Volatility*

Federico M. Bandi and Jeffrey R. Russell
Graduate School of Business, The University of Chicago

January 2005

Table of contents

*Preliminary and incomplete version. This review is still very much work in progress. Comments welcome.** *Forthcoming in the Handbook of Financial Engineering. Elsevier. Edited by John R. Birge and Vadim Linetsky.*

1 Introduction

Recorded asset prices deviate from their equilibrium values due to the presence of market microstructure frictions. Hence, the volatility of the observed prices depends on two distinct components, i.e., the volatility of the unobserved equilibrium price and the volatility of the equally unobserved market microstructure effects.

In keeping with this basic premise, this review starts from a model of price formation that allows for empirically relevant market microstructure effects to discuss current advances in the nonparametric estimation of both volatility notions using high-frequency price data.

Numerous insightful reviews have been written on volatility. The existing reviews concentrate on work that assumes observability of the equilibrium price and study its volatility properties in the absence of measurement error (see Andersen et al. (2002) and the references therein). Reviews have also been written on work that solely focuses on the measurement error and characterizes it in terms of frictions induced by the market fine grain dynamics (see Hasbrouck (1996), Stoll (2000), and the references in the special issue of the Journal of Financial Markets on execution costs, for example). Quantifying these frictions is of crucial importance to understand and measure the effective execution cost of trades.

The present review places equal emphasis on the volatilities of both unobserved components of a recorded price, i.e., equilibrium price and microstructure frictions. Specifically, we provide a unified framework to understand current advances in two important finance fields, namely equilibrium price volatility estimation and transaction cost evaluation.

We begin with a general price formation mechanism that expresses recorded asset prices as the sum of equilibrium prices and market microstructure effects.

2 A model of price formation with microstructure effects

Write the observed logarithmic price as

$$p = p^* + \eta, \tag{1}$$

where p^* denotes the logarithmic equilibrium price, i.e., the price that would prevail in the absence of market microstructure frictions,[1] and η denotes a microstructure contamination in the observed logarithmic price as induced by price discreteness and bid-ask bounce effects, for instance (see Stoll (2000)). Fix a certain time period h (a day, say) and assume availability of M high-frequency prices over h. Given Eq. (1) we can readily define continuously-compounded returns over any intra-period interval of length $\delta = \frac{h}{M}$ and write

$$\underbrace{p_{j\delta} - p_{(j-1)\delta}}_{r_{j\delta}} = \underbrace{p^*_{j\delta} - p^*_{(j-1)\delta}}_{r^*_{j\delta}} + \underbrace{\eta_{j\delta} - \eta_{(j-1)\delta}}_{\varepsilon_{j\delta}}. \tag{2}$$

[1]We start by being deliberately unspecific about the nature of the equilibrium (or fair) price. We will add more economic structure to the model when discussing transaction cost evaluation (Section 6).

The following assumptions are imposed on the price process and market microstructure effects.

Assumption 1. (The Price Process.) *The logarithmic price process p_t^* is a continuous stochastic volatility semimartingale. Specifically,*

(1)
$$p_t^* = \alpha_t + m_t, \tag{3}$$

where α_t (with $\alpha_0 = 0$) is a continuous drift process of finite variation defined as $\int_0^t \phi_s ds$ and m_t is a continuous local martingale defined as $\int_0^t \sigma_s dW_s$, with $\{W_t : t \geq 0\}$ denoting a standard Brownian motion.

(2) *The spot volatility process σ_t is càdlàg and bounded away from zero.*

(3) *The integrated variance process $\int_0^t \sigma_s^2 ds$ is bounded almost surely for all $t < \infty$.*

Assumption 2. (The Microstructure Noise.)

(1) *The microstructure frictions in the price process $\eta_{j\delta}'s$ have mean zero and are strictly stationary with joint density $f_M(.)$.*

(2) *The variance of $\varepsilon_{j\delta} = \eta_{j\delta} - \eta_{(j-1)\delta}$ is $O(1)$ for all j and all M.*

(3) *The $\eta_{j\delta}'s$ are independent of the $p_{j\delta}^{*\prime}s$ for all j and all M.*

In agreement with asset-pricing theory, Assumption 1 implies that the equilibrium return process evolves in time as a stochastic volatility martingale difference plus an adapted process of finite variation. The stochastic spot volatility can display jumps, diurnal effects, high-persistence (possibly of the long-memory type), and nonstationarities. Furthermore, leverage effects (i.e., dependence between σ and the Brownian motion W) are allowed.

Assumption 2 permits general dependence features for the microstructure noise components in the recorded prices. The correlation structure of the microstructure noise contaminations can, for instance, capture first-order negative autocorrelations in the recorded high-frequency returns as determined by bid-ask bounce effects (see Roll (1984), among others) as well as higher order dependences in the market frictions as induced by clustering in order flows. In general, the characteristics of the noise returns ε's may depend on the sampling frequency $\delta = \frac{h}{M}$. The joint density of the η's has a subscript M to make this dependence explicit. Similarly, the symbol $\mathbf{E}_M$ will be later used to denote expectations of the noise returns taken with respect to the measure $f_M(.)$.

While the equilibrium return process $r_{j\delta}^*$ is modelled as being $O_p\left(\sqrt{\delta}\right)$ over any intra-period time horizon of size $\delta = \frac{h}{M}$, the contaminations in the observed return process are $O_p(1)$. This result, which is a consequence of Assumptions 1(1) and 2(2), implies that longer period returns are less contaminated by noise than shorter period returns. On the other hand, the size of the contaminations does not decrease in probability with the distance between subsequent time stamps. Provided sampling does not occur between high-frequency price updates, the rounding of recorded prices to a grid (i.e., price discreteness) *alone* makes this feature of the set-up presented above empirically compelling. The different stochastic order of r_δ^* and ε_δ is an important aspect of some recent approaches to equilibrium price variance estimation as well as to transaction cost evaluation as we discuss below.

2.1 The *MA(1)* case

Sometimes the dependence structure of the microstructure noise process can be simplified. Specifically, one can modify Assumption 2 as follows:

Assumption 2b.

(1) *The microstructure frictions in the price process* $\eta'_{j\delta}s$ *are i.i.d. mean zero.*

(3) *The* $\eta'_{j\delta}s$ *are independent of the* $p^{*\prime}_{j\delta}s$ *for all* j *and all* M.

If the microstructure noise contaminations in the price process $\eta_{j\delta}$ are *i.i.d.*, then the noise returns $\varepsilon_{j\delta}$ display an $MA(1)$ structure and are negatively correlated. Importantly, the noise return moments do not depend on M, i.e., the number of observations over h or, equivalently, the sampling frequency $\frac{\delta}{M}$. This is an important feature of the $MA(1)$ model which, as we discuss below, has been exploited in recent work on volatility estimation.

The $MA(1)$ model, as typically justified by bid-ask bounce effects (Roll (1984)), is known to be a realistic approximation in decentralized markets where traders arrive in a random fashion with idiosyncratic price setting behavior, the foreign exchange market being a valid example (see Bai et al. (2004)). It can also be a good approximation in the case of equities when considering transaction prices or even quotes posted on multiple exchanges.

3 The variance of the equilibrium price

The recent availability of quality high-frequency financial data has motivated a growing literature devoted to the model-free measurement of variance. We refer the interested reader to the review paper by Andersen et al. (2002) and the references therein. The main idea is to aggregate intra-daily squared returns and compute $\widehat{V} = \sum_{j=1}^{M} r_{j\delta}^2$ over a period h. The quantity $\widehat{V}$, which has been termed "realized variance," is thought to approximate the daily increments of the quadratic variation of the semimartingale that drives the underlying logarithmic price process, i.e., $V = \int_0^h \sigma_s^2 ds$. The consistency result justifying this procedure is the convergence in probability of $\widehat{V}$ to V as returns are computed over intervals that are increasingly small asymptotically, that is as $\delta \to 0$ or, equivalently, as $M \to \infty$ for a fixed h. This result is a cornerstone in semimartingale process theory (see Chung and Williams (Theorem 4.1, page 76, 1990), for instance). More recently, the fundamental work of Andersen et al. (2001, 2003a) and Barndorff-Nielsen and Shephard (2002, 2004), BN-S hereafter, has championed empirical implementation of these ideas while providing a complete inferential theory to facilitate their application.

