Discrete Sine Transform for Multi-Scales Realized Volatility Measures *

Giuseppe Curci †, Fulvio Corsi ‡

This version: April 2010

Abstract

In this study we present a new realized volatility estimator based on the combination of the Multi-Scale regression and *Discrete Sine Transform* (DST) approaches. Multi-Scales estimators similar to that recently proposed by Zhang (2006) can, in fact, be constructed within a simple regression based approach by exploiting the linear relation existing between the market microstructure bias and the realized volatilities computed at different frequencies. We show how such powerful Multi-Scale regression approach can also be applied in the context of the Zhou (1998) or DST orthogonalization of the observed tick-by-tick returns. Providing a natural orthonormal basis decomposition of observed returns, the DST permits to optimally disentangle the volatility signal of the underlying price process from the market microstructure noise. Robustness of the DST approach with respect to more general dependent structure of the microstructure noise is also analytically shown. The combination of Multi-Scale regression approach with DST gives a Multi-Scales DST realized volatility estimator close in efficiency to the optimal Cramer-Rao bounds and robust against a wide class of noise contaminations and model misspecifications. Monte Carlo simulations based on realistic models for price dynamics and market microstructure effects, show the superiority of DST estimators, compared to alternative volatility proxies for a wide range of noise to signal ratios and different types of noise contaminations. Empirical analysis based on six years of tick-by-tick data for S&P 500 index-future, FIB 30, and 30 years U.S. Treasury Bond future, confirms the accuracy and robustness of DST estimators on different types of real data.

JEL classification: C13; C22; C50; C80
Keywords: High frequency data; Realized Volatility; Market Microstructure; Bias correction

* Earlier versions of this paper were circulated under the title “A Discrete Sine Transform Approach for Realized Volatility Measurement”

† Then Professor at Department of Physics, ”Enrico Fermi”, University of Pisa.

‡ University of Lugano and Swiss Finance Institute, Via Buffi 13, CH-6904 - Lugano. Email: fulvio.corsi@lu.unisi.ch.

1 Introduction

Asset returns volatility is a central feature of many prominent financial problems such as asset allocation, risk management and option pricing. Recently a nonparametric approach to develop ex-post observable proxies for the daily volatility has been proposed, the so called *Realized Volatility measures*.

In its standard form realized volatility is simply the sum of squared high-frequency returns over a discrete time interval of typically one day, i.e. the second uncentered sample moment of high-frequency returns. This idea traces back to the seminal work of Merton (1980) who showed that the integrated variance of a Brownian motion can be approximated to an arbitrary precision using the sum of intraday squared returns. More recently a series of papers (Andersen, Bollerslev, Diebold and Labys 2001, 2003 and Barndorff-Nielsen and Shephard 2001, 2002a 2002b, 2005 and Comte and Renault 1998) has formalized and generalized this intuition by applying the quadratic variation theory to the broad class of special (finite mean) semimartingales. In fact, under very general conditions the sum of intraday squared returns converges, as the sampling frequency increases, to the notional volatility over the day. Thus, realized volatility provides us, in principle, with a consistent nonparametric measure of the notional volatility.

In practice, however, empirical data differs in many ways from the frictionless continuous-time price process assumed in those theoretical studies. Beside the obvious consideration that a continuous record of prices is not available, the presence of market microstructure effects prevent the application of the limit theory necessary to achieve consistency of the realized volatility estimator. The main sources of microstructure effects are the bid-ask bounce[1] and price discreteness. As already noted by Roll (1984) and Blume and Stambaugh (1983), bid-ask spreads produce negative first-order autocovariances in observed price changes. Similarly, if one makes the assumption that observed prices are obtained by rounding underlying true values, Glottlieb and Kalay (1985) and Harris (1990) showed that price discreteness induces negative serial covariance in the observed returns.

Therefore, microstructure noise induces a non-zero autocorrelation in the returns process which makes no longer true that the variance of the sum is the sum of the variances. Hence, market microstructure introduces a bias that grows as the sampling frequency increases. Formal studies of the impact of microstructure noise on realized volatility measure have been made by Bandi and Russell (2005), Aït-Sahalia, Mykland and Zhang (2005) and Hansen and Lunde (2006).

Earlier attempts to directly correct for the microstructure effects at the tick-by-tick level were the first order serial covariance correction proposed by French and Roll (1986), Harris (1990) and Zhou (1996)[2] and the exponential moving average (EMA) filtering of Corsi, Zumbach, Müller and Dacorogna (2001) and Zumbach, Corsi and Trapletti (2002). However, the first type of estimator suffers from the possibility to become negative while the second one is a non-local estimator which adapts only slowly to changes in the properties of the pricing error component. Moreover, both estimators correct

[1] Studies on the bid-ask spread are largely developed within the framework of quote-driven markets. However, the bid-ask spread is not unique to the dealer markets: Cohen et al. (1981) and Glosten (1994) establish the existence of the bid-ask spread also in a limit-order market because of transaction costs and asymmetric information.

[2] More recently, this approach has been revived by Oomen (2005) and Hansen and Lunde (2006).

only for the bias deriving from the first lag of the return autocorrelation function, while they are very sensitive to non zero higher lag coefficients.

In fact, the presence of significant autocorrelation at lag lengths greater than one and the possibility that each trading day may be characterized by different autocorrelation structures make the filtering problem rather complex. In theory, this problem could be tackled by a fully parametric approach. Parametric higher order covariance correction has been proposed, for instance, by Bollen and Inder (2002) which makes use of a series of AR models selected on the basis of the Schwarz BIC criteria and by Hansen and Lunde (2004) which employ MA(q) filters where q changes with the returns frequency so to keep the time spanned by the autocorrelation window constant. More recently, Barndorff-Nielsen, Hansen, Lunde and Shephard (2008) proposed a modified kernel-based estimator which is asymptotically optimal. Concurrently, Zhang, Mykland and Aït-Sahalia (2005) proposed an estimator based on overlapping subsampling schemes and an appropriate combinantion of two realized volatilities computed at two different time scales. Recently, Zhang (2006) has generalized the Two Scales estimator to a multiple time scales estimator that combines realized volatilities computed at more than two return frequencies and reaches the same asymptotic efficiency of the kernel-based estimator. Our approach will follow the direction of this *Multi-Scales* methodology[3].

In this paper a new realized volatility measures based on the combination of the Multi-Scales linear regression approach and *Discrete Sine Transform* (DST) orthogonalization are presented. Multi-Scales estimators similar to that recently proposed by Zhang (2006) can, in fact, be constructed within a simple regression based approach (similar in spirit to the recent one suggested by Phillips and Yu 2006) by exploiting the linear relation existing between the market microstructure bias and realized volatilities computed at different frequencies. We show that the same Multi-Scales regression idea can be applied in the context of the Zhou (1998) or DST orthogonalization of the observed high frequency returns. The motivation for the employment of the DST approach rests on its ability to decorrelate signal for data exhibiting MA type of behaviour which arises naturally in discrete time models of tick-by-tick returns. In fact, DST diagonalizes MA(1) processes exactly and MA(q) approximately. Hence, this nonparametric DST approach, turns out to be very convenient as it provides an orthonormal basis which permits to optimally (in a linear sense) extract the volatility signal hidden in the noisy tick-by-tick return series. Moreover, thanks to the DST results, it is possible to derive a closed form expression for the Cramer-Rao bounds of an MA(1) process and hence to evaluate the absolute efficiency of the estimators. We show that this approach produces robust and accurate results even in the presence of not i.i.d. microstructure noise which leads to more general MA(q) processes for the tick-by-tick returns. It is then robust against a wide class of noise contaminations and model misspecifications.

The rest of the paper is organized as follows. Section 2 reviews a model for the tick-by-tick observed price process and its DST orthogonalization and presents both the Multi-Scales and Multi-Scales DST approaches. It also discusses absolute efficiency of volatility estimation for MA(1) process deriving the analytical expression of the Cramer-Rao bounds and analyzes the robustness of the DST

[3]Barndorff-Nielsen, Hansen, Lunde and Shephard (2004) show that a direct link between the Multi-Scales and the kernel-based estimators exists.

estimators with respect to more general noise contaminations. Section 3 outlines the setup for the Monte Carlo simulations and compares the performance of Multi-Scales DST estimators togheter with other alternative realized volatility estimators. Section 4 reports the results of the application of DST estimators on empirical data. Section 5 concludes.

2 Definitions and properties of DST volatility estimators

2.1 Price process with microstructure noise

As described in Hasbrouk (1993, 1996), a general way to model the impact of various sources of microstructure effects is to decompose the observed price into the sum of two unobservable components: a martingale component representing the informationally efficient price process and a stationary pricing error component expressing the discrepancy between the efficient price and the observed one. The dynamics of the true latent price can be modelled as a general continuous time Stochastic Volatility (SV) process[4]

$$dp^*(t) = \mu(t)dt + \sigma(t)dW(t) \tag{1}$$

where $p^*(t)$ is the logarithm of the true instantaneous price, $\mu(t)$ is the finite variation process of the drift, $dW(t)$ is a standard Brownian motion, and $\sigma(t)$ is the instantaneous volatility. For this diffusion, the notional or actual variance is equivalent to the integrated variance for the day t (IV_t)[5] which is the integral of the instantaneous variance of the underling true process $\sigma^2(t)$ over the one day interval $[t-1;t]$, i.e. $IV_t = \int_{t-1}^{t} \sigma^2(\omega)d\omega$.

The observed (logarithmic) price, being recorded only at certain intraday sampling times and contaminated by market microstructure effects, is instead a discrete time process described in the "intrinsic transaction time" or "tick time"[6] denoted with the integer index n:

$$p_{t,n} = {p^*}_{t,n} + \eta_t \omega_{t,n} \tag{2}$$

where $p^*_{t,n}$ is the unobserved true price at intraday sampling time n in day t and $\eta_t \omega_{t,n}$ represents the pricing error component with η_t the size of the perturbation and $\omega_{t,n}$ a standardized noise. Depending on the structure imposed on the pricing error component, many structural models for microstructure effects could be recovered. Here we take a more statistical perspective assuming ω to simply be a zero mean nuisance component independent of the price process. Initially the assumption of an i.i.d. noise process for ω is made while it will be subsequently relaxed to allow for more general dependence structure.

According to the Mixture of Distribution Hypothesis originally proposed by Clark (1973) and extended and refined in numerous subsequent works, the price process observed under the appropriate

[4]Alternatively, a pure jump process as the compound Poisson process proposed by Oomen (2006) could be employed to model the dynamics of the true price process.

[5]We use the notation (t) to indicate instantaneous variable while subscript t denote daily quantities.

