turn, lead to excess kurtosis and a standard deviation less than 1 using trades, and excess kurtosis and a standard deviation greater than 1 using quotes. Moreover, the standard deviation trends toward 1 and the kurtosis decreases below 3 at lower sampling frequencies as discretization effects become more dominant than microstructure effects.

Despite these pronounced features, it is interesting to note that, at the five-minute sampling frequency, the moments of the standardized returns are 'approximately' Gaussian. The standard deviation is between 0.9 and 1.0, the skewness is about 0.1, and the kurtosis is close to 3.0, using both trades and quotes. Nonetheless, statistical tests (not reported) reject the null hypothesis that the standardized returns conform to the standard normal.

4.1.2. Daily returns standardized using the realized kernel estimator
Table 1 presents summary statistics for returns standardized using the alternative BN-HLS (2008b) realized kernel estimator

described in (8). The realized kernel estimates are constructed using quotes, and using a 30 s sampling frequency, since the volatility signature plots (not shown) are flat, even at the highest sampling frequencies. The standard deviation estimates for the standardized returns are much closer to 1 than those shown in Fig. 2 for the highest sampling frequencies. The kurtosis estimates, however, are below 2.6 for all stocks, with a median value of 2.42. The table also reports *p*-values for tests against the standard normal distribution based on Hermite polynomials as suggested by Bontemps and Meddahi (2005).[12] The results indicate that daily returns standardized by the realized kernel estimator are clearly platykurtotic, with the test based on H_4 rejecting for all of the stocks at the 99% level. The test based on H_2 rejects for roughly half the stocks, and all of the joint tests based on the second and fourth, on the third and fourth, and on the second, third, and fourth Hermite polynomials reject for all of the stocks.

The QQ plots shown in Fig. 3 visually illustrate the inadequacy of the standard normal distribution. For most stocks, the plots display a marked 'S-shaped' pattern indicative of thin-tailed behavior relative to the standard normal. The thin-tailed behavior apparent in Table 1 and Fig. 3 accords with the reasoning of Andersen et al. (2007b), who argue that standardized returns will be thin-tailed in the presence of jumps.

We conduct two additional analyses to probe the robustness of our finding that returns standardized by a (nearly) conditionally unbiased estimator of realized variance are significantly thin-tailed relative to the Gaussian distribution. To conserve space, detailed results are not reported in the paper, but are available in the unpublished appendix. First, we replicate the analysis in Table 1 using trades instead of quotes. The results are extremely similar to those in Table 1 and Fig. 3, which is reassuring, since a properly implemented realized kernel estimator should deliver similar realized variance estimates using returns based on trades and returns based on quotes (BN-HLS (2009)). Second, we assess the impact of time variation in the nature of microstructure noise and, in particular, the impact of decimalization on January 29, 2001. We split the sample in half and compute summary statistics for standardized returns over each subperiod. The standard deviation of returns normalized using realized variance departs further from 1 during the first half of the sample for most stocks, indicating that microstructure noise is more severe during this period. For returns standardized using the realized kernel estimator, the results are comparable across subperiods. The variances are close to 1, the skewness values tend to be positive but small, and most of the kurtosis values are around 2.3 to 2.5, consistent with our full-sample results.

4.2. Do jumps account for the non-normality of standardized returns?

The pure diffusion specification in (2) appears to be an inadequate model for stock prices. Unresolved is the question of whether the arbitrage-free continuous time jump diffusive process in (1) constitutes a reasonable model for stock prices. Such a process may include leverage effects in addition to jump activity.

Consider the more general jump diffusion specification in (1) without leverage effects. Suppose momentarily that both IV_n and the jump process $\sum_{j=1}^{N(t)} \kappa_j$ are observable. An adjusted price process and associated return series may be constructed as