The theoretical validity of the procedure hinges on the observability of the equilibrium price process. However, it is widely accepted that the equilibrium price process and, as a consequence, the equilibrium return data are contaminated by market microstructure effects. Even though the realized variance literature is aware of the potential importance of market microstructure effects, it has largely abstracted from them. The theoretical and empirical consequences of the presence of market microstructure frictions in the observed price process have been explored only recently.

3.1 Inconsistency of the realized variance estimator

Under the price formation mechanism in Section 2, the realized variance estimates are asymptotically dominated by noise as the number of squared return data increases over a fixed time period. Write

$$\widehat{V} = \sum_{j=1}^{M} r_{j\delta}^2 = \sum_{j=1}^{M} r_{j\delta}^{*2} + \sum_{j=1}^{M} \varepsilon_{j\delta}^2 + 2\sum_{j=1}^{M} r_{j\delta}\varepsilon_{j\delta}. \tag{4}$$

Since $r_{j\delta}^*$ is $O_p\left(\sqrt{\delta}\right)$ and $\varepsilon_{j\delta}$ is $O_p(1)$, the term $\sum_{j=1}^{M} \varepsilon_{j\delta}^2$ is the dominating term in the sum. Specifically, $\sum_{j=1}^{M} \varepsilon_{j\delta}^2$ diverges to infinity almost surely as $M \to \infty$. The theoretical consequence of this effect is a realized variance estimator that fails to converge to the increment of the quadratic variation (or integrated variance) of the underlying logarithmic price process but, instead, increases without bound almost surely over any fixed period of time, however small: $\widehat{V} \overset{a.s.}{\to} \infty$ as $M \to \infty$ (or $\delta = \frac{h}{M} \to 0$ given h). This point has been made in independent and concurrent work by Bandi and Russell (2004a,b) and Zhang et al. (2004).[2]

The divergence to infinity of the realized variance estimator over any fixed time period is an asymptotic approximations to a rather pervasive empirical fact. When computing realized variance estimates for a variety of sampling frequencies δ, the resulting estimates tend to increase substantially as one moves to high frequencies (i.e., as $\delta \to 0$). In the terminology of Andersen et al. (1999, 2000), "the volatility signature plots," namely the plots of realized variance estimates versus different sampling frequencies, are generally upward sloping at high frequencies. Figure 1 shows the volatility signature plots constructed for IBM midquotes obtained from i) just NYSE quotes and ii) NYSE and midwest exchange quotes. Figure 2 presents volatility signature plots for IBM from using i) NYSE and NASDAQ quotes and ii) all quotes from the consolidated market. Figure 3 presents volatility signature plots for midquotes obtained from two NASDAQ stocks (Cisco Systems and Microsoft). In all cases the realized variance estimates increase as the sampling interval decreases.

[2]This theoretical result does not hinge on the independence between the price process and the noise, as implied by Assumption 2(iii). Hence, Assumption 2(iii) can be relaxed. Also, it does not hinge on an *MA(1)* structure for the noise return component ε. Bandi and Russell (2004a) clarify both statements.

3.2 The mean-squared error of the realized variance estimator

The presence of market microstructure contaminations induces a bias/variance trade-off in integrated variance estimation through realized variance. When the equilibrium price process is observable, higher sampling frequencies over a fixed period of time result in more precise estimates of the integrated variance of the logarithmic price (see Andersen et al. (2003a) and BN-S (2002)). When the equilibrium price process is not observable, as is the case in the presence of microstructure frictions, frequency increases provide information about the underlying integrated variance but, inevitably, entail accumulation of noise that affects both the bias and the variance of the estimator (Bandi and Russell (2004a,b) and Zhang et al. (2004)).

Under Assumptions 1 and 2, absence of leverage effects and unpredictability of the equilibrium returns (i.e., $\alpha_t = 0$),[3] Bandi and Russell (2004a) provide an expression for the conditional (on the underlying volatility path) mean-squared error (MSE) of the realized variance estimator as a function of the sampling frequency δ (or, equivalently, as a function of the number of observations M), i.e.,

$$\mathbf{E}_M\left(\widehat{V}-V\right)^2 = 2\frac{h}{M}\left(Q+o(1)\right)+\Lambda_M, \tag{5}$$

where

$$\Lambda_M = M\mathbf{E}_M\left(\varepsilon^4\right)+2\sum_{j=1}^{M}(M-j)\mathbf{E}_M\left(\varepsilon^2\varepsilon_{-j}^2\right)+4\mathbf{E}_M(\varepsilon^2)V \tag{6}$$

and $Q=\int_0^h \sigma_s^4 ds$ is the so-called quarticity as introduced by BN-S (2002). Notice that the bias of the estimator can be easily deduced by taking the expectation of $\widehat{V}$ in Eq. (4), i.e.,

$$\mathbf{E}_M\left(\widehat{V}-V\right) = M\mathbf{E}_M\left(\varepsilon^2\right). \tag{7}$$

As for the variance of $\widehat{V}$, we can write

$$\mathbf{E}_M\left(\widehat{V}-\mathbf{E}_M(\widehat{V})\right)^2 = 2\frac{h}{M}\left(Q+o(1)\right)+\Lambda_M - M^2\left(\mathbf{E}_M\left(\varepsilon^2\right)\right)^2. \tag{8}$$

The conditional MSE of $\widehat{V}$ can serve as the basis for an optimal sampling theory designed to choose M in order to balance bias and variance as we discuss below.

[3]Both additional assumptions, namely absence of leverage and unpredictability of the equilibrium returns, can be justified.

In the case of the latter, Bandi and Russell (2004) argue that the drift component α_t is rather negligible in practise at the sampling frequencies considered in the realized variance literature. They provide an example based on IBM. Assume a realistic annual constant drift of 0.08. The magnitude of the drift over a minute interval would be $0.08/(365*24*60) = 1.52\times10^{-7}$. Using IBM transaction price data from the TAQ data set for the month of February 2002, Bandi and Russell (2004a) compute a standard deviation of IBM return data over the same horizon equal to 9.5×10^{-4}. Hence, at the one minute interval, the drift component is 1.6×10^{-4} or nearly $1/10,000$ the magnitude of the return standard deviation.

Assuming absence of leverage effects is empirically reasonable in the case of exchange rate data. The same condition appears restrictive when examining high frequency stock returns. However, some recent work uses tractable parametric models to show that the effect of leverage on the *unconditional* MSE of the realized variance estimator in the absence of market microstructure noise is asymptotically negligible (Meddahi (2002) and Andersen et al. (2003b)). This work provides some justification for the standard assumption of no-leverage in the literature (see the review paper by Andersen et al. (2002)).

4 Solutions to the inconsistency problem

4.1 The early approaches: sparse sampling and filtering

Thorough theoretical and empirical treatments of the consequences of market microstructure contaminations in realized variance estimation are recent phenomena. However, while abstracting from in-depth analysis of the implications of frictions for variance estimation, the early realized variance literature is concerned about the presence of microstructure noise in recorded asset prices.

In order to avoid substantial noise contaminations at high-sampling frequencies, Andersen et al. (2001), for example, suggest sampling at frequencies that are lower than the highest frequencies at which the data arrives. The 5-minute interval was recommended as a valid approximate choice. Relying on the levelling off of the volatility signature plots at frequencies around 15 minutes, Andersen et al. (1999, 2000) suggest using 15 to 20-minute intervals in practise.

If the equilibrium returns are unpredictable, the correlation structure of the observed returns must be imputed to microstructure noise. Andersen et al. (2001, 2003a) filter the data using an $MA(1)$ filter. An $AR(1)$ filter is employed in Bollen and Inden (2002).

4.2 MSE-based optimal sampling

More recently, an MSE-based optimal sampling theory has been suggested by Bandi and Russell (2004a,b). Specifically, in the case of the model laid out above, the optimal frequency $\delta^* = \frac{h}{M^*}$ at which to sample continuously-compounded returns for the purpose of realized variance estimation can be chosen as the minimizer of the MSE expansion in the previous section.