[6]That is, a time scale having the number of trades as its directing process (here we don't make the distinction between tick time and transaction time).

transaction (or tick) time can be interpreted as a subordinate stochastic process which has been shown to properly accommodate for many empirical regularities.[7] In the following, we assume that observing the process in tick time, instead of the common physical time, determines a stochastic time change (see Ané and Geman 2000) that transforms a stochastic volatility in physical time into a process in tick time where volatility is locally approximately constant. In other words, we consider (as in Ané and Geman 2000) the number of trades as a good approximation to the true stochastic clock which allows one to recover the normality and homoskedasticity of asset returns. For our purposes, we need to assume that this approximation holds only locally over a small window of few ticks. In fact, although for the ease of exposition we will describe the model having the efficient price following a Brownian motion in tick time with volatility changing from day to day, the actual implementation of DST approach would only formally require the volatility in tick time to be constant over a small window of very few ticks (20 or 30). Moreover, the robustness of the DST approach against violation of the assumption of homoskedasticity in tick time is checked in the simulation study by explicitly employing an heteroskedastic DGP for the true tick-by-tick price process.

Computing daily volatility in tick or transaction time also presents several practical advantages (see Oomen 2006 for a detailed comparison of different sampling schemes). Intuitively, in tick time all observations are used so that no information is wasted. The interpolation error and noise arising from the construction of the artificial regular grid is avoided. Moreover, using a tick time grid the underlying price process tends to be sampled more frequently when the market is more active, that is, when it is needed more because the price moves more.

The observed tick-by-tick return $r_{t,n}$ of day t at time n can then be decomposed as

$$r_{t,n} = \sigma_t {\epsilon^*}_{t,n} + \eta_t \left(\omega_{t,n} - \omega_{t,n-1}\right) \tag{3}$$

where the unobserved innovation of the efficient price ${\epsilon^*}_{t,n}$ and the pricing error $\omega_{t,n}$, are independent IID $(0,1)$ processes. Hence, for each day the tick-by-tick return process is a MA(1) with $\mathbb{E}\left(r_{t,n}\right) = 0$ and autocovariance function given by

$$\mathbb{E}\left[r_{t,n} r_{t,n-h}\right] = \begin{cases} \sigma_t^2 \quad + \quad 2\eta_t^2 & \text{for } h = 0 \\ - \ \eta_t^2 & \text{for } h = 1 \\ 0 & \text{for } h \geq 2 \end{cases} \tag{4}$$

where σ_t^2 represents the tick-by-tick variance of the unobserved true price for day t and η_t^2 the extra variance in the observed returns coming from the market microstructure noise observed during that day.

Under this assumption, the integrated variance of day t simply becomes $IV_t = N_t \sigma_t^2$, with N_t the number of ticks occurred in day t. Since we will compute the Realized Volatility for each day separately, in the following, the daily subscript t will be suppressed to simplify the notation.

[7] In tick time even a simple constant volatility process can reproduce stylized facts observed in physical time such as heteroskedasticity, volatility clustering, fat tails and others. Hence, the hypothesis of homoskedastic processes in tick time is far less restrictive compared to the same hypothesis made for processes defined in physical time where this assumption would be clearly violated by the empirical data.

2.2 The Discrete Sine Transform

Considering the vector of M tick-by-tick observed returns $R(M,n) = [\ r_n \quad r_{n-1} \quad \cdots \quad r_{n-M+1}\ ]^\top$, we develop a Principal Component Analysis[8] (PCA) of the associated variance-covariance matrix $\Omega^{(M)} = \mathbb{E}\left(R(M,n)\, R(M,n)^\top\right)$, which is a tridiagonal matrix of the form:

$$\Omega^{(M)} = \begin{bmatrix} \sigma^2 + 2\eta^2 & -\eta^2 & & \\ -\eta^2 & \sigma^2 + 2\eta^2 & \ddots & \\ & \ddots & \ddots & -\eta^2 \\ & & -\eta^2 & \sigma^2 + 2\eta^2 \end{bmatrix}$$

By solving the eigenvalue equation $\Omega^{(M)}\varphi_m^{(M)} = \lambda_m^{(M)}\varphi_m^{(M)}$ with $m = 1, 2, \ldots, M$, it can be shown (Elliot 1953 and Gregory and Karney 1969) that the eigenvalues of $\Omega^{(M)}$ are given by

$$\lambda_m^{(M)} = \sigma^2 + 4\eta^2 \sin^2 \frac{\pi m}{2(M+1)} \tag{5}$$

with $0 < \lambda_1^{(M)} < \lambda_2^{(M)} < \cdots < \lambda_M^{(M)}$. Therefore, the eigenvalues of the DST components are ordered, separated and all non degenerate. The corresponding eigenvectors are

$$\varphi_m^{(M)}(k) = \sqrt{\frac{2}{M+1}} \sin \frac{\pi m k}{M+1} \qquad k = 1, 2, \ldots, M \tag{6}$$

The remarkable fact is that, unlike common situations, the eigenvectors ($\varphi_m^{(M)}$) of a MA (1) process are universal and they coincide with the orthonormal basis used in the Discrete Sine Transform (DST). Hence, such nonparametric orthogonalization is very useful for the analysis of high frequency return data as it provides an universal basis to optimally (in a linear sense) decorrelate the price signal from market microstructure noise. This orthogonalization has been applied on high frequency data also by Zhou (1998).

According to the PCA interpretation, the simple and computationally fast DST of the returns

$$c_m^{(M)}(n) = \sum_{k=1}^{M} \varphi_m^{(M)}(k)\, r_{n-k+1}$$

acts as a projector of the signal into its principal components and the variance of the DST components are directly the eigenvalues of the variance-covariance matrix:

$$\mathbb{E}\left(c_m^{(M)}(n)\, c_m^{(M)}(n)\right) = \left(\varphi_m^{(M)}\right)^\top \Omega^{(M)}\varphi_m^{(M)} = \lambda_m^{(M)} = \sigma^2 + 4\eta^2 \sin^2 \frac{\pi m}{2(M+1)}$$

Since we are interested in the permanent component of volatility the idea is to consider the projection of the returns on the minimal principal component which is the one less contaminated by the transient volatility coming from the microstructure noise. Therefore, in place of measuring volatility on the raw

[8] Also known as Karhunen-Loéve expansion or Hotelling transformation.

return series, we compute it on the minimal component series obtained by the DST filter, i.e. taking the mean value of the square of the DST component associated with the minimal eigenvalue of the covariance matrix:

$$\hat{\sigma}^2_{min-DST} = \left(\sum_{k=1}^{M} \varphi^{(M)}_{min}(k)\, r_{n-k+1} \right)^2$$

Having an estimate of the average volatility of the tick-by-tick returns for a given day, the corresponding daily volatility is readily obtained by rescaling σ^2 with the number of ticks occurred in that day. We term this volatility estimator the *Minimal DST estimator* and we will use it as the building block for our pre-filtered multi-scales estimator defined later.

Being

$$\mathbb{E}\left[\hat{\sigma}^2_{min-DST}\right] = \sigma^2 + 4\eta^2 \sin^2 \frac{\pi}{2\,(M+1)} \tag{7}$$

$\hat{\sigma}^2_{min-DST}$ is asymptotically unbiased for σ^2 since for large M the bias coming from the microstructure noise vanishes as $\sigma^2_M \simeq \sigma^2 + \eta^2 \dfrac{\pi^2}{M^2}$. This clearly shows how the aggregation on the minimal component decreases the impact of the microstructure noise at a much higher speed compared with the standard aggregation of returns. In fact, in this second case, the bias is reduced at the rate M while on the minimal DST component the bias is cut down at rate M^2, allowing to substantially increase the "unbiased return frequency" and then improving the precision of the volatility estimation.

To judge the stability and robustness of the DST filter with respect to time-dependent noise, we now relax the i.i.d. assumption and analyze the behavior of the DST filter under a more general MA$(q-1)$ structure for the noise. In the presence of an MA$(q-1)$ dependent noise, the observed return in tick time becomes an MA(q) process which can be written (by generalizing equation 3) as

$$r_n = \sigma {\epsilon^*}_n + \sum_{i=1}^{q} \eta_i \left(\omega^{(i)}_n - \omega^{(i)}_{n-i} \right)$$

with ${\epsilon^*}_n \sim \mathrm{IID}\,(0,1)$ and $\omega^{(i)}_n \sim \mathrm{IID}\,(0,1)$. It can be shown (see Appendix A) that in this more general case the variance of the Minimal DST component can be approximated as $\sigma^2_M \simeq \sigma^2 + \frac{\pi^2}{M^2} \sum_{i=1}^{q} (i\,\eta_i)^2$ which, as before, clearly shows that the bias coming from higher order autocorrelations is also cut down at the same rate M^2, guaranteeing the robustness of the DST estimators respect to a wide class of noise contaminations and model misspecifications.

It should be noted however, that the presence of a dependent noise process could, in some cases, be an artificial result of the construction of the equidistant series in physical time. In fact, the time deformation induced by the transformation from a tick time scale to a physical one, can transform an MA(1) process into an MA(q) or ARMA(p,q). In other words, the time deformation induced by the equidistant grid construction could have the effect of spreading the mass of the first autocorrelation lag onto higher order lags[9]. This possible artificial increase of the autocorrelation order induced by

[9] See Corsi et al. (2001) for an empirical example and Oomen (2006) for a detailed theoretical analysis.

the regular grid construction is, in fact, an additional important reason to favor, in the computation of the realized volatility, the use of a tick time scale instead of the commonly used regular grid in physical time.

2.3 The Multi-Scales Least Square estimator

Recently Zhang, Mykland and Aït-Sahalia (2004) have introduced the "Two-Scales" estimator while Zhang (2006) has generalized this approach to a "Multi-Scales" estimator that combine realized volatilities computed at more than two return frequencies. Here, we present a different approach to the construction of realized volatility estimators computed with multiple time scales. This approach is more similar in spirit to the one recently proposed by Phillips and Yu (2006). It is important to note that throughout the paper, in order to reduce estimation errors and assure consistency of the estimators, the computation of any variance estimator at any level of aggregation, is always performed by adopting a full overlapping scheme i.e. (using the terminology introduced by Zhang et al. 2004) by subsampling and averaging.

Under the assumption of i.i.d. noise the conditional expectation of the daily realized variance $RV^{(k_j)}$ computed with observed returns of different tick-lengths k_j is

$$\mathbb{E}\left[RV^{(k_j)}\right] = IV + 2N^{(k_j)}\,\eta^2 \tag{8}$$

where $N^{(k_j)}$ is the number of k_j-returns in the day. Hence, a consistent and unbiased estimator can be obtained by computing the realized variance at different frequencies[10] k_j and then estimating IV and η^2 by means of a simple linear regression of $RV^{(k_j)}$ on $N^{(k_j)}$[11]. We will denote this class of estimators as *Multi-Scales Least Square* estimators.

It is interesting to note that applying this Multi-Scales Least Square approach to only two different frequencies k_1 and k_2, one gets a very simple linear system of two equations in two unknowns (IV and η^2) which can be directly solved, giving as estimator of IV

$$TS = \frac{\alpha RV^{(k_2)} - RV^{(k_1)}}{\alpha - 1} \tag{9}$$

where $\alpha = N^{(k_1)}/N^{(k_2)}$ is the ratio between the number of returns sampled at the "base" frequency k_1 and that at the lower "auxiliary" frequency k_2, while TS is exactly the expression of the Two-Scales estimator for serially dependent noise (with the small sample bias correction) recently proposed by Aït-Sahalia, Mykland and Zhang (2006) as an extention of the Zhang et al. (2005) estimators. Therefore, this alternative "Jack Knife style" derivation of the Aït-Sahalia et al. (2006) estimator shows that the Multi-Scales Least Square approach can be seen as another natural generalization of the Two-Scales estimator to more than just two sampling frequencies.