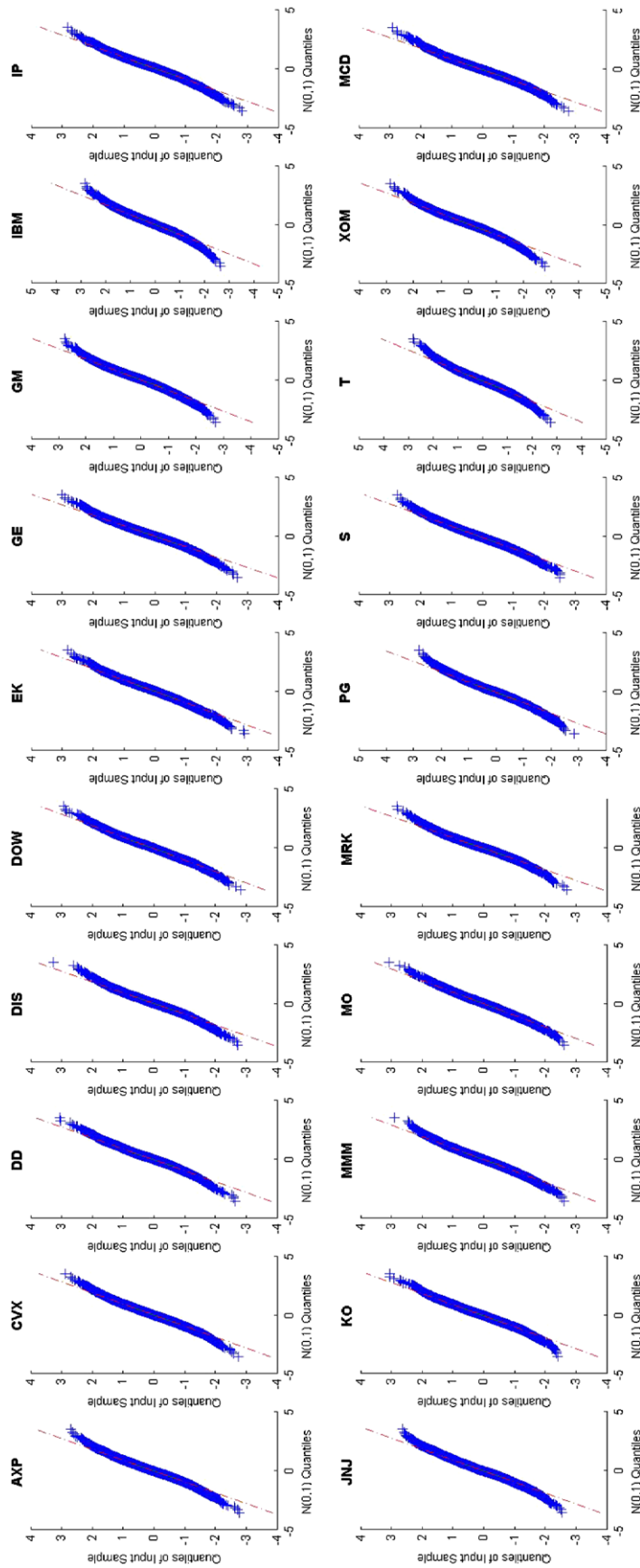$$p_t^* = p_t - \sum_{j=1}^{N(t)} \kappa_j \tag{9}$$

Fig. 3. QQ plots for standardized returns using the realized kernel estimator. The figure presents QQ plots of daily returns standardized using the BN–HLS realized kernel estimator of integrated variance. The sample period is January 4, 1993, to December 31, 2003 (2770 observations for most stocks).

Table 1
Returns standardized using the realized kernel estimator. The table reports summary statistics and the results of normality tests for daily open-to-close returns standardized using the BN-HLS realized kernel estimator of integrated variance. For each stock, we report the mean, standard deviation, skewness, and kurtosis of the standardized returns, and the *p*-values for normality tests based on the second, third and fourth Hermite polynomials (Bontemps and Meddahi, 2005), and for joint tests based on the second and fourth Hermite polynomials, the third and fourth Hermite polynomials, the second, third and fourth Hermite polynomials, and the third and fifth Hermite polynomials. The sample period is January 4, 1993, to December 31, 2003 (2770 trading days for most stocks).