Bandi and Russell's theoretical framework clarifies outstanding issues in the extant empirical literature having to do with sparse sampling and filtering. We start with the former. The volatility signature plots provide very useful insights about the bias of the realized variance estimates. The bias generally manifests itself in an upward sloping pattern as the sampling interval becomes short, i.e., the bias increases with M (see Eq. (7)). However, it would be theoretically difficult to choose a single optimal frequency based on the bias, as implied by the volatility signature plots. While it is empirically sensible to focus on low frequencies for the purpose of bias reduction, the bias is only one of the components of the estimator's estimation error. At sufficiently low frequencies the bias can be negligible. However, at the same frequencies, the variability of the estimates might be substantial (see Eq. (8)). Figure 4 is a picture from simulations for parameter values consistent with IBM. The MSE-based sampling in Bandi and Russell (2004a,b) trades-off bias and variance optimally. As for filtering, while the dependence that the noise induces in the data can be reduced by filtering, residual contaminations are bound to remain in the data. These contaminations continue to give rise to inconsistent realized variance estimates. Bandi and Russell (2004a) make this point while studying the theoretical properties of both filtering at the highest frequencies at which observations arrive and filtering at all frequencies.

The MSE criterion in Subsection 3.2 can be evaluated for the purpose of obtaining an optimal sampling frequency. Bandi and Russell (2004a) discuss evaluation of the MSE under Assumption 1 and 2 as well as in the $MA(1)$ case (i.e., under Assumptions 1 and 2b). When empirically justifiable, the $MA(1)$ case is very convenient in that the moments of the noise do not depend on the sampling frequency. Furthermore, the MSE simplifies substantially:

$$\mathbf{E}_M \left(\widehat{V} - V\right)^2 = 2\frac{1}{M}\left(Q + o(1)\right) + M\beta + M^2\alpha + \gamma, \tag{9}$$

where the parameters α, β, and γ are defined as

$$\alpha = \left(\mathbf{E}(\varepsilon^2)\right)^2, \tag{10}$$

$$\beta = 2\mathbf{E}\left(\varepsilon^4\right) - 3\left(\mathbf{E}(\varepsilon^2)\right)^2, \tag{11}$$

and

$$\gamma = 4\mathbf{E}(\varepsilon^2)V - \mathbf{E}(\varepsilon^4) + 2\left(\mathbf{E}(\varepsilon^2)\right)^2. \tag{12}$$

If M^* is large, the following approximation to the optimal sampling frequency applies

$$M^* \approx \left(\frac{hQ}{\left(\mathbf{E}(\varepsilon^2)\right)^2}\right)^{1/3}. \tag{13}$$

In the $MA(1)$ case, evaluation of the MSE does not need to be conducted on a grid of frequencies and simply relies on the consistent estimation of the frequency-independent moments of the noise ($\mathbf{E}(\varepsilon^2)$ and $\mathbf{E}(\varepsilon^4)$) as well as on the estimation of the quartic term Q.[4] In this case, Bandi and Russell (2004a,b) show that sample moments of the *observable* contaminated return data can be employed to identify the moments of the *unobservable* noise process at *all* frequencies. Thus, while realized variance is inconsistent in the presence of microstructure noise, appropriately defined arithmetic averages of the observed returns consistently estimate the moments of the noise. Under $\mathbf{E}\left(\eta^8\right) < \infty$, the following result holds

$$\frac{1}{M}\sum_{j=1}^{M} r_{j\delta}^q - \mathbf{E}\left(\varepsilon^q\right) \xrightarrow{p} 0 \qquad 1 \leq q \leq 4 \tag{14}$$

as $M \to \infty$. We provide intuition for this finding in the case $q = 2$. The sum of the squared contaminated returns can be written as in Eq. (4) above, namely as the sum of the squared equilibrium returns plus the sum of the squared noise returns and a cross-product term. The price

[4]The quartic term can be identified using the estimator proposed by BN-S (2002), namely

$$\widehat{Q} = \frac{M}{3h}\sum_{j=1}^{M} \widetilde{r}_{j\delta}^4.$$

However, $\widehat{Q}$ is not a consistent estimate of Q in the presence of noise. One could then sample the observed returns to be used in the definition of $\widehat{Q}$ at a lower frequency than the highest frequency at which observations arrive. Bandi and Russell (2004a) show by simulation that sub-optimal sampling for the quartic term does not give rise to imprecise sampling choices for realized variance. They suggest using 15 minute frequencies in practise. Using real data, Bandi and Russell (2004b) also show that sampling intervals for the quarticity between 10 and 20 minutes have virtually no effect on the resulting optimal frequencies of the realized variance estimator.

formation mechanism in Section 2 is such that the orders of magnitude of the three terms in Eq. (4) above differ since $r^*_{j\delta} = O_p\left(\sqrt{\delta}\right)$ and $\varepsilon_{j\delta} = O_p(1)$. Thus, the microstructure noise component dominates the equilibrium return process at very high frequencies, i.e., for values of δ that are small. This effect determines the diverging behavior of $\widehat{V}$. By the same logic, when we average the contaminated squared returns as in Eq. (14), the sum of the squared noises constitutes the dominating term in the average. Naturally, then, while the remaining terms in the average vanish asymptotically due to the asymptotic order of the equilibrium returns, i.e., $O_p\left(\sqrt{\delta}\right)$, the average of the squared noises converge to the second moment of the noise returns as implied by Eq. (14).

Using a sample of mid-quotes for the S&P 100 stocks over the month of February 2002, Bandi and Russell (2004a,b) report (average) daily optimal sampling frequencies that are between 1 minute and 13 minutes with a median value of about 4 minutes. The MSE improvements that the optimal MSE-based frequencies guarantee over the 5 or 15-minute frequency can be substantial. Not only do the optimal frequencies vary cross-sectionally, they also change over-time. Using mid-quotes going back to 1993 for three stocks with various liquidity features, namely EXXON Mobile Corporation (XOM), SBC communications (SBC), and Merrill Lynch (MEL), Bandi and Russell (2004b) show that the daily optimal frequencies have substantially decreased in recent times, generally due to decreases in the magnitude of the noise moments. In the context of an asset allocation strategy relying on volatility timing as in Fleming et al. (2001, 2003), Bandi and Russell (2004b) show that the economic benefit of optimally sampling realized variance over time versus sampling every 5 or 15 minutes can be considerable.

In agreement with the analysis in Bandi and Russell (2004a,b), Oomen (2004a) discusses an MSE approach to optimal sampling for the purpose of realized variance estimation. However, some important novelties characterize Oomen's work. First, the underlying equilibrium price is not modelled as in Section 2 but as a compound Poisson process. Second, Oomen explores the relative benefits of business time sampling versus calendar time sampling. The logarithmic price process in Oomen (2004a) can be expressed as

$$p_t = p_0 + \sum_{j=1}^{N(t)} \xi_j + \sum_{j=1}^{N(t)} \eta_j, \tag{15}$$

where $\xi_j \sim i.i.d\ N(\mu_\xi, \sigma_\xi^2)$, $\eta_j = \rho_0 \nu_j + \rho_1 \nu_{j-1} + ... + \rho_q \nu_{j-q}$ and $\nu_j \sim i.i.d.\ N(\mu_\nu, \sigma_\nu^2)$, with $N(t)$ denoting a Poisson process with instantaneous intensity $\lambda(t)$. The equilibrium price p_t^* is equal to $p_0 + \sum_{j=1}^{N(t)} \xi_j$ in this model. Hence, it is a jump process of finite variation in the tradition of Press (1967). The microstructure noise contaminations η_j have an $MA(q)$ structure.

Oomen (2004a) provides closed-form expressions for the MSE of the realized variance estimator under both calendar time sampling, as in the approach described above, and business time sampling. Define the average integrated intensity λ_M as

$$\lambda_M = \frac{1}{Mh} \int_0^h \lambda(s) ds. \tag{16}$$

Given M (the total number of observations), business time sampling is obtained by sampling the price process every time λ_M realizes. Oomen (2004a) discusses optimal choice of M in an MSE sense. Using IBM transaction prices from the consolidated market over the period between January 1, 2000, and August 31, 2003, he finds that business time sampling generally outperforms calendar time sampling. In his sample the average increase in MSE that calendar time sampling induces is about 3%. The largest gains are obtained for days with irregular trading patterns, early market closures, and sudden changes in market activity.