Irrespective of its simplicity, this regression approach represents a powerful and flexible way of exploiting the information contained at many different data frequency. It also provides an alternative

[10] As aforementioned, for returns with a length in ticks $k_j > 1$, subsampling and averaging (i.e. a full overlapping scheme) is adopted.

[11] This regression approach has been recently applied also by Nole and Voev (2008).

way of estimating the variance of the microstructure noise η^2 to the commonly used first lag tick return covariance (i.e. $\widehat{\eta}^2 = -\mathbb{E}\left[r_{t,n} r_{t,n-h}\right]$) or the $\widetilde{\eta}^2 = RV/2M$.

Moreover, it is robust to the presence of endogenous noise. In fact, it can be shown (Appendix B) that in the presence of correlation between microstructure noise and true price process modeled in the form $p_n = p_n^* + \eta_n \omega_n + \rho\, p_n^*$ the regression equation (8) becomes

$$\mathbb{E}\left[RV^{(k_j)}\right] = IV + N^{(k_j)}(\rho(2+\rho)\sigma^2 + 2\eta^2) \tag{10}$$

hence, the Multi-scale regression based approach still provide consistent and unbiased estimator of IV (although the estimator of η^2 is now biased).

It would be possible to also deal with not i.i.d. noise by cutting from the regression those higher frequencies k_j not belonging to the linear part in the plot (analogous to the volatility signature plot) of $\mathbb{E}\left[RV^{(k_j)}\right]$ against the number of observations $N^{(k_j)}$.

About the choice of scales in the Multi-Scale RV, in theory using all scales should lead to the maximum gathering of information and hence to optimality. In practice however, neighborhood frequencies contain very similar information, and collecting all of them can lead to multicollinearity problem and other computational burdens. Therefore, for practical purposes, an heuristic choice of sparse frequencies considerably simplify the implementations of the estimator and does not lead to a significant loss of information already with a moderate number of scales.[12]

2.4 The Multi-Scales DST estimator

The powerful idea of exploiting linear relations among realized volatility measures computed at different aggregation frequencies by means of simple linear regressions, could also be applied in the DST approach. The resulting Multi-Scales DST estimators will have the further advantage, compared to simple Multi-Scales estimators, to be "robustified" by the use of the DST decomposition which will optimally prefilter market microstructure noise. This is made possible by the linear relation existing between the realized variance of the minimal principal component $c_{min}^{(M)}$ (i.e. the Minimal DST estimator) and the window length of the PCA M.

In fact, denoting $RV_{min}^{(M)} = \widehat{V[c_{min}^{(M)}]}$ we have:

$$\mathbb{E}\left[RV_{min}^{(M_j)}\right] = \sigma^2 + \eta^2 \mathcal{N}^{(M_j)} \tag{11}$$

where $\mathcal{N}^{(M_j)} = 4\sin^2 \frac{\pi}{2(M_j+1)}$.

Therefore, analogously to the Multi-Scales regression, a similar approach can be employed for the DST decomposition by evaluating the Minimal DST estimator $RV_{min}^{(M_j)}$ for different values of M_j and then performing a simple linear regression. Then, the intercept is an unbiased (not only asymptotically

[12] Both in simulation and empirical applications, we found that the performance of multi-scales estimator are very robust (giving very similar numerical results) to the precise choice of the frequencies provided that they are reasonably distributed along the frequency spectrum. In our applications we chose the following set of frequency: $\{1, 4, 8, 12, 16, 20, 25, 30, 60, 90, 120\}$

	σ^2	η^2	std(σ^2)	std(η^2)
True values & Cramer-Rao bounds	1	4	0.095	0.169
MS-DST	0.996	4.004	0.095	0.203
MS-DST + NR	0.998	4.000	0.093	0.168

Table 1: Evaluation of the absolute efficiency of the Multi-Scales DST estimator and the Multi-Scales DST estimator + Newton-Raphson algorithm (MS-DST + NR) with true volatility of 1, noise to signal ratio $\frac{\eta}{\sigma} = 2$, 2048 observations per day and 5,000 simulations.

but also in finite sample) and consistent estimator of the tick-by-tick volatility σ^2, while the slope is an estimate of η^2. Once appropriately rescaled by the number of ticks per day, the resulting IV estimator will be called the *Multi-Scales DST estimator*.

The contemporaneous presence of DST filtering, subsampling and mulstiscale regression makes the analytical computation of asymptotic results very difficult. However, from simulations (Table 1) it turns out that the Multi-Scales DST estimator for σ^2 possesses a finite sample variance very close to the Cramer-Rao bounds which now, thanks to the DST results, can be analytically computed (see Appendix C). Moreover, if desired, this small loss of efficiency could be eliminated by using the Multi-Scales DST estimator as initial value in a ML numerical optimization performed, for example, with Newton-Raphson method (as shown in Table 1). A more realistic simulation set up will be employed in the next section to compare different realized volatility estimators.

3 Monte Carlo Simulations

The DGP used in the simulations is a combination of the Heston (1993) SV model for the dynamics of the true price process and the model proposed by Hasbrouck (1999) for the microstructure effects.

In the Heston model the true log price assumes the following continuous time dynamics

$$dp^*(t) = (\mu - v(t)/2)dt + \sigma(t)dB(t) \quad (12)$$

$$dv(t) = k(\alpha - v(t))dt + \gamma v(t)^{1/2}dW(t) \quad (13)$$

where $v = \sigma^2$ and the initial value $v(0)$ is drawn from the unconditional Gamma distribution of v. The value of the parameters are the same as in Zhang et al. (2005), which in annualized terms are: $\mu = 5\%$, $k = 5$, $\alpha = 0.04$ corresponding to an expected annualized volatility of 20%, $\gamma = 0.5$ and the correlation coefficient between the two Brownian motions $\rho = -0.5$. Those parameters, who are reasonable for stocks, will be held constant throughout the simulations. The continuous time model of the true price is simulated at the usual Euler clock of one second.

To this SV model for the dynamics of the true price, we add the Hasbrouck bid-ask model for the observed price. The Hasbrouck model views the discrete bid and ask quotes as arising from the efficient price plus the quote-exposure costs (information and processing costs). Then the bid price is the efficient price less the bid cost rounded down to the next tick and the ask quote is the efficient price plus the ask cost rounded up to the next tick. As in Alizadeh et al. (2002) the model is simplified by assuming that the bid cost and the ask cost are both equal to the minimum tick size.

Then, according to the Hasbrouck model the bid and ask prices are respectively

$$B_n = \Delta \lfloor P_n^*/\Delta - 1 \rfloor \qquad \text{and} \qquad A_n = \Delta \lceil P_n^*/\Delta + 1 \rceil \tag{14}$$

where Δ represents the tick size, $\lfloor x \rfloor$ is the floor function, $\lceil x \rceil$ the ceiling one and the unobserved efficient price is $P_n^* = e^{p_n^*}$.

Hence, the observed price is given by the following bid-ask model

$$P_n = B_n q_n + A_n (1 - q_n) \tag{15}$$

with $q_n \sim \text{Bernoulli}(1/2)$. Therefore, the observed logarithmic return can be written as

$$r_n = \ln \frac{P_n}{P_{n-1}} = \ln \frac{\lceil P_n/\Delta + 1 \rceil}{\lceil P_{n-1}/\Delta + 1 \rceil} + q_n \ln \frac{\lfloor P_n/\Delta - 1 \rfloor}{\lceil P_n/\Delta + 1 \rceil} - q_{n-1} \ln \frac{\lfloor P_{n-1}/\Delta - 1 \rfloor}{\lceil P_{n-1}/\Delta + 1 \rceil} \tag{16}$$

which would be an MA(1) process in case the true price followed a Brownian motion. Notice that, by adopting the Heston model for the dynamics of the true price, the observed prices do not follow an MA(1) process anymore, making the DST approach formally misspecified. We choose this misspecified simulation setting expressly to show the robustness of the DST approach against more general heteroskedastic process in tick time as discussed in section 2.

We first follow Hasbrouck and Alizadeh et al. and choose parameter values which imply a high level of the noise to signal ratio: $\Delta = 1/16$ and $P_0 = 45$. These values, together with the average annualized volatility of 20% given by the Heston model for the true price, induce an average noise to signal ratio of about 3.5[13]. Such high level of noise manifests itself as a strong price fluctuation between bid and ask quotes, which generates a highly negative first order-autocorrelation $\rho(1) \approx -48\%$ for the tick-by-tick returns r_n.

This noise to signal ratio reflects a microstructure impact on the return process which is remarkably large and rarely observed on real data. However, such an extreme setting provides a useful stress test for realized volatility measures and hardens the competition versus daily range-based estimators which are favored under these circumstances.

[13]Following (Oomen 2006) we define the noise to signal ratio as the standard deviation of the noise divided by the average standard deviation per tick of the true price process. This standardization has been chosen first because it seems reasonable to normalize both the noise and signal standard deviation respect to the same time interval and second because doing that at the tick-by-tick level facilitates comparison across different assets and over time, makes such a ratio not affected by the different market activity.

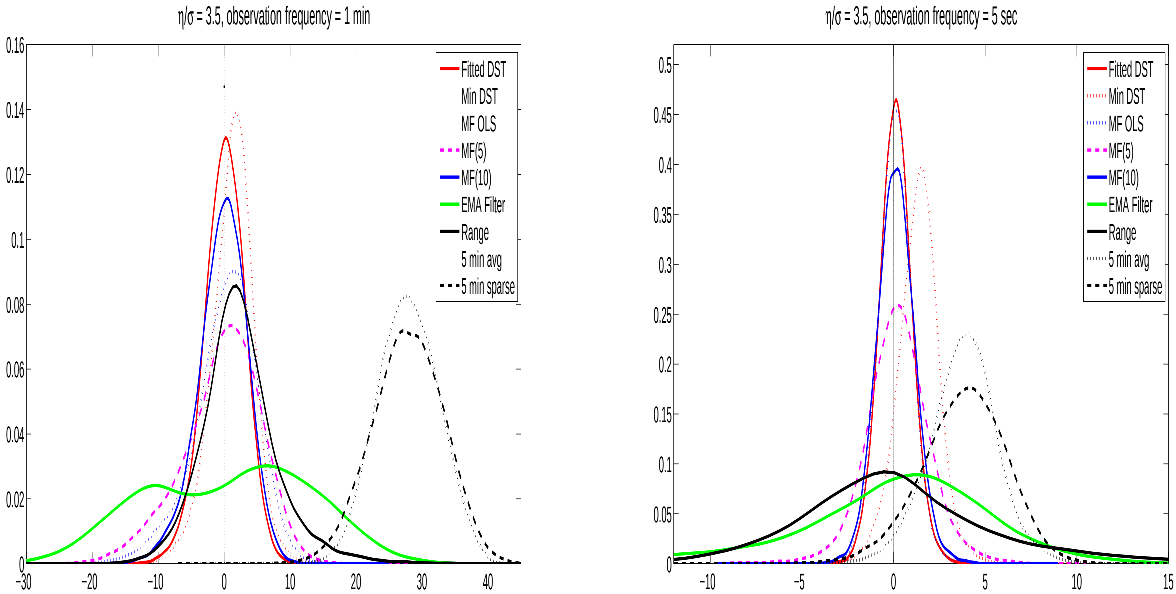


Figure 1: Comparison of the pdf of the estimation errors on the annualized percentage volatility (on average 20%) obtained with an average observation frequency of 1 minute (left panel) and 5 seconds (right panel) and a noise to signal ratio $\frac{\eta}{\sigma} = 3.5$.