Ticker	Summary statistics				Test statistics (*p*-values)						
	Mean	Std.	Skew	Kurt	H_2	H_3	H_4	H_{24}	H_{34}	H_{234}	H_{35}
AXP	0.02	0.98	0.04	2.40	0.08	0.43	0.00	0.00	0.00	0.00	0.45
CVX	−0.01	0.98	0.03	2.43	0.10	0.52	0.00	0.00	0.00	0.00	0.79
DD	0.03	0.97	0.13	2.47	0.01	0.02	0.00	0.00	0.00	0.00	0.01
DIS	−0.03	0.97	0.05	2.41	0.05	0.30	0.00	0.00	0.00	0.00	0.54
DOW	0.04	0.97	0.02	2.53	0.05	0.85	0.00	0.00	0.00	0.00	0.79
EK	−0.04	0.96	0.08	2.55	0.00	0.10	0.00	0.00	0.00	0.00	0.10
GE	0.02	1.00	0.11	2.41	0.92	0.03	0.00	0.00	0.00	0.00	0.01
GM	−0.06	1.03	0.12	2.41	0.02	0.01	0.00	0.00	0.00	0.00	0.00
IBM	−0.02	1.05	0.09	2.34	0.00	0.03	0.00	0.00	0.00	0.00	0.06
IP	−0.05	0.97	0.15	2.50	0.03	0.00	0.00	0.00	0.00	0.00	0.00
JNJ	0.02	0.96	0.03	2.45	0.00	0.69	0.00	0.00	0.00	0.00	0.89
KO	0.06	0.98	0.06	2.48	0.19	0.33	0.00	0.00	0.00	0.00	0.35
MCD	0.03	0.99	0.01	2.40	0.72	0.84	0.00	0.00	0.00	0.00	0.85
MMM	0.02	0.94	0.01	2.50	0.00	0.90	0.00	0.00	0.00	0.00	0.80
MO	−0.02	0.94	0.04	2.58	0.00	0.45	0.00	0.00	0.01	0.00	0.73
MRK	0.03	1.00	0.06	2.34	0.99	0.22	0.00	0.00	0.00	0.00	0.13
PG	0.04	1.01	0.05	2.39	0.37	0.26	0.00	0.00	0.00	0.00	0.17
S	0.11	0.95	0.00	2.47	0.00	0.52	0.00	0.00	0.00	0.00	0.43
T	0.01	1.02	0.00	2.32	0.10	0.92	0.00	0.00	0.00	0.00	0.90
XOM	−0.08	1.02	0.13	2.39	0.12	0.01	0.00	0.00	0.00	0.00	0.01

$$r_n^* \equiv p_{n^C}^* - p_{n^O}^* = r_n - \sum_{j=N(n^O)+1}^{N(n^C)} \kappa_j. \tag{10}$$

The series in (10) is simply the daily return less the return component arising from jumps. If (1) holds then daily returns r_n^*, standardized by the square root of IV_n, will follow a standard normal distribution.

In the presence of jumps, the realized variance converges not to the integrated variance, but to the quadratic variation, which also includes the jump contribution to variance. Here we employ an alternative estimator, *realized bipower variation*, that is (asymptotically) capable of extracting the integrated variance component from total price variability:

$$BV_n^m = \frac{\pi}{2} \frac{m}{m-1} \sum_{i=2}^{m} |r_n^i| \, |r_n^{i-1}|. \tag{11}$$

If log prices follow (1), then under mild conditions Barndorff-Nielsen and Shephard (2006) and Barndorff-Nielsen et al. (2006) show that

$$BV_n^m \to \int_{n^O}^{n^C} \sigma^2(s)\mathrm{d}s.$$

We can therefore use BV_n^m as a jump-robust estimator for IV_n to implement a test of whether $r_n^*/\sqrt{IV_n}$ follows a standard normal distribution. Implementing the test also requires an estimator for the jump component, J_n, where $r_n^* = r_n - J_n$. As a simple, exploratory approach, we take $r_n^* = r_n$. This approach sidesteps potentially thorny issues in testing for and extracting jumps; however, it also suffers from the obvious weakness that jumps remain unfiltered in daily returns. Nonetheless, our ambition in the present analysis is simply to illustrate the extent to which implicitly accounting for jumps by using a jump-robust estimator of integrated variance to standardize returns alleviates obvious signs of model misspecification.

We use bipower variation signature plots to select an appropriate sampling frequency for constructing realized bipower variation.[13] As in the case of realized variance, the 'optimal' sampling

frequency is the highest frequency at which the BV_n^m estimates appear to be unbiased. The signature plots (not shown) indicate that the optimal sampling frequency is three minutes for the majority of the stocks.[14]

Fig. 4 displays QQ plots for daily returns standardized by realized bipower variation. It is visually clear that these returns provide a much better approximation to the standard normal distribution. Table 2 presents estimates of the moments for scaled returns and *p*-values for tests against the standard normal. Most significantly, the kurtosis estimates for all stocks are substantially higher than in Table 1 and close to the value of 2.85 theoretically expected using the distributional result in (7). There is still, however, some evidence of misspecification. The standard deviations tend to be slightly less than 1 and the standardized returns for some stocks continue to be positively skewed, potentially indicative of leverage effects. In addition, the H_3–H_5 test fails to reject for many stocks. To the extent that microstructure noise does not affect the odd moments, these test results are consistent with the notion that microstructure noise is a remaining issue.