4.3 Bias-correcting

Hansen and Lunde (2004a) propose to account for microstructure noise contaminations by providing a bias-adjustment to the conventional realized variance estimator. The estimator that they suggest is in the tradition of robust covariance estimators such as those of Newey and West (1987) and Andrews and Monahan (1992). Its form is

$$\widehat{V}^{db} = \sum_{j=1}^{M} r_{j\delta}^2 + 2\sum_{h=1}^{q_M} \frac{M}{M-h} \sum_{j=1}^{M-h} r_{j\delta} r_{(j+h)\delta}, \tag{17}$$

where q_M is a frequency-dependent number of covariance terms. If the correlation structure of the noise return is such that the covariances of the noise terms of order higher than q_M are equal to zero (and $\alpha_t = 0$), then the estimator in Eq. (17) is unbiased for the underlying integrated variance over a period, i.e., $\mathbf{E}_M\left(\widehat{V}^{db}\right) = \int_0^h \sigma_s^2 ds$.

Interestingly, the finite sample unbiasedness of Hansen and Lunde's estimator is robust to the presence of dependence between the underlying local martingale price process and market microstructure noise, i.e., Assumption 2(3) is not required.

In the $MA(1)$ case Hansen and Lunde's estimator simplifies and can be written as

$$\widehat{V}^{db(MA(1))} = \sum_{j=1}^{M} r_{j\delta}^2 + 2\frac{M}{M-1} \sum_{j=1}^{M-1} r_{j\delta} r_{(j+1)\delta}. \tag{18}$$

The logic of the bias-correction is apparent in this case and worth emphasizing.. Under Assumption 2b, the correlation between $r_{j\delta}$ and $r_{(j+1)\delta}$, i.e., $\mathbf{E}_M(r_{j\delta} r_{(j+1)\delta})$, is the same at all frequencies and equal to $-\mathbf{E}(\eta^2)$. Hence, $\mathbf{E}\left(2\frac{M}{M-1}\sum_{j=1}^{M-1} r_{j\delta} r_{(j+1)\delta}\right) = -2M\mathbf{E}(\eta^2)$. However, the bias of the estimator $\widehat{V}$ is equal to $M\mathbf{E}(\varepsilon^2) = 2M\mathbf{E}(\eta^2)$ (see Eq. (7)). Therefore, the second term in Eq.(18) provides the required adjustment.

Under an assumed $MA(1)$ structure, Zhou (1996) is the first to use the estimator in Eq. (18) in the context of variance estimation through high-frequency data. He also obtains the variance of the estimator assuming a *constant* return variance and Gaussian market microstructure noise. He concludes that the variance of the estimator can be minimized for a finite M.

Hansen and Lunde (2004b) have recently further studied the MSE properties of their bias-corrected estimator in the $MA(1)$ case (i.e., the estimator in Zhou's 1996 study). Working under Assumption 2b, they find that bias-correcting permits optimal sampling at higher frequencies than those obtained by Bandi and Russell (2004a,b) when employing the standard realized variance estimator of Andersen et al. (2003) and BN-S (2002). In addition, MSE improvements can be achieved. Using 5 years of Alcoa (AA) transaction price data from January 2, 1998, to December 31, 2002, Hansen and Lunde (2004b) report an (average) daily optimal sampling frequency for the their bias-corrected estimator equal to about 5 seconds. Their reported optimal frequency for the realized variance estimator is 2 minutes. The ratio between the MSE of the realized variance estimator at 2 minutes and the MSE of the first-order bias-corrected realized variance estimator in Eq. (18) is about 5.

Bandi and Russell (2004a) provide an alternative bias-correction in both the correlated noise case and in the $MA(1)$ case. For conciseness, here we only discuss the $MA(1)$ case (see also Zhang et al. (2004) in this case). As we point out above (see Eq. (14)), the bias of the realized variance estimator can be estimated consistently by computing an arithmetic average of the observed return data sampled at the highest frequencies. The bias-adjusted realized variance estimator is equal to

$$\widehat{V}^{db} = \widehat{V} - M \frac{1}{\widetilde{M}} \sum_{j=1}^{\widetilde{M}} r_{j\delta}^2, \tag{19}$$

where $\widetilde{M}$ is the number of observations in the full sample.[5] Bandi and Russell (2004) obtain the MSE of the estimator in Eq. (19) and compute the optimal sampling frequency M_{db}^* of the bias-corrected estimator in closed-form, i.e.,

$$M_{db}^* = \left(\frac{hQ}{2\mathbf{E}\left(\varepsilon^4\right) - 3\left(\mathbf{E}(\varepsilon^2)\right)^2} \right)^{1/2}. \tag{20}$$

Bandi and Russell (2004a) confirm Hansen and Lund's result that bias-correcting allows optimal sampling at higher frequencies than in the biased case while offering MSE improvements.

Oomen (2004b) extends the framework in Oomen (2004a) to the case of bias-corrected realized variance. Specifically, he studies the MSE properties of Zhou's estimator in Eq. (18) to the case of an underlying jump process of finite variation (as in Eq. (15)) and business time sampling. Using IBM and SPY transaction data over the period from January 2, 2003, to August 31, 2003, he confirms that *(i)* business time sampling can be beneficial in practise (Oomen (2004a)) and *(ii)* bias-correcting can induce a drop in the optimal sampling frequency as well as MSE gains. He finds optimal frequencies around 10 second and MSE gains around 60%. His results also suggest

[5] The sample second moment of the noise, i.e.,

$$\frac{1}{\widetilde{M}} \sum_{j=1}^{\widetilde{M}} r_{j\delta}^2$$

can be purged of residual contaminations induced by the equilibrium price variance by substracting from it a quantity defined as

$$\frac{1}{\widetilde{M}} \sum_{j=1}^{\widetilde{P}} r_{j\delta}^2,$$

where $\widetilde{P}$ is an appropriate number of low frequency returns calculated using 15 or 20-minute intervals, for instance.

that business time sampling might have a second-order impact on the estimation error of variance estimates, as measured by the estimator's MSE, when compared to a first-order bias correction (as in Eq. (18)).

4.4 Sub-sampling

Zhang et al. (2004) propose a methodology to consistently estimate integrated variance in the presence of $MA(1)$ microstructure noise.[6] Their method relies on sub-sampling. They begin by defining K non-overlapping sub-grids $G^{(i)}$ of the full grid of n arrival times with $i = 1, ..., K$. The first sub-grid starts from t_0 and takes every K-th arrival time, i.e., $G^{(1)} = (t_0, t_{0+K}, t_{0+2K}, ...,)$, the second sub-grid starts from t_1 and takes every K-th arrival time, i.e., $G^{(2)} = (t_1, t_{1+K}, t_{1+2K}, ...,)$, and so on. Given the ith sub-grid of arrival times, one can define the corresponding realized variance estimator as

$$\widehat{V}^{(i)} = \sum_{t_j, t_{j+} \in G^{(i)}} \left(p_{t_{j+}} - p_{t_j}\right)^2 \tag{21}$$

where t_j and t_{j+} denote consecutive elements in $G^{(i)}$. Estimation entails averaging the realized variance estimates obtained by using sub-grids and bias-correcting them. Define

$$\widehat{V}^{sub} = \frac{\sum_{i=1}^{K} \widehat{V}^{(i)}}{K} - 2\overline{n}\widehat{\mathbf{E}}(\varepsilon^2), \tag{22}$$

where $\overline{n} = \frac{n-K+1}{K}$, $\widehat{\mathbf{E}}(\varepsilon^2) = \frac{\sum_{j=1}^{n} \left(p_{t_{j+}} - p_{t_j}\right)^2}{n}$ is a consistent estimate of the second moment of the noise return, and $2\overline{n}\widehat{\mathbf{E}}(\varepsilon^2)$ is the required bias-correction (see the discussion in Section 3). Under Assumption 1 with $\alpha_t = 0$ and Assumption 2b (i.e., the case of $MA(1)$ noise), Zhang et al. (2004) show that, as $n \to \infty$ with $\frac{n}{K} \to \infty$, $\widehat{V}^{sub}$ is a consistent estimator of the integrated variance V over h. The rate of convergence of $\widehat{V}^{sub}$ to V is $n^{-1/6}$ and the asymptotic distribution is mixed-normal with an estimable asymptotic variance. Zhang et al. (2004) also provide an expression for the optimal number of sub-grids K, i.e.,

$$K = cn^{2/3} \tag{23}$$

with

$$c = \left(\frac{16 \left(\mathbf{E}(\varepsilon^2)\right)^2}{h \frac{8}{3} Q} \right)^{1/3}. \tag{24}$$

As illustrated by Zhang et al. (2004), both components of the proportionality factor c, namely $\mathbf{E}(\varepsilon^2)$ and Q, can be evaluated from the data. Specifically, $\mathbf{E}(\varepsilon^2)$ can be estimated by using a sample average of squared continuously-compounded returns sampled at the highest frequencies as indicated above. The quarticity term Q can be identified by using the procedure that Zhang et al.