We simulate one-day sample paths of 6.5 hours (the typical opening time for stock markets) for 25,000 days. The simulation is repeated for two different values of the total number of price observations per day: $M = 390$ which corresponds to an average intertrade duration of one minute, and $M = 4,680$ which corresponds to an average tick arrival time of 5 seconds.

The competing estimators are:

- the two DST estimators: the Minimal DST (Min-DST) is computed with a window length of 30 ticks, while, for the Multi-Scales DST (MS-DST), we construct a series of minimal DST estimator using a sequence of M_j ranging from 2 to 20 ticks and then fit equation (11) through simple OLS.

- three "simple" Multi-Scales estimators: two Two-Scales estimators with frequency ratio α of 5 and 10, denoted respectively TS(5) and TS(10), and the Multi-Scales Least Square (MS-LS) estimator.[14]

- the local EMA filter (i.e. calibrated on a single day), which then simply corresponds to a daily MA(1) filter;

[14]About the values chosen for the frequency ratio α, Bandi and Russel (2009) show that for realistic sample sizes encountered in practical applications, the asymptotic criteria suggested in Zhang et al. (2005) do not give sufficient guidance for practical implementation, as they provide unsatisfactory representations of the finite sample properties of the estimators. We then preferred to determine the optimal scales in the Two-Scales estimator performing, by mean of simulations, an extensive grid search over the return frequencies of the slower scale.

VOLATILITIES ESTIMATES WITH $\frac{\eta}{\sigma}$ =3.5

	1 min			5 sec		
	mean	std	RMSE	mean	std	RMSE
MS-DST	-0.052	3.103	3.103	0.090	0.891	0.895
Min-DST	1.183	3.206	3.418	1.295	1.113	1.708
MS-LS	0.080	4.818	4.819	0.066	0.909	0.911
TS(5)	-0.561	5.929	5.955	0.148	1.777	1.783
TS(10)	-0.371	3.711	3.730	0.090	1.041	1.044
EMA Filter	0.393	12.401	12.408	-0.018	5.222	5.222
daily Range	2.252	5.925	6.339	0.149	5.907	5.909
5 min avg	27.703	5.375	28.220	3.735	2.319	4.396
5 min sparse	27.752	4.620	28.134	3.724	1.855	4.160

Table 2: The table report the mean, standard deviation and RMSE of the estimation errors on the annualized percentage volatility (on average 20%) obtained with an average observation frequency of 1 minute (left panel) and 5 seconds (right panel), and a noise to signal ratio $\frac{\eta}{\sigma} = 3.5$.

- two standard realized volatility measures both computed with 5 minute returns but one sampled with an overlapping scheme and then averaged;
- the daily range, as proposed by Parkinson (1980) and recently advocated by Alizadeh et al. (2002) in the context of SV models estimation[15] .

We first consider the case of having 390 observations per day (corresponding to an average one minute frequency) and a noise to signal ratio of 3.5. Table 2 reports the mean, standard deviation and Root Mean Square Error (RMSE) of the estimation errors on the annualized volatility (express as a percentage). Figure 1 shows the probability density functions of those volatility estimation errors.

Given the high level of noise and the relatively small number of observations per day, the estimation of the first order autocorrelation required to calibrate the EMA filter, is very noisy and does not always satisfy the theoretical bound for MA(1) process $\mid \rho(1) \mid < 1/2$ (in the 30% of the cases), leading to a complex MA(1) coefficient θ. In such cases, the EMA filter would fail and we are then forced to impose an artificial floor to $\rho(1)$. But, besides its arbitrariness, this procedure induces unreasonably low volatility estimates (responsible for the left bump presents in the EMA estimator pdf on the left panel of Figure 1). Moreover, under these conditions, the variance of the estimator is extremely large.

[15] Comparison with the recently proposed high frequency range, the so called "Realized Range-based Variance" (Christensen and Podolskij 2006 and Martens and van Dijk 2006) will be deferred to future research.

For the 5 minute realized volatility, the fact that the aggregation from 1 to 5 minute returns is not able to eliminate all the negative autocorrelation, makes this estimator strongly upward biased. In the case of the Minimal DST estimator instead, the aggregation works much better but, due to a relatively low window length of 30 ticks, a small upward bias is still present. Even the daily range suffers from a significant bias but it also has a much larger variance (both, the bias and the variance, are about two times those of the Minimal DST one). Among the three simple Multi-Scales estimators the TS(5) and TS(10) have small negative biases while the MS-LS is virtually unbiased. However, the variance and the RMSE of the TS(10) are lower than the other two simple multi-frequency estimators. Under this extreme setting, the only measure which is still able to remain unbiased and sufficiently precise is the MS-DST estimator, which has, in fact, the lowest RMSE. Moreover, comparing the realized volatility estimators with the one based on the daily range shows that, even in the most unfavorable setting for the realized volatilities, they remain much more accurate than the daily range: the best realized volatility estimator, the MS-DST, possesses, in fact, a RMSE 48% smaller than that of the daily range.

Keeping the same level of noise, we repeat the simulation at 5 second frequency (which means 4,680 observations per day). With twelve times more data the realized volatility measures are much more precise: the local EMA filter has less failings (5%) and lower variance, while the 5 minute realized volatilities (thanks to the longer aggregation period) have smaller, but still significant, biases. Although smaller than the biases of the 5 minute realized volatilities, the Minimal DST still shows a bias with this high level of noise. The Zhang et al. estimators become both unbiased with the TS(10) having a smaller variance than the TS(5). The MS-DST and the MS-LS estimators are both unbiased and equally very accurate, remaining the best choices among the estimators considered.

In practice, however, financial time series present a noise to signal ratio at tick-by-tick level that usually lies between 0.5 and 2. But, even with such a moderate level of noise, a naive high frequency realized volatility measure would be from one to three times the actual one. We then repeat the simulation with a more realistic noise to signal ratio of 1.5 for both observation frequencies. Table 3 and Figure 2 summarize the results.

At the 1 minute frequency the daily range and the local EMA filter are unbiased but quite inaccurate while the realized volatilities based on 5–minute have again a large bias. The MS-DST and the other Multi-Scales estimators are the most accurate with the MS-LS having a slightly smaller bias and variance than the others.

At the 5 second frequency, with a moderate level of noise and a large number of data, the EMA filter starts to have a much lower variance and the 5 minute measures much lower biases. Nevertheless, they can still not compete with the MS-DST and the three other Multi-Scales estimators which become extremely precise and accurate under this setting.

Empirical studies on the autocorrelation of tick-by-tick data often show significant values not only for the first order but also for higher order lags (though, usually, of much smaller amplitude). A possible explanation, and way to model it, is by relaxing the assumption of i.i.d. microstructure noise

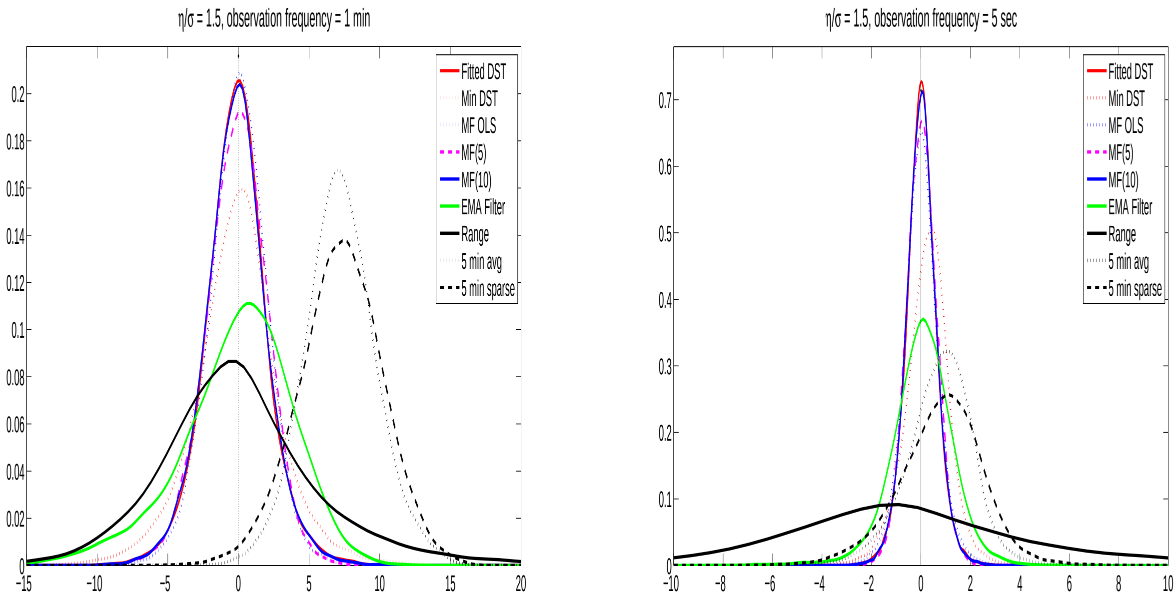


Figure 2: Comparison of the pdf of the estimation errors on the annualized percentage volatility (on average 20%) obtained with an average observation frequency of 1 minute (left panel) and 5 seconds (right panel) and a noise to signal ratio $\frac{\eta}{\sigma} = 1.5$.