Independently of our work, Andersen et al. (2007b) and ABFN (2009) extend the empirical approach described above to extract the jump component of daily returns J_n and robustify the analysis to the presence of leverage effects. To extract the jump component, they apply a test based on the difference between RV_n^m and BV_n^m to detect jumps. If the null of no jumps is rejected, they apply a procedure to estimate J_n. To handle leverage effects, they follow Peters and de Vilder (2006) and sample returns in 'financial time', corresponding to equal increments in the quadratic variation process. If prices follow a pure diffusion process (after jumps have been removed), then, by the Dambis–Dubins–Schwartz theorem, normalized returns sampled in financial time will be standard normal even in the presence of leverage effects. Andersen et al. (2007b) and ABFN (2009) find that, after applying this procedure, normalized returns accord well with the standard normal.

---

[13] ABFN (2006) suggest using signature plots to determine the sampling frequency for power variation measures.

[14] We also investigated 'skipped' versions of realized bipower variation (see, e.g., Huang and Tauchen (2005) and Barndorff-Nielsen and Shephard (2006)), but found that these methods only partially mitigate the bias at high sampling frequencies.
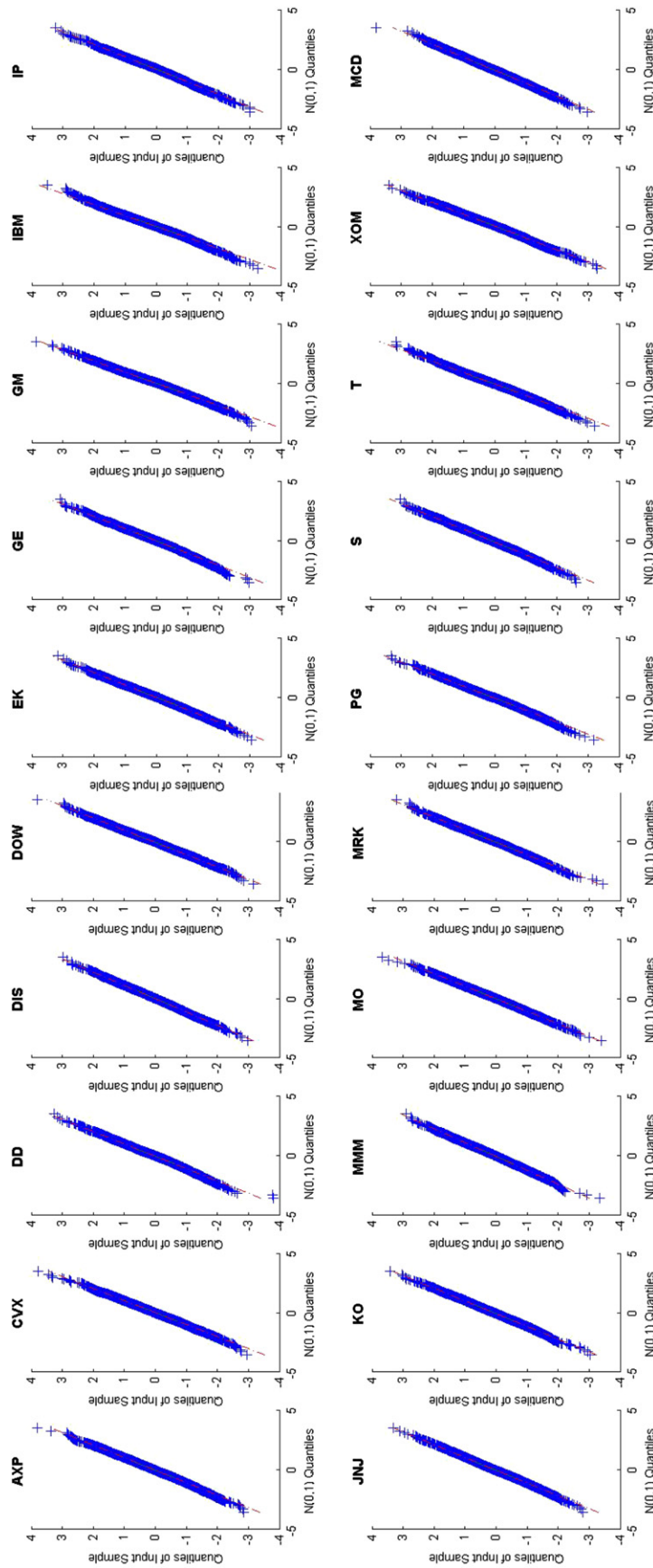
Fig. 4. QQ plots for standardized returns using realized bipower variation. The figure presents QQ plots of daily returns standardized using the realized bipower variation estimator of integrated variance. The sample period is January 4, 1993, to December 31, 2003 (2770 observations for most stocks).