[6] See also Aït-Sahalia et al. (2003) for a discussion of consistent maximum likelihood estimation of the *constant* variance of scalar diffusion processes in parametric models with microstructure noise.

(2004) lay out in their Section 6. Alternatively, as we discussed earlier, one could employ the BN-S quartic estimator, namely $\widehat{Q} = \frac{M}{3h} \sum_{j=1}^{M} \widetilde{r}_{j\delta}^{4}$ (BN-S (2002)), with $\delta = \frac{h}{M}$ chosen to coincide with either 15 or 20 minutes.

In recent work, under an assumed $MA(1)$ structure, Hansen et al. (2005) provide an analysis of the properties of kernel-based estimators for integrated variance that are similar to the estimator in Eq. (18). They show that these estimators are always inconsistent. Furthermore, they relate the kernel-based estimators to the sub-sampling estimator of Zhang et al. (2004) and show that the difference between the two is due to end effects (which are irrelevant in the context of stationary time series a là Newey-West, but appear to matter in this context). According to their new results, the end effects allow the sub-sampling estimator to be consistent at rate $n^{-1/6}$. Finally, Hansen et al. (2005) show how to modify kernel-based estimators to be consistent with a conjectured rate equal to $n^{-1/4}$.

In concurrent work, Zhang (2005) has extended the sub-sampling estimator of Zhang et al. (2004). Her new estimator achieves the best attainable rate, namely $n^{-1/4}$, and is robust to noise dependence.

5 The variance of microstructure noise: a consistency result

Even though the standard realized variance estimator is not a consistent estimator of the variance of the underlying equilibrium price, a rescaled version of the standard realized variance estimator is consistent for the variance of the noise return component. More generally, sample moments of the observed return data estimate moments of the underlying noise return process at high-frequencies (see Eq. (14) above). Bandi and Russell (2004a) discuss this result and use it to characterize the MSE of the conventional realized variance estimator.

While the realized variance literature focuses on the volatility features of the underlying equilibrium price, the empirical market microstructure research places emphasis on the other component of the observed price process in Eq. (1), namely the price frictions η. Such frictions can be interpreted in terms of transaction costs in that they constitute the difference between the observed price p and the corresponding equilibrium price p^*.[7] Hasbrouck (1993) and Bandi and Russell (2004c) provide related but different frameworks to use high-frequency transaction price data in order to estimate the second moment of the transaction cost η (rather than moments of ε as needed in the realized variance literature) under mild assumptions on the features of the price formation mechanism in Section 2. The implications of their results in measuring transaction costs are discussed in the following section. We start with a discussion of traditional approaches to transaction cost evaluation.

[7] Measuring the execution costs of stock market transactions and understanding their determinants is of importance to a variety of market participants, such as individual investors and portfolio managers, as well as regulators. In November 2000, the Security and Exchange Commission issued Rule 11 Ac. 1-5 requesting market venues to widely distribute (in electronic format) execution quality statistics regarding their trades.

6 The benefit of consistency: measuring market quality

6.1 Transaction cost estimates

Following Perold (1988), it is generally believed that an ideal measure of the execution cost of a trade should be based on the comparison between the trade price for an investor's order and the equilibrium price prevailing at the time of the trading decision. Although individual investors can plausibly construct this measure, researchers and regulators do not have enough information to do so (see Bessembinder (2003) for a discussion).

Most available estimates of transaction costs relying on high-frequency data hinge on the basic logic behind Perold's original intuition. Specifically, there are three measures of execution costs that have drawn attention in recent years, i.e., the so-called quoted *bid-ask half spread*, the *effective half spread*, and the *realized half spread.* The quoted bid-ask half spread is defined as half the difference between ask quote and bid quote. The effective half spread is the (signed[8]) difference between the price at which a trade is executed and the mid-point of the reference bid-ask quotes. As for the realized half spread, this measure is defined as the (signed) difference between the transaction price and the mid-point of a quote in effect some time after the trade.[9] In all cases, an appropriately chosen mid-point bid-ask quote is used as an approximation for the relevant equilibrium price.

The limitations of these measures of the cost of trade have been pointed out in the literature (the interested reader is referred to the special issue of the Journal of Financial Markets on transaction cost evaluation for updated discussions). The quoted bid-ask half spread, for example, is known to overestimate the true cost of trade in that trades are often executed at prices within the posted quotes. As for the effective and realized spreads, not only do they require the trades to be signed as buyer or seller-initiated but they also require the relevant quotes and transaction prices to be matched.

The first issue (i.e., assigning the trade direction) arises due to the fact that commonly used high-frequency data sets (the TAQ database, for instance) do not contain information about whether a trade is buyer or seller-initiated. Some data sets do provide this information (the TORQ database being an example) but the length of their time series is often insufficient. Naturally, then, a considerable amount of work has been devoted to the construction of algorithms intended to classify trades as being buyer of seller-initiated simply on the basis of transaction prices and quotes (see, for example, Lee and Ready (1991) and Ellis et al. (2000)). The existing algorithms can of course missclassify trades (the Lee and Ready method, for example, is known to categorize incorrectly about 15% of the trades), thereby inducing biases in the final estimates. Bessembinder (2003) and Peterson and Sirri (2003) contain a thorough discussion of the relevant issues.

The second issue (i.e., matching quotes and transaction prices) requires potentially arbitrary judgment calls. Since the trade reports are often delayed, when computing the effective spreads, for example, it seems sensible to compare the trade prices to mid-quotes occurring before the trade re-

[8]Positive for buy orders and negative for sell orders.

[9]The idea is that the traders possess private information about the security value and the trading costs should be assessed based on the trades' non-informational price impacts.

port time. The usual allowance is 5 seconds (see Lee and Ready (1991)) but longer lags can of course be entertained. As pointed out by Bessembinder (2003), it would appear appropriate to compare the trade prices to earlier quotes even if there were no delays in the reporting. This comparison would somehow incorporate the temporal difference between trading decision and implementation of the trade as in Perold's recommendation.

This said, there is a well-known measure, which can be computed using low frequency data, that does not require either the signing of the trades or the matching of quotes and transaction prices, i.e., *Roll's effective spread estimator* (Roll (1984)). Roll's estimator does not even rely on the assumption that the mid-point bid-ask quotes are good proxies for the unobserved equilibrium prices. The idea behind Roll's measure can be easily laid out using the model in Section 2. Write the model in transaction time. Assume

$$\eta_i = sI_i \tag{25}$$

where I_i equals 1 for a buyer-initiated trade and -1 for a seller-initiated trade with $p(I_i = 1) = p(I_i = -1) = \frac{1}{2}$. If $\alpha_t = 0$ in Assumption 1 and Assumption 2b is satisfied, then

$$\mathbf{E}(r, r_{-1}) = -s^2. \tag{26}$$

Equivalently,

$$s = \sqrt{-\mathbf{E}(r, r_{-1})}. \tag{27}$$

Thus, the constant width of the spread can be estimated consistently based on the negative first-order autocovariance of recorded stock returns.

Roll's estimator hinges on potentially restrictive assumptions. The equilibrium returns r^* are assumed to be serially uncorrelated. In addition, the microstructure frictions in the observed returns r follow a simplified $MA(1)$ structure with a constant cost of trade s. Finally, the estimator relies on the microstructure noise components being uncorrelated with the equilibrium prices.

6.2 Hasbrouck's pricing errors

Hasbrouck (1993) assumes a price formation mechanism which is identical to the mechanism in Section 2, Eq. (1). However, his set-up is in discrete-time and time is measured in terms of transaction arrival times. Specifically, the equilibrium price p^* is modelled as a random-walk while the η's, which may or may not be correlated with p^*, are mean-zero covariance stationary processes. Hence, Hasbrouck (1993) considerably relaxes the assumptions that are necessary to derive Roll's effective spread estimator.