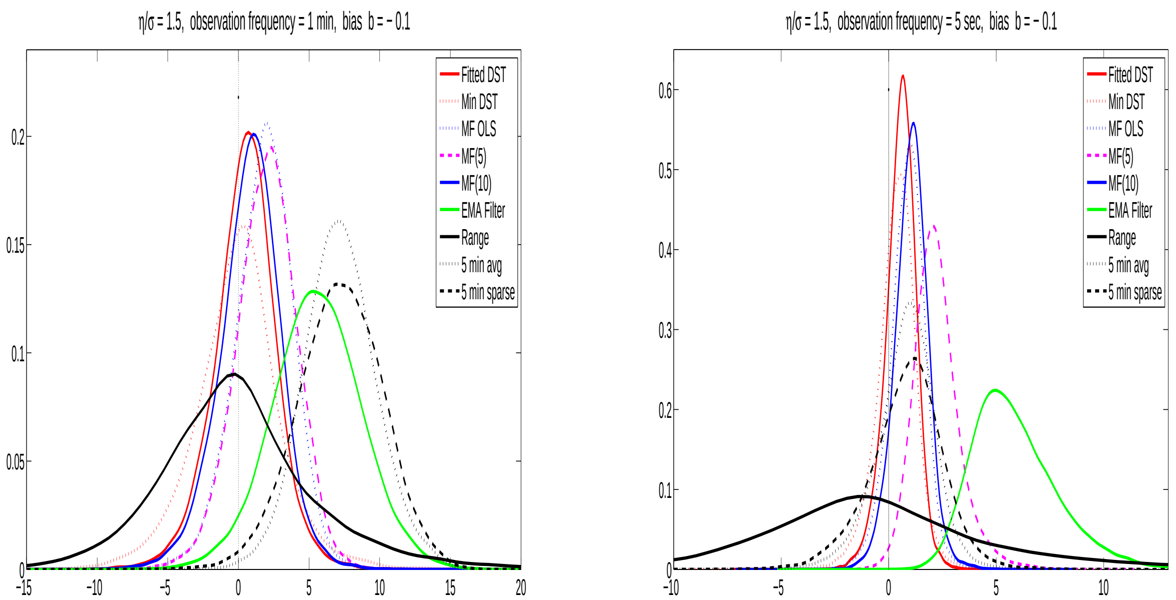


Figure 3: Comparison of the pdf of the estimation errors on the annualized percentage volatility (on average 20%) obtained with an average observation frequency of 1 minute (left panel) and 5 seconds (right panel), a noise to signal ratio $\frac{\eta}{\sigma} = 1.5$ and a biased Bernoulli process with bias $b = -0.1$.

VOLATILITIES ESTIMATES WITH $\frac{\eta}{\sigma} = 1.5$

Unbiased Bernoulli						
	1 min			5 sec		
	mean	std	RMSE	mean	std	RMSE
MS-DST	-0.220	2.213	2.224	-0.012	0.627	0.627
Min-DST	-0.226	3.068	3.076	0.210	0.899	0.923
MS-LS	-0.096	2.030	2.032	-0.028	0.711	0.712
TS(5)	-0.207	2.155	2.165	-0.004	0.612	0.612
TS(10)	-0.238	2.212	2.225	-0.013	0.625	0.625
EMA Filter	-0.296	4.046	4.056	-0.030	1.268	1.268
daily Range	-0.009	5.998	5.998	-0.471	5.824	5.843
5 min avg	7.126	2.971	7.720	0.732	1.704	1.855
5 min sparse	7.121	2.528	7.557	0.700	1.345	1.517

Biased Bernoulli						
	1 min			5 sec		
MS-DST	0.389	2.234	2.267	0.675	0.717	0.984
Min-DST	-0.153	3.077	3.081	0.329	0.904	0.963
MS-LS	1.722	2.070	2.693	0.831	0.817	1.166
TS(5)	1.888	2.150	2.862	2.200	1.081	2.451
TS(10)	0.726	2.214	2.330	1.023	0.791	1.294
EMA Filter	5.467	3.136	6.303	5.840	1.979	6.167
daily Range	-0.128	5.823	5.825	-0.457	5.926	5.944
5 min avg	7.103	2.978	7.702	0.754	1.741	1.898
5 min sparse	7.105	2.549	7.548	0.719	1.370	1.548

Table 3: The table reports the mean, standard deviation and RMSE of the estimation errors on the annualized percentage volatility (on average 20%) obtained with an average observation frequency of 1 minute (left panel) and 5 seconds (right panel), a noise to signal ratio $\frac{\eta}{\sigma} = 1.5$ and, for the bottom panel, a biased Bernoulli process with bias $b = -0.1$.

by introducing a correlation in the sequence at which bid and ask prices arrive[16]. Hence, instead of having an "unbiased" Bernoulli(1/2) for the q_n process, we construct a Bernoulli process which produces autocorrelation in q_n. This "biased" Bernoulli is obtained by taking $q_n = \text{Bernoulli}\,(1/2 + b)$ if $q_{n-1} = 1$ and $q_n = \text{Bernoulli}\,(1/2 - b)$ if $q_{n-1} = 0$. We choose $b = -0.10$ which induces a second order autocorrelation of about -6%.

Now, in the presence of not i.i.d. microstructure noise, the local EMA filter, which was unbiased, becomes highly biased at both frequencies (see Figure 3). Also the TS and MS-LS estimators are now showing a positive bias. Among all the realized volatility estimators the DST measures are the ones with the smallest bias and smallest RMSE, showing a high degree of robustness against more general microstructure noise contaminations (as analytically described in the previous section).

Summarizing the results of the simulation study, we can draw the following conclusions. The daily range estimator is always inferior to the realized volatility ones. The realized volatilities with 5 minute returns are often significantly biased and inaccurate. The local EMA filter gives satisfactory results only in the presence of a high number of observations and a low level of i.i.d. noise. Although more precise in general, similar considerations can be made for the simple MS estimators (TS and MS-LS). When the microstructure noise is moderate and i.i.d. the simple MS estimators are almost as accurate as the MS-DST and hence close to the optimal Cramer-Rao efficiency bound (since, as shown in section 2.4, the MS-DST is very close to the full efficiency of the Cramer-Rao bounds). In particular, the MS-LS seems to be particularly efficient in exploiting the information contained in the data when a relatively small number of observations is available (perhaps due to its ability to extract information from many frequencies), while the TS(5) and TS(10) are at their best when the number of observations increases. However, when the microstructure noise increases and deviations from the i.i.d. structure arise, the discrepancy between the simple MS estimators and the MS-DST starts to increase due to a higher level of robustness of the DST approach. Therefore, the overall winner that seems to arise from this volatility estimation "horse race" is the MS-DST which shows the highest level of precision and robustness across a wide range of microstructure noise contaminations.

4 Empirical application

To verify the behavior of volatility estimators when the microstructure noise is not an i.i.d. process, we analyze six years of tick-by-tick data (from January 1998 to October 2003) for the following three future contracts: the S&P 500 stock index future, the 30 years U.S. Treasury Bond future and the Italian stock index future FIB 30. In a base asset mapping approach (as the one of RiskMetrics), those three major future contracts can be seen as the reference liquid base assets for, respectively, the US stock and bond market and the Italian stock market.

In order to analyze the dependence structure of the microstructure noise in those series, we in-

[16] Hasbrouck and Ho (1987) suggest that positive autocorrelation at lag lengths greater than one may be the result of traders working an order: "a trader may distribute purchases or sales over time". However also significant negative autocorrelation at lag two are often observed.

vestigate the behaviour of the autocorrelation of tick-by-tick returns. This tick-time autocorrelation analysis shows significant departure from the standard i.i.d. assumption for the microstructure noise. In fact, more complex structures than that of a simple MA(1) expected under the standard i.i.d. assumption, were found in all the three series. These patterns are independent of the inclusion or censoring of all zero trade-by-trade returns which, in all the three assets, usually represent only a small percentage of the total number of trade-by-trade returns.

Such autocorrelation patterns of the tick-by-tick returns are instead consistent with more complex ARMA structure for the microstructure noise. In fact, simulating the Hasbrouck model with those ARMA structures for the noise, leads to exactly the same autocorrelation functions observed in the data. In particular, those patterns are consistent with a microstructure noise having an MA(1) structure for the FIB, an MA(2) (at least) for the S&P and a strong oscillatory AR(1) for the US bond[17] (see Figure 4).

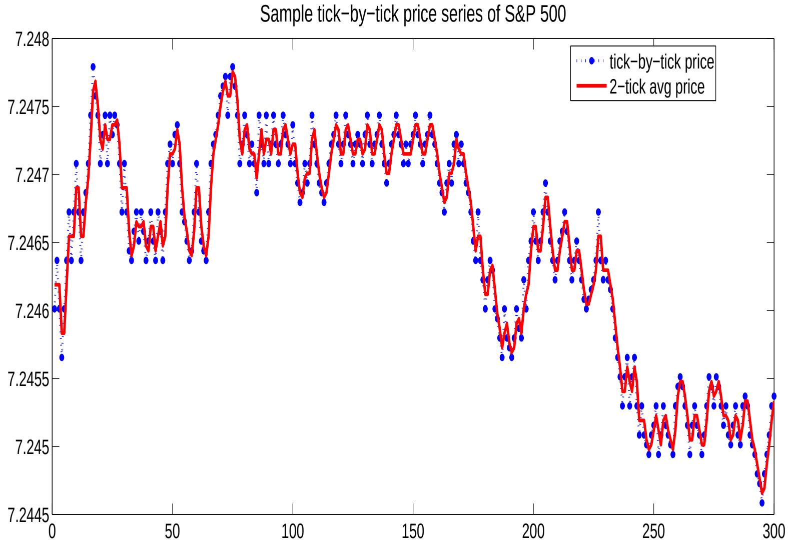
To overcome the problem of a complex ARMA structure in the autocorrelation of tick-by-tick returns of the S&P and US bond, we first notice that, in both cases, a simple aggregation of two ticks returns almost restore the MA(1) autocorrelation pattern typical of the i.i.d. assumption for the microstructure noise (see Figure 5). Therefore, applying the MS-DST estimator to the two-ticks returns of the S&P and US bond series, we can still obtain a highly precise evaluation of the realized volatilities of the two assets and closely follow their time series dynamics (see Figure 6).

Obviously, in empirical analysis the true volatility is not observable, hence no direct evaluation criteria of the quality of the volatility estimators exist. However, general indirect criteria can be employed.

First of all, the unconditional mean of daily volatilities obtained with high frequency estimators should not be significantly different from the unconditional mean volatility obtained with lower frequency returns. We asses this property for the MS-DST estimator by computing its volatility signature plot. Figure 7 shows the volatility signature plot of the standard and MS-DST realized volatility measures for the three assets, averaged over the whole six years period. Ideally, for estimators which are robust against microstructure effects the scaling should appear as a flat line in the volatility signature plot. The top panel of Figure 7 refers to the scaling of S&P 500 future showing a moderate but clear impact of the market microstructure on the standard realized volatility measure and the presence of a mild lower frequency autocorrelation. The MS-DST estimator correctly discounts market microstructure effects on volatility while it retains the residual lower frequency autocorrelation which is responsible for its scaling behavior to be not completely flat. The middle and bottom panels are respectively the FIB 30 and U.S. Bond future. In both cases market microstructure noise has a strong impact on the standard measure of realized volatility inducing larger biases as the frequency increases while the MS-DST estimator, remaining reasonably flat at any frequency, confirms its ability to properly filter out market microstructure effects.

Under the hypothesis of an underlying continuous time diffusion process for the logarithm price,

[17] The analysis of the market microstructure determinants and specific institutional constraints that would lead to such empirical evidences is beyond the scope of the present study.