Table 2
Returns standardized by realized bipower variation. The table reports summary statistics and the results of normality tests for daily open-to-close returns standardized by realized bipower variation. For each stock, we report the mean, standard deviation, skewness, and kurtosis of the standardized returns, and the *p*-values for normality tests based on the second, third and fourth Hermite polynomials (Bontemps and Meddahi, 2005), and for joint tests based on the second and fourth Hermite polynomials, the third and fourth Hermite polynomials, the second, third and fourth Hermite polynomials, and the third and fifth Hermite polynomials. The sample period is January 4, 1993, to December 31, 2003 (2770 trading days for most stocks).

Ticker	Summary statistics				Test statistics (*p*-values)						
	Mean	Std.	Skew	Kurt	H_2	H_3	H_4	H_{24}	H_{34}	H_{234}	H_{35}
AXP	0.02	0.93	0.13	2.99	0.00	0.04	0.53	0.00	0.09	0.00	0.10
CVX	−0.01	0.97	0.10	3.00	0.02	0.06	0.92	0.05	0.16	0.02	0.13
DD	0.03	0.93	0.13	2.95	0.00	0.04	0.76	0.00	0.13	0.00	0.00
DIS	−0.01	0.87	0.07	2.99	0.00	0.25	0.09	0.00	0.12	0.00	0.28
DOW	0.05	0.98	0.06	2.91	0.19	0.28	0.46	0.32	0.42	0.32	0.49
EK	−0.03	0.95	0.12	2.92	0.00	0.02	0.63	0.00	0.05	0.00	0.02
GE	0.02	0.92	0.15	2.83	0.00	0.02	0.60	0.00	0.05	0.00	0.01
GM	−0.05	1.04	0.14	2.83	0.00	0.00	0.03	0.00	0.00	0.00	0.01
IBM	−0.01	1.01	0.11	2.73	0.43	0.02	0.00	0.01	0.00	0.00	0.02
IP	−0.04	0.94	0.16	2.97	0.00	0.00	0.99	0.00	0.01	0.00	0.00
JNJ	0.02	0.90	0.05	2.89	0.00	0.57	0.73	0.00	0.80	0.00	0.84
KO	0.06	0.91	0.09	2.99	0.00	0.39	0.32	0.00	0.42	0.00	0.58
MCD	0.03	0.88	0.05	2.93	0.00	0.75	0.26	0.00	0.51	0.00	0.95
MMM	0.03	0.87	0.10	2.96	0.00	0.40	0.06	0.00	0.13	0.00	0.27
MO	−0.01	0.93	0.09	3.03	0.00	0.10	0.42	0.00	0.18	0.00	0.22
MRK	0.03	0.94	0.08	2.90	0.00	0.22	0.78	0.00	0.45	0.00	0.07
PG	0.04	0.95	0.10	2.84	0.00	0.11	0.32	0.00	0.17	0.00	0.20
S	0.10	0.90	0.09	2.84	0.00	0.88	0.84	0.00	0.97	0.00	0.77
T	0.02	0.99	0.06	2.72	0.55	0.26	0.00	0.01	0.01	0.02	0.32
XOM	−0.06	0.97	0.12	3.05	0.02	0.01	0.80	0.08	0.03	0.01	0.01

5. Conclusion

Microstructure noise can distort the variance of standardized returns in a manner directly related to the upward (or potentially downward, when using quotes) bias in realized variance and artificially inflate the kurtosis of standardized returns, especially at very fine sampling frequencies. Without controlling for these effects, daily returns standardized by realized volatility may artificially appear to be approximately standard normal. Using a realized kernel estimator of integrated variance that is (nearly) unbiased, the standardized returns are clearly platykurtotic and the normal distribution is soundly rejected. Further, returns normalized using bipower variation, an estimator of integrated variance robust to jumps, show a dramatic improvement in the closeness of fit to the standard normal distribution.