He interprets the difference η between the transaction price p and the equilibrium price p^* as a *pricing error* impounding microstructure effects. The unconditional expectation of the pricing error is zero. However, conditional on a trader's identity, the pricing error is not fair game. A positive pricing error is a cost to a buyer while a negative pricing error is a cost to a seller.

Hasbrouck (1993) focuses on the standard deviation of the pricing error σ_η. He interprets it as a natural measure of market quality in that stocks whose transaction prices track the equilibrium price can be regarded as being stocks that are less affected by barriers to trade.

Using techniques that the macroeconometric literature has introduced to study nonstationary time series (like the observed price p) which can be expressed as the sum of a nonstationary component (here the equilibrium price p^*) and a residual stationary component (here the pricing error η), Hasbrouck (1993) provides lower bounds for σ_η. His empirical work focuses on NYSE stocks and employes transaction data collected from the Institute for the Study of Securities Markets (ISSM) tape for the first quarter of 1989. His estimated average bound for σ_η is equal to about 33 basis points. Under an assumption of normality, the corresponding average bound for the expected transaction costs $\mathbf{E}\,|\eta|' s$ is equal to about 26 basis points (namely $\frac{2}{\sqrt{\pi}}\sigma_\eta \approx 0.8\sigma_\eta$).

6.3 Full-information transaction costs

Bandi and Russell (2004c) define a notion of transaction cost (or pricing error in Hasbrouck's terminology) which they name *full-information transaction cost* or *FITC.* Their approach requires imposing more economic structure on the model in Section 2. They begin by noting that in a rational expectation set-up with asymmetric information two equilibrium prices can be defined: "the efficient price," i.e., the price that would prevail in equilibrium given public information, and the "full-information price," the price that would prevail in equilibrium given all private and public information. Both the efficient price and the full-information price are unobservable. The econometrician only observes transaction prices.

In this setting there are two sources of market inefficiency to consider. First, transaction prices deviate from the efficient price due to classical market microstructure frictions (See Stoll's presidential address to the AFA - Stoll (2000)). Second, the presence of asymmetric information induces deviations between the efficient price and the full-information price. Hasbrouck's approach (as described in Subsection 6.2), just like traditional approaches to transaction cost evaluation (Subsection 6.1), refer to the efficient price as the relevant equilibrium price. Hence, the above-mentioned methods account for the first source of market inefficiency. Full-information transaction costs are designed to account for both sources of inefficiency.

A cornerstone of market microstructure theory is that uninformed agents learn about existing private information from observed order flow (the interested reader is referred to the discussions in O'Hara (1995)). Since each trade carries information, meaningful revisions to the efficient price will be made regardless of the time interval between trade arrivals. Hence the efficient price is naturally thought of as a process changing discretely at transaction times. On the other hand, the full-information set, by definition, contains all information used by agents in their decisions to transact. Hence the full-information price is unaffected by past order flow. Barring occasional news arrivals to the informed agents the dynamic behavior of the full information price is expected to be relatively "smooth." As for the microstructure frictions, separate prices for buyers and sellers and discreteness of prices alone suggest that changes in the microstructure frictions from trade to trade are discrete in nature.

Bandi and Russell (2004c) write the model in Section 2 in transaction time. They add structure to the specification in Eq. (1) in order to account for the previously described properties of efficient price, full-information price, and microstructure noise. Specifically, they write

$$\begin{aligned} p_i &= p^*_{t_i} + \eta_i && (28) \\ &= p^*_{t_i} + \eta^{asy}_i + \eta^{fri}_i, && (29) \end{aligned}$$

where $p^*_{t_i}$ is now the full-information price, $p^*_{t_i} + \eta^{asy}_i$ is the discretely-evolving efficient price, and η^{fri}_i denotes conventional (discrete) microstructure frictions. The quantity η_i is a full-information transaction cost. It includes a standard friction component η^{fri}_i and an asymmetric information component η^{asy}_i.

As earlier in Section 2, one can rewrite the model in terms of observed returns, i.e.,

$$r_i = r^*_{t_i} + \varepsilon_i \tag{30}$$

where $r_i = p_i - p_{i-1}$, $r^*_{t_i} = p^*_{t_i} - p^*_{t_{i-1}}$ and $\varepsilon_i = \eta_i - \eta_{i-1}$. At very high-frequencies, the continuously-compounded return data (the r_i's) are dominated by return components that are induced by microstructure effects since the underlying full-information returns evolve smoothly in time. In fact, $r^*_{t_i} = O_p\left(\sqrt{\max|t_i - t_{i-1}|}\right)$ and $\varepsilon_i = O_p(1)$. In this context, Bandi and Russell (2004c) employ sample moments of the *observed* high-frequency return data to learn about moments of the *unobserved* effective cost of trade. They do so by using the informational content of observed return data whose full-information return component $r^*_{t_i}$ is largely swamped by the transaction cost component ε_i when sampling is conducted at the high frequencies at which transactions occur in practice. Assume that the covariance structure of the η's is such that $\mathbf{E}(\eta\eta_{-j}) = \theta_j \neq 0$ for $j = 1, ..., k < \infty$ and $\mathbf{E}(\eta\eta_{-j}) = 0$ for $j > k$. One can show that

$$\sigma_\eta = \sqrt{\left(\frac{1+k}{2}\right)\mathbf{E}(\varepsilon^2) + \sum_{s=0}^{k-1}(s+1)\mathbf{E}(\varepsilon\varepsilon_{-k+s})}. \tag{31}$$

As said, sample moments of the *unobserved* noise components can be estimated using high-frequency return data:

$$\widehat{\sigma}_\eta = \sqrt{\left(\frac{k+1}{2}\right)\left(\frac{\sum_{i=1}^{\widetilde{M}} r_i^2}{\widetilde{M}}\right) + \sum_{s=0}^{k-1}(s+1)\left(\frac{\sum_{i=k-s+1}^{\widetilde{M}} r_i r_{i-k+s}}{\widetilde{M}}\right)} \underset{\widetilde{M}\to\infty}{\overset{p}{\to}} \sigma_\eta, \tag{32}$$

where $\widetilde{M}$ is now the total number of transactions over a period. This result is robust to correlatedness in the underlying full-information price, presence of jumps in the full-information price, correlatedness between the full-information price and the remaining frictions as well as time-dependence in the frictions. Bandi and Russell (2004c) also suggest a finite sample adjustment to $\widehat{\sigma}_\eta$ in order to purge the estimates of the potential presence of a residual (full-information) variance component.

Bandi and Russell (2004c) call the quantity $\widehat{\sigma}_\eta$ *FITC*. In general, the *FITC's* are standard deviations. However, one can either assume normality of the η's (as in Hasbrouck (1993)) or use the approach in Roll (1984) to derive expected costs. In the former case, a consistent estimate of $\mathbf{E}\,|\eta|$ can be provided by $\frac{2}{\sqrt{\pi}}\widehat{\sigma}_\eta$. In the latter case, assume $\eta = sI$, where the random variable I, defined as in Subsection 6.1, represents now the direction (i.e., higher or lower) of the transaction price with respect to the full-information price and s is the full-information transaction cost. Then, $\widehat{\sigma}_{\overline{\eta}}$ consistently estimates s.

Bandi and Russell (2004c) find that the deviations of the efficient prices from the full-information levels, as determined by the existence of private information in the market place, can be as large as the departures of the transaction prices from the efficient prices. The latter, which are to be imputed to standard market microstructure frictions, are the focus of more conventional measures, like effective spreads. Using a sample of S&P 100 stocks over the month of February 2002, Bandi and Russell (2004c) report an average value for $\widehat{\sigma}_\eta$ equal to 12 basis points. Their average estimated $\mathbf{E}\,|\eta|$ is equal to about 9.6 basis points. This value is considerably larger that the corresponding average effective spread (i.e., about 6 basis points).