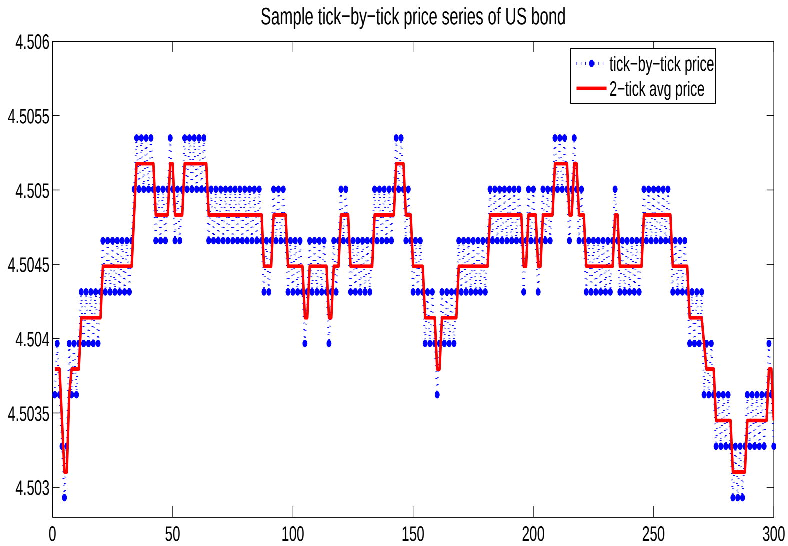


Figure 4: Sample path of the tick-by-tick price process (dotted line) with its two ticks moving average (solid line) for the S&P 500 (top panel) and US Bond (bottom panel) future.

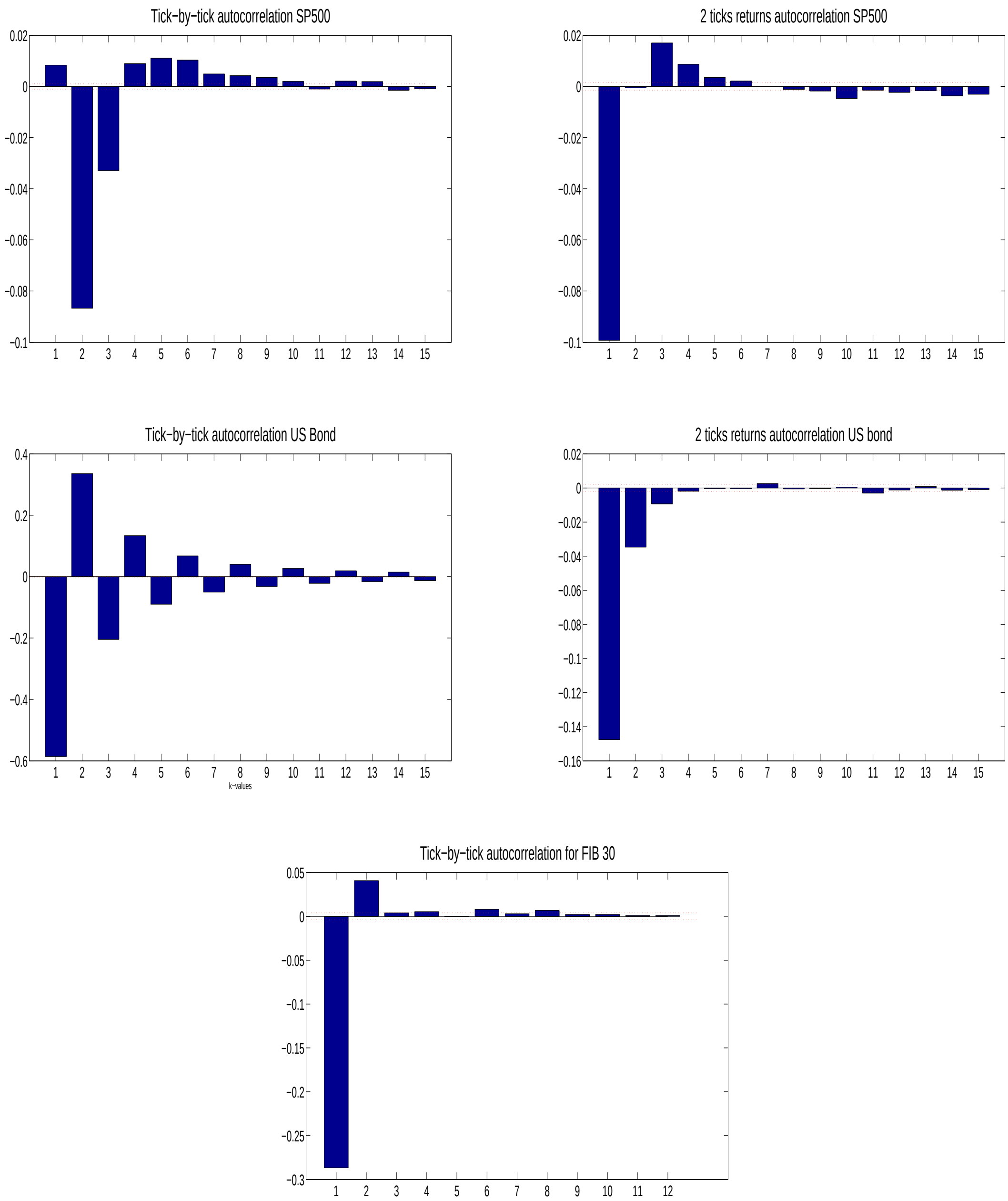


Figure 5: Sample autocorrelation of S&P (top) and US Bond (middle) for tick-by-tick returns (left) and 2-ticks returns (right) together with the tick-by-tick autocorrelation of FIB30 (bottom). All autocorrelation functions are computed over the six year sample, from 1998 to 2003.

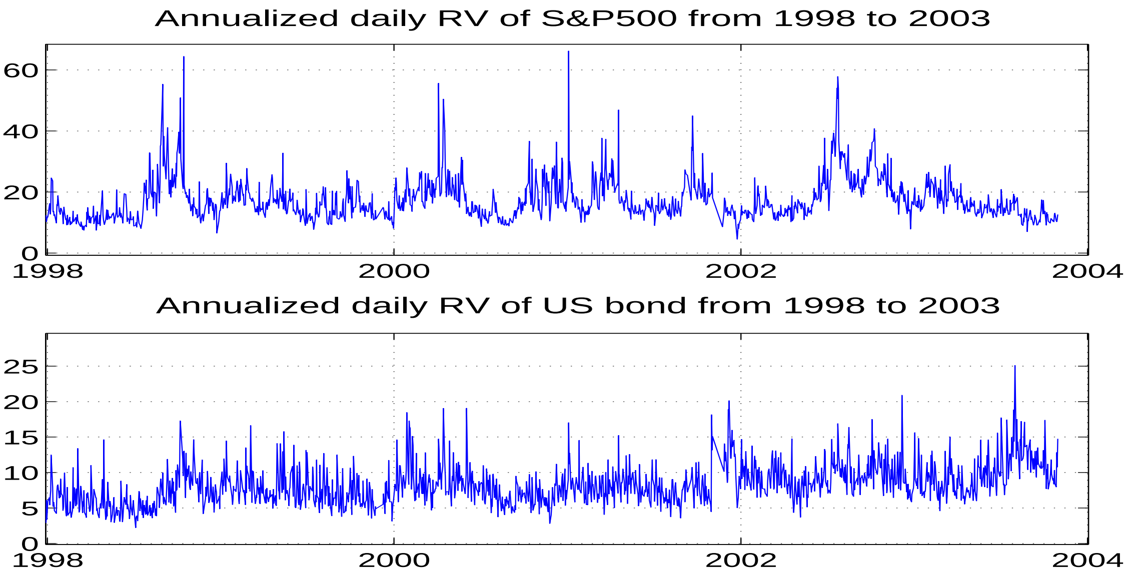


Figure 6: Time series of the annualized realized volatility computed with the MS-DST estimator on the two ticks returns for the S&P 500 (top panel) and US Bond (bottom panel) futures from 1998 to 2003.

another indirect criterion can be considered to asses the quality of realized volatility measures in empirical applications. In fact, if the log-price follows a SV diffusion the model for daily returns could be written as $r_t = \sigma_t z_t$ where $z_t \sim$ i.i.d. $N(0, 1)$. Hence, the 1-day return would be conditionally Gaussian with variance equal to the integrated variance. The normality of z_t is justified by appealing to the Central Limit Theorem for mixing process aggregated over a reasonable length of time (such as daily for highly traded assets). Therefore, if a volatility measure adequately estimates the integrated volatility, the corresponding standardized returns should be normally distributed. We test this condition using the Jarque-Bera normality test on returns standardized by the 30 minute realized volatility[18], MS-DST realized volatility and the daily range[19].

Table 4 reports the results. In all three cases, daily raw returns are highly leptocurktic as expected, while returns standardized by daily ranges become highly thin tailed and remain far from normal. Returns standardized by 30 minute realized volatilities become excessively thin tailed for the S&P and US Bond while remaining too fat tailed for the FIB 30 and clearly failing the Jarque-Bera tests

[18] This choice of a somewhat lower frequency of 30–minute instead of an higher one, is motivated by the need of having an unbiased estimator of the daily volatility. Higher frequency realized volatility measures (such as with 1–minute or 5–minute returns) would in fact suffers, on this data set, of a large positive bias that by itself would heavily distort the criterion here considered.

[19] The EMA filter estimator has not been included here because, as shown in the simulations, it is sensitive to the presence of significant higher order autocorrelation in the tick-by-tick returns which results to be significantly different from zero in all the three series considered here. While, on this kind of data, the simple MS estimators give results that under these weak empirical tests are almost indistinguishable from the ones of the MS-DST, thus confirming the results of the Monte Carlo simulations where the MS-DST and the simple MS estimators were all very close when the level of noise were moderate and the number of observations relatively high.