On the basis of these results, we conclude that a primary restriction of the mixture of normals hypothesis (i.e., continuous price paths) is violated in the data, which is consistent with findings by other researchers. This conclusion complements those of ABFN (2009), who find that accounting for both jumps and leverage effects is important for stock returns, and Andersen et al. (2007b) who obtain a similar finding for S&P 500 index returns.

Acknowledgements

References

Ait-Sahalia, Y., Mykland, P., Zhang, L., 2005. Ultra-high frequency volatility estimation with dependent microstructure noise. Working Paper w11380, NBER.

Alexander, C.O., Narayanan, S., 2001. Option pricing with normal mixture returns: modelling excess kurtosis and uncertainty in volatility. Working Paper, University of Reading.

Andersen, T.G., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. Journal of Finance 57, 1239–1284.

Andersen, T.G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115–158.

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

Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H., 2001a. The distribution of realized stock return volatility. Journal of Financial Economics 61, 43–76.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2000. Great realizations. Risk Magazine 13, 105–108.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001b. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42–55.

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

Andersen, T.G., Bollerslev, T., Dobrev, D., 2007b. No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: theory and testable distributional implications. Journal of Econometrics 138, 125–180.

Andersen, T.G., Bollerslev, T., Frederiksen, P.H., Nielsen, M.O., 2006. Comment on P.R. Hansen and A. Lunde: realized variance and market microstructure noise. Journal of Business and Economic Statistics 24, 173–179.

Andersen, T.G., Bollerslev, T., Frederiksen, P.H., Nielsen, M.O., 2009. Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. Journal of Applied Econometrics 25, 233–261.

Back, K., 1991. Asset prices for general processes. Journal of Mathematical Economics 20, 317–395.

Bandi, F.M., Russell, J.R., 2006. Separating market microstructure noise from volatility. Journal of Financial Economics 79, 655–692.

Bandi, F.M., Russell, J.R., 2008. Microstructure noise, realized variance, and optimal sampling. Review of Economic Studies 75, 339–369.

Barndorff-Nielsen, O.E., Graversen, S.E., Jacod, J., Shephard, N., 2006. Limit theorems for bipower variation in financial econometrics. Econometric Theory 22, 677–719.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008a. Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise. Econometrica 76, 1481–1536.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008b. Multivariate realized kernels: consistent positive semi-definite estimators of the covariance of equity prices with noise and non-synchronous trading. Working Paper, Stanford University.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2009. Realized kernels in practice: trades and quotes. Econometrics Journal 12, C1–C32.

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

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

Bontemps, C., Meddahi, N., 2005. Testing normality: a GMM approach. Journal of Econometrics 124, 149–186.

Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics. Journal of Econometrics 116, 225–257.

Dacorogna, M., Gencay, R., Muller, U., Olsen, R.B., Pictet, O.V., 2001. An Introduction to High-Frequency Finance. Academic Press, London.

Hansen, P.R., Lunde, A., 2003. An optimal and unbiased measure of realized variance based on intermittent high-frequency data. Working Paper, Mimeo prepared for the CIREQ-CIRANO Conference: realized volatility.

Hansen, P.R., Lunde, A., 2006. Realized variance and market microstructure noise. Journal of Business and Economic Statistics 24, 127–218.

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. Journal of Financial Econometrics 3, 456–499.

Jiang, G.J., Oomen, R.C.A., 2007. Estimating latent variables and jump diffusion models using high-frequency data. Journal of Financial Econometrics 5, 1–30.

Peters, R.T., de Vilder, R.G., 2006. Testing the continuous semimartingale hypothesis for the S&P 500. Journal of Business and Economic Statistics 24, 444–454.

Tauchen, G., Zhou, H., 2011. Realized jumps on financial markets and predicting credit spreads. Journal of Econometrics 160 (1), 102–118.

Zhang, L., 2006. Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. Bernoulli 12, 1019–1043.

Zhang, L., Mykland, P.A., Ait-Sahalia, Y., 2005. A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100, 1394–1411.