7 Directions for future work

7.1 The dynamic features of microstructure noise volatility

In keeping with the logic behind the vibrant and successful realized variance literature initiated by Andersen et al. (2001, 2003a) and BN-S (2002), the methods in Bandi and Russell (2004c) effectively render the volatility of the microstructure noise component observable. While the realized variance literature has placed emphasis on the volatility of the underlying true price process, one can focus on the other volatility component of the observed returns, i.e., the microstructure noise volatility. Treating the volatility of the noise component of the observed prices as being directly observable can allow one to address a broad array of fundamental issues. Some have a statistical flavor having to do with the distributional and dynamic properties of the noise variance and its relationship with the time-varying variance of the underlying price process. Some have an economic importance having to do with the dynamic determinants of the cost of trade. Since the most salient feature of the quality of a market is how much one has to pay in order of transact, much can be learned about the genuine market dynamics by exploiting the informational content of the estimated noise variances.

7.2 Portfolio choice and risk-management

Considerable importance has been recently placed on nonparametric variance estimation both in the absence and in the presence of market microstructure noise frictions. Surprisingly, with the exception of a thorough theoretical treatment in the frictionless case (BN-S (2004)), little theoretical work exists on the high-frequency estimation of covariances and betas when noise plays a role. The provision of methods that are intended to purge estimated covariances and betas of the impact of frictions represents a necessary next step for the practise of effective portfolio choice and risk management through high-frequency data.

7.3 Volatility and asset-pricing

Some recent work has been devoted to assessing whether stock market volatility is priced in the cross-section of stock returns. Being innovations in volatility correlated with changes in investment opportunities, this is a relevant study to undertake. Using model-free measures of volatility based on low frequency realized variance estimates, Ang et al. (2004) and Moise (2004) find that stocks with relatively higher exposure to volatility earn a negative risk premium. Volatility is high during recessions. Stocks whose returns covary with volatility are stocks which pay off during bad times. Investors are willing to pay a premium to hold them. The results in Ang et al. (2004) and Moise (2004) are robust to the use of alternative volatility measures. However, barring complications induced by the shorter observation spans of price data sampled at high frequencies, it is of interest, mainly for efficiency reasons, to evaluate the cross-sectional determinants of stock returns using high-frequency volatility estimates. In this context, microstructure issues ought to be accounted for.

Since individuals are likely to take into account the cost of acquiring and rebalancing their portfolios, expected stock returns should also embed transaction costs in equilibrium. This observation has given rise to a convergence between market microstructure work on price determination and asset pricing in recent years (the interested reader is referred to the recent survey of Easley and O'Hara (2002)). The current attempts to characterize the cross-sectional relationship between expected stock returns and cost of trade rely on liquidity-based theories of transaction cost determination (Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), Datar, Naik, and Radcliffe (1998), Hasbrouck (2003), and Pastor and Stambaugh (2003), among others). Alternatively, they rely on information-based approaches to the same issue (Easley et al. (2002)). Much remains to be done. Full-information transaction costs, for example, can be regarded as providing a bridge between both arguments. Bandi and Russell (2005) are analyzing the cross-sectional dependence between expected stock returns and full-information transaction costs in current work (Bandi and Russell (2005)).

Generally speaking, the convergence between market microstructure theory and methods and asset-pricing is in its infancy. We are convinced that the recent interest in microstructure issues in the context of volatility estimation is providing and will continue to provide a strong boost to this inevitable process of convergence.

References

[1] Aït-Sahalia, Y., P. Mykland, and L. Zhang (2004). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies.* Forthcoming.

[2] Amihud, Y., and H. Mendelson (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17, 223-249.

[3] Andersen, T.G., T. Bollerslev, and F.X. Diebold (2002). Parametric and nonparametric measurements of volatility. In Y. Aït-Sahalia and L.P. Hansen (Eds.) *Handbook of Financial Econometrics.* Elsevier, North-Holland. Forthcoming.

[4] Andersen, T.G., T. Bollerslev, F.X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics,* 61, 43-76.

[5] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (1999). (Understanding, optimizing, using and forecasting) Realized volatility and correlation. *Working paper.*

[6] Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (2000). Great Realizations. *Risk Magazine,* 105-108.

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

[8] Andersen, T.G., T. Bollerslev, and N. Meddahi (2003b). Correcting the errors: volatility forecast evaluation based on high frequency data and realized volatilities. *Working paper.*

[9] Andrews, D.W.K. and J.C. Monahan (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica,* 60, 953-966.

[10] Ang, A., R. Hodrick, Xing, and X. Zhang (2004). The cross-section of volatility and expected returns. *Journal of Finance.* Forthcoming.

[11] Bai, X., J.R. Russell, and G. Tiao (2004). Effects on non-normality and dependence on the precision of variance estimates using high-frequency data. *Working paper.*

[12] Bandi, F.M., and J.R. Russell (2004a). Realized variance, microstructure noise, and optimal sampling. *Working paper.*

[13] Bandi, F.M., and J.R. Russell (2004b). Separating market microstructure noise from volatility. *Working paper.*

[14] Bandi, F.M., and J.R. Russell (2004c). Full-information transaction costs. *Working paper.*

[15] Bandi, F.M., and J.R. Russell (2005). Expected stock returns and full-information transaction costs. *Working paper.*

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

[17] Barndorff-Nielsen, O.E., and N. Shephard (2004). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics. *Econometrica,* 72, 885-925.

[18] Barndorff-Nielsen, O.E., P. Hansen, A. Lunde and N. Shephard (2005). Regular and modified kernel-based estimators of integrated variance: the case with independent noise. *Working paper.*

[19] Bessembinder, H. (2003). Issues in assessing trade execution costs. *Journal of Financial Markets,* 6, 233-257.

[20] Bollen, B., and B. Inder (2002). Estimating daily volatility in financial markets utilizing intraday data. *Journal of Empirical Finance,* 9, 551-562.

[21] Brennan, M.J., and A. Subrahmanyam (1996). Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. *Journal of Financial Economics,* 41, 441-464.

[22] Chung, K.L., and R.J. Williams (1990). *Introduction to Stochastic Integration.* Second Edition. Birkhäuser.

[23] Datar, V.T., N.Y. Naik and R. Radcliffe (1998). Liquidity and stock returns: an alternative test. *Journal of Financial Markets,* 1, 203-219.

[24] Easley D., and M. O'Hara (2002). Microstructure and asset pricing. In G. Constantinides, M. Harris, and R. Stulz (Eds.). *Handbook of Financial Economics.* Elsevier, North Holland. New York.

[25] Easley D., S. Hvidkjaer, and M. O'Hara (2002). Is information risk a determinant of asset returns? *Journal of Finance*, 57, 2185-2221.

[26] Ellis, K., R. Michaely, and M. O'Hara (2000). The accuracy of trade classification rules: evidence from Nasdaq. *Journal of Financial and Quantitative Analysis,* 35, 529-552.

[27] Fleming, J., C. Kirby, and B. Ostdiek (2001). The economic value of volatility timing. *Journal of Finance,* 56, 329-352.

[28] Fleming, J., C. Kirby, and B. Ostdiek (2003). The economic value of volatility timing using "realized volatility." *Journal of Financial Economics,* 67, 473-509.

[29] Hansen, P.R. and A. Lunde (2004a). An unbiased measure of realized variance. *Working paper.*

[30] Hansen, P.R., and A. Lunde (2004b). Realized variance and market microstructure noise. *Working paper.*

[31] Hasbrouck, J. (1993). Assessing the quality of a security market: a new approach to transaction cost measurement. *Review of Financial Studies*, 6, 191-212 .

[32] Hasbrouck, J. (1996). Modelling market microstructure time series. In G.S. Maddala and C.R. Rao (Eds.) *Handbook of Statistics, Vol. 14, 647-692.* Elsevier, North-Holland. Amsterdam.

[33] Hasbrouck, J. (2003). Trading costs and returns for U.S. securities: the evidence from daily data. *Working paper.*

[34] Lee, C. and M. Ready (1991). Inferring trade direction from intraday data. *Journal of Finance,* 46, 733-746.

[35] Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics,* 17, 475-508.

[36] Moise, C. (2003). Is market volatility priced? *Working paper.*

[37] Newey, W. and K. West (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica,* 55, 703-708.

[38] O'Hara, M. (1995). *Market Microstructure Theory.* Blackwell Publishers Ltd, Oxford.