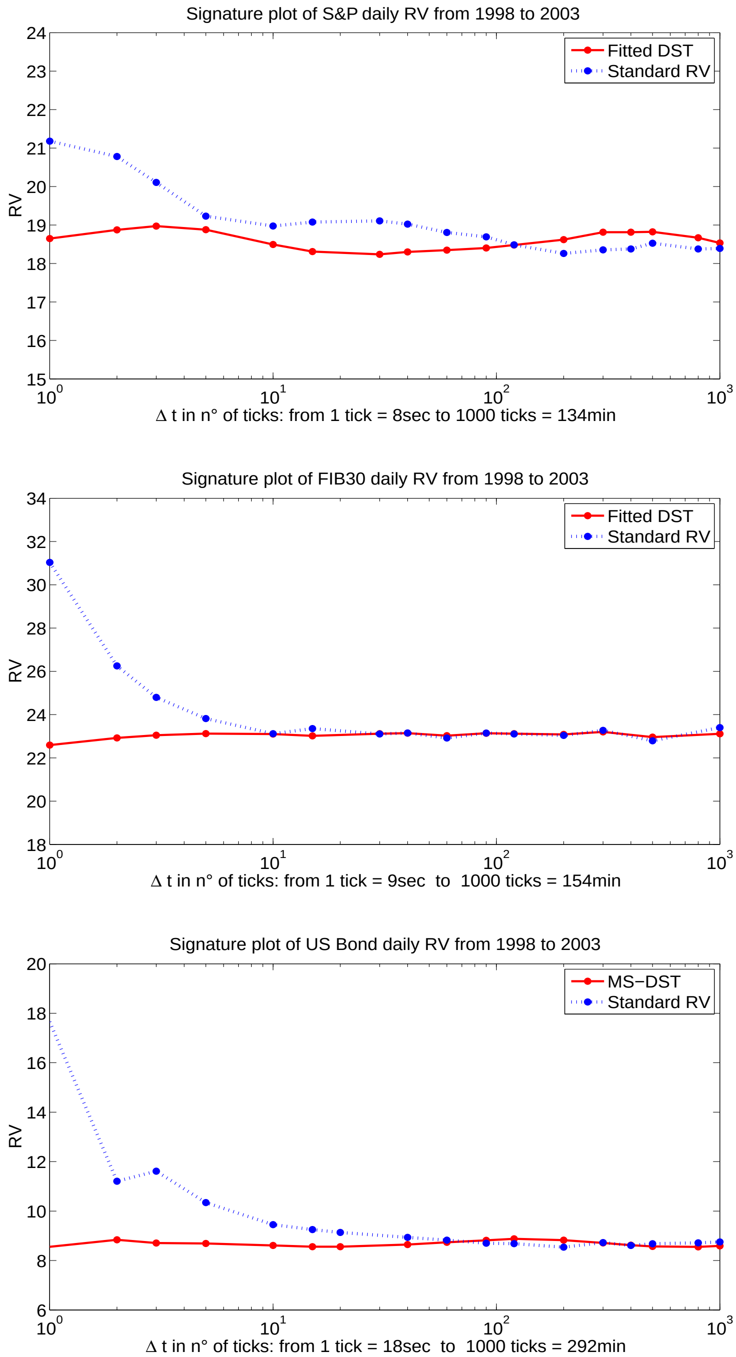


Figure 7: Signature Plot in semilog scale of the standard (dotted) and DST (solid) realized volatility for S&P 500, FIB 30 and US Bond from January 1998 to October 2003.

in all the three cases. Whereas for the MS-DST standardized returns, Jarque-Bera test cannot be rejected for both the stock index future S&P and FIB. However, for the US Bond future, even though among the three competing estimators the MS-DST standardized returns remain by far the closest to the standard normal, the Jarque-Bera test is rejected. The rejection is due to a value of the kurtosis excessively smaller than three, meaning that the MS-DST measure tends to overestimate the "true" integrated volatility of the Bond future process. However, since the realized volatility consistently estimates the quadratic variation (which includes the contributions of jumps) and not the integrated volatility (which only considers the contribution of the continuous part), such overestimation could be due to the presence of a large jump components in the Bond future series. Indeed, the fact that the relative contribution of jumps is higher in bond series compared to stock indices, has been recently found by Andersen Bollerslev and Diebold (2007) and is consistent with the empirical evidence of the fixed income market being the most responsive to macroeconomic news announcements (Andersen, Bollerslev, Diebold and Vega 2003).

In summary, the analysis conducted on the empirical data confirms the ability of the MS-DST estimators to accurately and reliably estimate daily realized volatility, thus confirming the results obtained in the Monte Carlo simulation analysis.

5 Conclusions

The presence of microstructure effects represent a challenging problem for realized volatility measures making the naive realized volatility computed at short time intervals highly biased. In this study new realized volatility measures based on Multi-Scale regression and *Discrete Sine Transform* (DST) approaches are presented. We show that Multi-Scales estimators similar to that recently proposed by Zhang (2006) can be constructed within a simple regression based approach by exploiting the linear relation existing between the market microstructure bias and the realized volatilities computed at different frequencies. These regression based estimators can be further improved and robustified by using the DST approach to filter market microstructure noise. This approach is justified by the theoretical result regarding the ability of the DST to diagonalize exactly an MA(1) process and approximately an MA(q) one. Hence, we utilize the DST orthonormal basis decomposition to optimally disentangle the underlying efficient price signal from the time-varying nuisance component contained in tick-by-tick return series. The robustness of the DST approach with respect to more general dependent structures of the microstructure noise is also analytically shown.

The combination of such a Multi-Scale regression approach with the DST gives us a Multi-Scales DST realized volatility estimator which is then robust against a wide class of noise contaminations and model misspecifications. Thanks to the DST orthogonalization, which also allows us to analytically derive closed form expressions for the Cramer-Rao bounds of MA(1) processes, an evaluation of the absolute efficiency of volatility estimators under the i.i.d. noise assumption becomes available, indicating that the Multi-Scales DST estimator possesses a finite sample variance very close to the optimal Cramer-Rao bounds.

	Std. Dev	Kurtosis	Skewness	Jarque-Bera	Probability
S&P 500					
Raw returns	19.395	6.540	-0.010	734.828	0.000
MS-DST-std. returns	1.024	2.764	0.019	3.344	0.187
30 min-std. returns	1.066	2.432	-0.026	19.044	0.000
Range-std. returns	0.922	1.754	-0.037	91.316	0.000
FIB 30					
Raw returns	24.4786	6.853	-0.236	829.114	0.000
MS-DST-std. returns	1.106	2.890	0.146	5.370	0.068
30 min-std. returns	1.577	4.171	-0.036	75.720	0.000
Range-std. returns	0.912	1.752	0.049	86.137	0.000
US Bond					
Raw returns	8.6578	4.101	-0.423	113.056	0.000
MS-DST-std. returns	0.966	2.517	-0.110	16.464	0.000
30 min-std. returns	1.000	2.283	-0.094	32.211	0.000
Range-std. returns	0.887	1.766	-0.094	91.326	0.000

Table 4: Comparison of sample distribution properties of daily raw and standardized returns of FIB 30, S&P 500 and thirty years Bond futures from 1998 to 2003. Standardized returns are computed using MS-DST, 30 minute realized volatility and daily range.

Monte Carlo simulations based on a realistic model for microstructure effects and volatility dynamics, show the superiority of MS-DST estimators compared to alternative local volatility proxies such as the TS and MS-LS estimators, the daily range, the EMA filter and 5 minute realized volatilities. The MS-DST estimator results to be the most accurate and robust across a wide range of noise to signal ratios and types of microstructure noise contaminations. The empirical analysis based on six years of tick-by-tick data for S&P 500 index-future, FIB 30, and 30 years U.S. Tresaury Bond future, seems to confirm Monte Carlo results.

Appendix A: Dependent Noise

In the presence of MA$(q-1)$ dependent noise, the observed returns in tick time become an MA(q) process which can be written as

$$r_n = \sigma \tilde{\epsilon}_n + \sum_{i=1}^{q} \eta_i \left(\omega_n^{(i)} - \omega_{n-i}^{(i)} \right)$$

with $\tilde{\epsilon}_n \sim \text{IID}(0,1)$ and $\omega_n^{(i)} \sim \text{IID}(0,1)$. In this more general case the variance of the Minimal DST component becomes

$$\sigma_M^2 = \mathbb{E}\left(c_{min}^{(M)}(n)\, c_{min}^{(M)}(n)\right) = \sigma^2 + \sum_{i=1}^{q} \eta_i^2 F(M,i)$$

where

$$F(M,i) = \frac{2}{M+1}\left[M+1-(M+1-i)\cos\frac{\pi i}{M+1} - \cot\frac{\pi}{M+1}\sin\frac{\pi i}{M+1}\right]$$

.

Because $F(M,i)$ can be approximated as

$$F(M,i) = \pi^2\frac{i^2}{M^2} - 2\pi^2\frac{i^3}{M^3}\left(1+\frac{3}{i}-\frac{1}{i^2}\right) + O\left(\frac{i^4}{M^4}\right)$$

when $M/q \to \infty$, we obtain that

$$\sigma_M^2 \simeq \sigma^2 + \frac{\pi^2}{M^2}\sum_{i=1}^{q}(i\,\eta_i)^2$$

Appendix B: Multi-Scales Regression with Endogenous Noise

Writing the observed tick-by-tick return as $r_n = r_n^* + u_n$ with $r_n^* = p_n^* - p_{n-1}^*$ the efficient return and u_n the endogenous noise, we model the correlation between the microstructure noise and the dynamics of the efficient price by assuming the following form for u_n (as in Hansen and Lunde 2006):

$$u_n = \rho(p_n^* - p_{n-1}^*) + \eta(\omega_n - \omega_{n-1}) = \rho r_n^* + \eta(\omega_n - \omega_{n-1}). \tag{B-1}$$

We then have that $E[u_n] = 0$, $E[u_t^2] = \rho^2\sigma_n^2 + 2\eta^2$ and $E[r_n^* u_n] = \rho\sigma_n^2$, while $E[r_{n-s}^* u_n] = 0$ for $s > 0$.

Therefore, the conditional expectation of the daily realized variance $RV^{(k_j)}$ computed with observed returns of different tick-lengths k_j is

$$\mathbb{E}\left[RV^{(k_j)}\right] = IV + N^{(k_j)}(\rho^2\sigma^2 + 2\eta^2) + 2N^{(k_j)}\rho\sigma^2 = IV + N^{(k_j)}(\rho(2+\rho)\sigma^2 + 2\eta^2) \tag{B-2}$$

where $N^{(k_j)}$ is the number of k_j-returns in the day and $\sigma^2 = IV/N$ the average variance per tick. The terms proportional to $N^{(k_j)}$ stem from the facts that the endogenous noise arise only when an observation is made (thus the term $N^{(k_j)}$) and the correlation between k_j tick return ($r^{*(k_j)}$) and microstructure noise remains the same irrespective of their tick lengths since

$$E[r_n^{*(k_j)}\, u_n] = E[(r_{n-k_j}^* + r_{n-k_j+1}^* + \dots + r_{n-1}^* + r_n^*)\, u_n] = E[r_n^* u_n] = \rho\sigma_n^2. \tag{B-3}$$

Appendix C: Exact MA(1) Likelihood, Cramer-Rao bounds and absolute efficiency

Thanks to the universality of the eigenvectors, we can obtain a diagonalization of the variance-covariance matrix of MA(1) processes which does not depend on the parameters to be estimated. Collecting the M eigenvectors of Ω in the MxM characteristic matrix $\Psi = [\varphi_1\varphi_2...,\varphi_M]$, we can project the return vector onto the orthogonal space of the principal component $C = \Psi^\top R$, which is a Mx1 vector distributed as $C \sim N(0,\Lambda)$, where Λ is the MxM diagonal matrix containing the M eigenvalues of the tridiagonal matrix Ω[20]. Therefore, the likelihood function of R can be rewritten in terms of the principal components vector C as

$$f(C) = \frac{1}{\sqrt{(2\pi)^M \det\Lambda}}\exp\left[-\frac{1}{2}C^\intercal\Lambda^{-1}C\right] = \frac{1}{\sqrt{(2\pi)^M\prod_{n=1}^M \lambda_n}}\exp\left[-\frac{1}{2}\sum_{n=1}^{M}\frac{c_n^2}{\lambda_n}\right]$$