[39] Oomen, R.C.A. (2004a). Properties of realized variance for a pure jump process: calendar time sampling versus business time sampling. *Working paper.*

[40] Oomen, R.C.A. (2004b). Properties of bias-corrected realized variance in calendar time and business. *Working paper.*

[41] Pastor, L. and R.F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy,* 111, 642-685.

[42] Perold, A (1988). The implementation shortfall. *Journal of Portfolio Management,* 14, 4-9.

[43] Peterson, M. and E. Sirri (2002). Evaluation of the biases in execution cost estimates using trade and quote data. *Journal of Financial Markets,* 6, 259-280.

[44] Press, J.S. (1967). A compound events model for security prices. *Journal of Business,* 40, 317-335.

[45] Roll, R. (1984). A simple measure of the effective bid-ask spread in an efficient market. *Journal of Finance,* 39, 1127-1139.

[46] Stoll, H.R. (2000). Friction. *Journal of Finance,* 55, 1479-1514.

[47] Zhang, L., P. Mykland, and Y. Aït-Sahalia (2004). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association.* Forthcoming.

[48] Zhang, L. (2005). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Working paper.*

[49] Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics,* 14, 45-52.

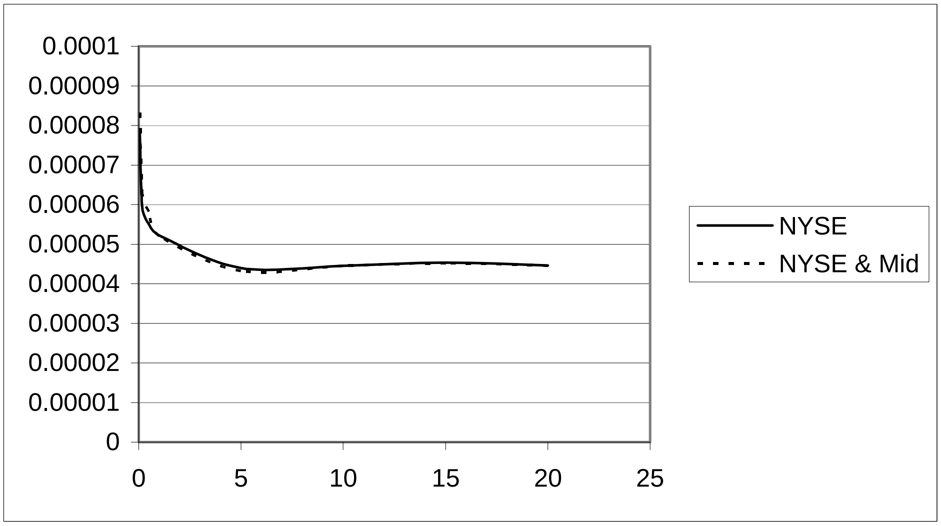

Figure 1. Volatility signature plot for IBM from midquotes using i) NYSE only and ii) NYSE and Midwest.

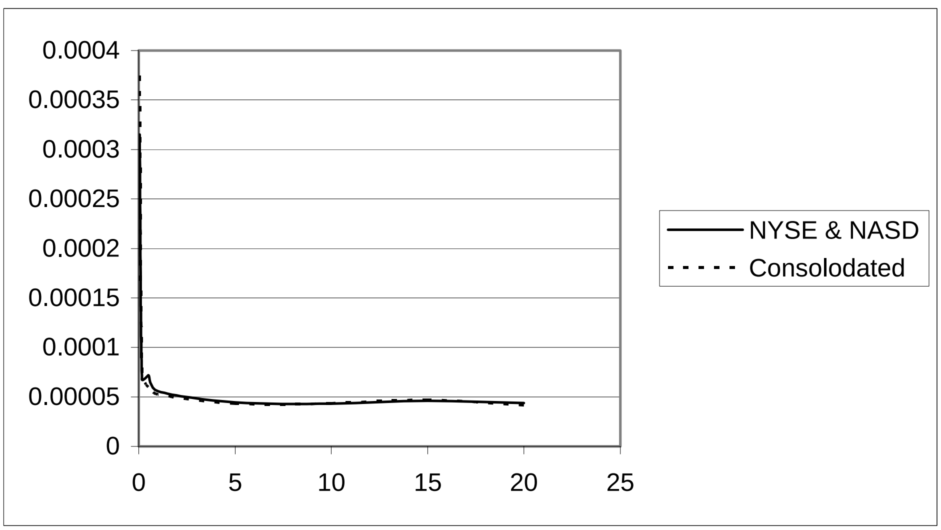

Figure 2. Volatility signature plot for IBM from mid-quotes using i) NYSE and NASDAQ and ii) the consolidated market.

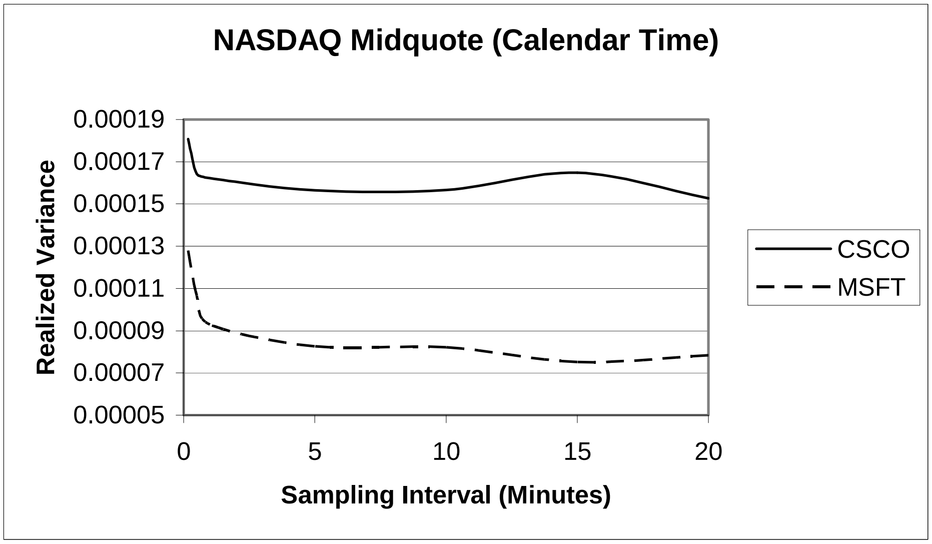

Figure 3. Volatility signature plots for the two NASDAQ stocks Cisco Systems and Microsoft using mid-quotes.

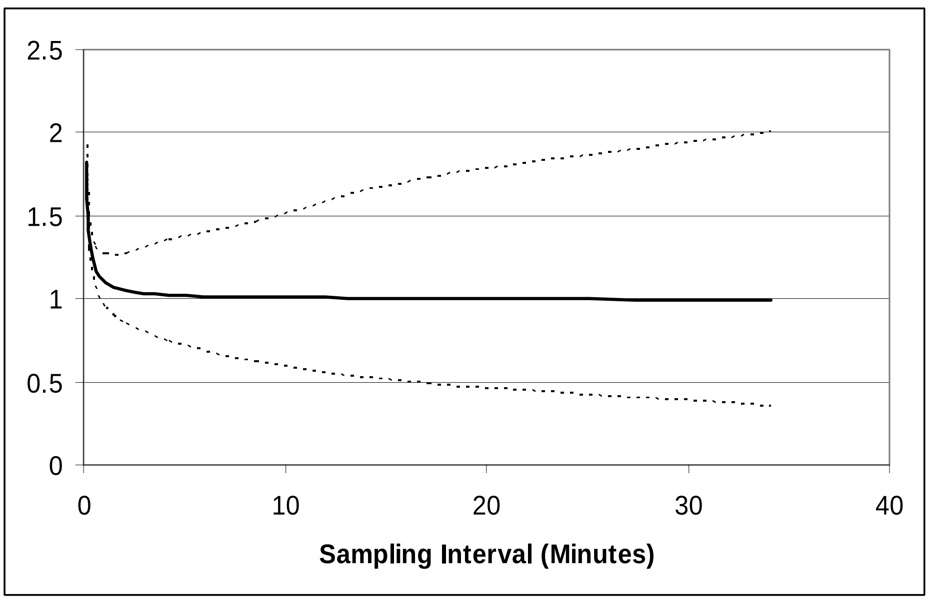

Figure 4. Realized Variance signature plot and empirical 95% interval for the simulated data.