[20] In fact, $\mathbb{E}[CC^\top] = \mathbb{E}[\Psi^\top RR^\top\Psi] = \Psi^\top\mathbb{E}[RR^\top]\Psi = \Psi^\top\Omega\Psi = \Lambda$.

and

$$\ln f_\theta(C) \quad = \quad -\frac{M}{2}\ln 2\pi - \frac{1}{2}\sum_{n=1}^{M}\ln\lambda_n - \frac{1}{2}\sum_{n=1}^{M}\frac{c_n^2}{\lambda_n}$$

Then, from the linear equation (5) we readily obtain

$$\frac{\partial\lambda_n}{\partial\theta} = \left(\begin{array}{c} 1 \\ 4\sin^2\frac{\pi n}{2(M+1)} \end{array}\right) \qquad \text{and} \qquad \frac{\partial^2\lambda_n}{\partial\theta_i\partial\theta_k} = 0 \qquad \text{for} \qquad i,k = 1,2$$

and hence, we are now able to analytically derive the equations for the Score and the Hessian

$$\frac{\partial\ln f_\theta(C)}{\partial\theta_i} = \frac{1}{2}\sum_{n=1}^{M}\left(\frac{c_n^2}{\lambda_n^2} - \frac{1}{\lambda_n}\right)\frac{\partial\lambda_n}{\partial\theta_i}\ ; \qquad \frac{\partial^2\ln f_\theta(C)}{\partial\theta_i\partial\theta_k} = -\sum_{n=1}^{M}\left(\frac{c_n^2}{\lambda_n^3} - \frac{1}{2\lambda_n^2}\right)\frac{\partial\lambda_n}{\partial\theta_i}\frac{\partial\lambda_n}{\partial\theta_k}$$

Therefore, thanks to equation (5) and (6) we are able to explicitly compute the Fisher Information matrix of an MA(1) process, which reads

$$\mathcal{I}_{ik} = -\mathbb{E}\left(\frac{\partial^2\ln f_\theta(C)}{\partial\theta_i\partial\theta_k}\right) = \frac{1}{2}\sum_{n=1}^{M}\frac{1}{\lambda_n^2}\frac{\partial\lambda_n}{\partial\theta_i}\frac{\partial\lambda_n}{\partial\theta_k}$$

With each element of the matrix given by

$$\mathcal{I}_{11} = \frac{1}{2}\sum_{n=1}^{M}\frac{1}{\lambda_n^2}, \quad \mathcal{I}_{22} = 8\sum_{n=1}^{M}\frac{1}{\lambda_n^2}\sin^4\left(\frac{\pi n}{2(M+1)}\right), \quad \mathcal{I}_{12} = \mathcal{I}_{21} = 2\sum_{n=1}^{M}\frac{1}{\lambda_n^2}\sin^2\left(\frac{\pi n}{2(M+1)}\right)$$

Then the Cramer-Rao bounds of $\hat{\sigma}^2$ and $\hat{\eta}^2$ can now be given in closed form as

$$\operatorname{var}\left(\hat{\sigma}^2\right) \geqslant \frac{\mathcal{I}_{22}}{\mathcal{I}_{11}\mathcal{I}_{22} - \mathcal{I}_{12}^2} \qquad \text{and} \qquad \operatorname{var}\left(\hat{\eta}^2\right) \geqslant \frac{\mathcal{I}_{11}}{\mathcal{I}_{11}\mathcal{I}_{22} - \mathcal{I}_{12}^2}$$

These results have two important implications. First they obviously permit to evaluate the absolute efficiency of volatility estimators. Second numerical optimization of the exact likelihood is greatly simplified. In fact, given that the principal components do not depend on the parameter, the orthogonalization of the returns process needs to be done only once rather then at each iteration as it occurs using Cholesky factorization (see Hamilton 1994).

References

Aït-Sahalia, Y., Mykland, P. A. and Zhang, L.: 2005, How often to sample a continuous-time process in the presence of market microstructure noise, *Review of Financial Studies* **18**, 351–416.

Aït-Sahalia, Y., Mykland, P. A. and Zhang, L.: 2006, Ultra high frequency volatility estimation with dependent microstructure noise, *Manuscript.* Princeton University, University of Chicago and University of Illinois.

Ané, T. and Geman, H.: 2000, Order flow, transaction clock, and normality of asset returns, *Journal of Finance* **55**(5), 2259–84.

Alizadeh, S., Brandt, M. W. and Diebold, F. X.: 2002, Range-based estimation of stochastic volatility models, *Journal of Finance* **57**(3), 1047–90.

Andersen, T. G., Bollerslev, T. and Diebold, F. X.: 2005, Parametric and nonparametric volatility measurement, *in* Y. Aït-Sahalia and L. P. Hansen (eds), *Handbook of Financial Econometrics*, North Holland, Amsterdam.

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

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

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P.: 2001b, Great realizations, *Risk* **13**, 105–108.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Vega, C.: 2003, Micro effects of macro announcements: Real-time price discovery in foreign exchange, *American Economic Review* **93**, 38–62.

Bandi, F. and Russell, J. R.: 2005, Microstructure noise, realized volatility, and optimal sampling. Working Paper, Graduate School of Business, University of Chicago.

Bandi, F. and Russell, J. R.: 2009, Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations, *Journal of Econometrics*, forthcoming.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N.: 2004, Regular and modified kernel-based estimators of integrated variance: The case with independent noise. Working paper, http://ssrn.com/abstract=620203.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N.: 2008, Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise, *Econometrica* . Forthcoming.

Barndorff-Nielsen, O. E. and Shephard, N.: 2001, Non-gaussian ornstein-uhlembech-based models and some of their uses in financial economics, *Journal of the Royal Statistical Society* **Series B**(63), 167–241.

Barndorff-Nielsen, O. E. and Shephard, N.: 2002a, 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. and Shephard, N.: 2002b, Estimating quadratic variation using realized variance, *Journal of Applied Econometrics* **17**(5), 457–477.

Barndorff-Nielsen, O. E. and Shephard, N.: 2005, How accurate is the asymptotic approximation to the distribution of realized volatility?, *in* D. W. F. Andrews and J. H. Stock (eds), *Identification and Inference for Econometric Models. A Festschrift in Honour of T.J. Rothenberg*, Cambridge University Press, pp. 306–331.

Blume, M. and Stambaugh, S.: 1983, Biases in computed returns, *Journal of Financial Economics* **12**, 387–404.

Bollen, B. and Inder, B.: 2002, Estimating daily volatility in financial markets utilizing intraday data, *Journal of Empirical Finance* **186**.

Christensen, K. and Podolskij, M.: 2006, Realized range-based estimation of integrated variance, *Journal of Econometrics* . To appear in.

Clark, P. K.: 1973, A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* **41**(1), 135–55.

Cohen, K., S., M., R., S. and D., W.: 1981, Transaction costs, order placement strategy, and existence of the bid-ask spread, *Journal of Political Economy* **89**, 287–305.

Comte, F. and Renault, E.: 2001, Long memory in continuous time stochastic volatility models, *Mathematical Finance* **8**, 291–323.

Corsi, F., Zumbach, G., Müller, U. A. and Dacorogna, M.: 2001, Consistent high-precision volatility from high-frequency data, *Economic Notes* **30**(2), 183–204.

Elliott, J., F.: 1953, The characteristic roots of certain real symmetric matrices, *Technical report*, University of Tennessee. Mater Thesis.

French, K. R. and Roll, R.: 1986, Stock return variances : The arrival of information and the reaction of traders, *Journal of Financial Economics* **17**(1), 5–26.

Glosten, L.: 1994, Is the electronic open limit-order book inevitable?, *Journal of Finance* **49**, 1127–1161.

Glottlieb, G. and Kalay, A.: 1985, Implication of the discreteness of observed stock prices, *Journal of Finance* **40**, 135–153.

Gregory, Robert, T. and Karney, D.: 1969, A collection of matrices for testing computational algorithms, *Wiley-Interscience* .

Hamilton, J.: 1994, *Time Series Analysis*, Princeton: Princeton University Press.

Hansen, P. R. and Lunde, A.: 2004, An unbiased measure of realized variance. Working Paper, Department of Economics, Stanford University.

Hansen, P. R. and Lunde, A.: 2006, Realized variance and market microstructure noise, *Journal of Business and Economic Statistics* **24**(127-218).

Harris, L.: 1990, Estimation of stock price variances and serial covariances from discrete observations, *Journal of Financial and Quantitative Analysis* **25**(3), 291–306.

Hasbrouck, J.: 1993, Assessing the quality of a security market: a new approach to transaction-cost measurement, *Review of Financial Studies* **6**(1), 191–212.

Hasbrouck, J.: 1996, Modelling market microstructure time series, *in* G. S. Maddala and C. R. Rao (eds), *Statistical Methods in Fianance*, Vol. 14 of *Handbook of Statistics*, North Holland, Amsterdam, pp. 647–692.

Hasbrouck, J.: 1999, The dynamichs of discrete bid and ask quotes, *Journal of Finance* **54**, 2109–42.

Hasbrouck, J. and Ho, T.: 1987, Order arrival, quote behavior, and the return-generating process, *Journal of Finance* **42**, 1035–48.

Heston, S.: 1993, A closed-form solution for options with stochastic volatility with applications to bonds and currency options, *Review of Financial Studies* **6**, 327–343.

Lan, Z.: 2006, Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach., *Bernoulli* **12**, 1019–1043.

Martens, M. and van Dijk, D.: 2007, Measuring volatility with the realized range, *Journal of Econometrics* **138**(1), 181–207.

Merton, R. C.: 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* **8**, 323–61.

Nolte, I. and Voev, V.: 2008, Estimating high-frequency based (co-) variances: A unified approach. Working Paper.

Oomen, R. C.: 2005, Properties of bias-corrected realized variance under alternative sampling schemes, *Journal of Financial Econometrics* **3**(4), 555–577.

Oomen, R. C.: 2006, Properties of realized variance under alternative sampling schemes, *Journal of Business and Economic Statistics* **24**, 219–237.

Parkinson, M.: 1980, The extreme value method for estimating the variance of the rate of return, *Journal of Business* **53**, 61–65.

Phillips, P. C. B. and Yu, J.: 2006, Discussion of hansen & lunde (2006), *Journal of Business and Economic Statistics* **24**, 202–208.

Roll, R.: 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* **39**, 1127–1139.

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

Zhou, B.: 1996, High-frequency data and volatility in foreign-exchange rates, *Journal of Business & Economic Statistics* **14**(1), 45–52.

Zhou, B.: 1998, Parametric and nonparametric volatility measurement, *in* C. L. Dunis and B. Zhou (eds), *Nonlinear Modelling of High Frequency Financial Time Series*, John Wiley Sons Ltd, New York, pp. 109–123.

Zumbach, G. O., Corsi, F. and Trapletti, A.: 2002, Efficient estimation of volatility using high frequency data, *Working Paper, Olsen & Associates* .