Forecasting Volatility with the Realized Range in the Presence of Noise and Non–Trading

Karim Bannouh   Martin Martens   Dick van Dijk

25–10–2012

Abstract

We introduce a heuristic bias-adjustment for the transaction price-based realized range estimator of daily volatility in the presence of bid-ask bounce and non-trading. The adjustment is an extension of the estimator proposed in Christensen et al. (2009). We relax the assumption that all intra-day high (low) transaction prices are at the ask (bid) quote. Using data-based simulations we obtain estimates of the probability that a given intraday range is (upward or downward) biased or not, which we use for a more refined bias-adjustment of the realized range estimator. Both Monte Carlo simulations and an empirical application involving a liquid and a relatively illiquid SP500 constituent demonstrate that ex post measures and ex ante forecasts based on the heuristically adjusted realized range compare favorably to existing bias-adjusted (two time scales) realized range and (two time scales) realized variance estimators.

Keywords: Realized variance; realized range; two time scales; high frequency data; market microstructure noise; forecasting

JEL codes: C58; C53; G17

ERIM Report Series *Research in Management*	
ERIM Report Series reference number	ERS–2012–018-F&A
Date of publication	2012–10–25
Version	25–10–2012
Number of pages	27
Persistent URL for paper	`http://hdl.handle.net/1765/37538`
Email address corresponding author	djvandijk@ese.eur.nl
Address	Erasmus Research Institute of Management (ERIM) RSM Erasmus University / Erasmus School of Economics Erasmus University Rotterdam PO Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31104081182 Fax: +31104089640 Email: info@erim.eur.nl Internet: `http://www.erim.eur.nl`
Availability	The ERIM Report Series is distributed through the following platforms: RePub, the EUR institutional repository Social Science Research Network (SSRN) Research Papers in Economics (RePEc)
Classifications	The electronic versions of the papers in the ERIM Report Series contain bibliographic metadata from the following classification systems: Library of Congress Classification (LCC) Journal of Economic Literature (JEL) ACM Computing Classification System Inspec Classification Scheme (ICS)

Forecasting Volatility with the Realized Range in the Presence of Noise and Non-Trading

Karim Bannouh

Econometric Institute
Erasmus University Rotterdam

Martin Martens

Department of Finance
Erasmus University Rotterdam

Dick van Dijk*

Econometric Institute
Erasmus University Rotterdam

October 25, 2012

Abstract

We introduce a heuristic bias-adjustment for the transaction price-based realized range estimator of daily volatility in the presence of bid-ask bounce and non-trading. The adjustment is an extension of the estimator proposed in Christensen et al. (2009). We relax the assumption that all intra-day high (low) transaction prices are at the ask (bid) quote. Using data-based simulations we obtain estimates of the probability that a given intraday range is (upward or downward) biased or not, which we use for a more refined bias-adjustment of the realized range estimator. Both Monte Carlo simulations and an empirical application involving a liquid and a relatively illiquid S&P500 constituent demonstrate that *ex post* measures and *ex ante* forecasts based on the heuristically adjusted realized range compare favorably to existing bias-adjusted (two time scales) realized range and (two time scales) realized variance estimators.

Keywords: Realized variance; realized range; two time scales; high frequency data; market microstructure noise; forecasting

JEL Classification: C58, C53, G17

---

*Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, Netherlands, e-mail: `djvandijk@ese.eur.nl` (corresponding author).

1 Introduction

Measuring and forecasting the volatility of asset returns plays a key role in various areas of financial economics, including portfolio management, risk management and the pricing of derivatives. The increasing availability of high-frequency asset price data has triggered a vast amount of academic studies proposing volatility estimators that exploit intraday prices to estimate and forecast daily volatility measures.

The realized variance (RV) estimator sums squared non-overlapping intraday returns to estimate the daily variance, see e.g. Andersen et al. (2001). In a frictionless market with continuous trading, RV converges to the integrated variance (IV) as the sampling frequency of the intraday returns increases. In practice, however, high-frequency asset prices are contaminated with market microstructure noise. This causes potentially severe problems in terms of consistent estimation of the daily IV by means of realized measures, see McAleer and Medeiros (2008) for a review. For RV estimators based on intraday returns obtained from transaction prices the dominant source of market microstructure noise is bid-ask bounce. Transactions take place at bid and ask prices causing an upward bias in the RV estimator. The magnitude of the bias increases with the sampling frequency.

A pragmatic solution to circumvent the problems arising from bid-ask bounce is to sample returns more sparsely by using longer intraday intervals; examples include the popular 5- and 30-minute frequencies. While lowering the sampling frequency reduces the bias in RV estimators, it also increases the variance. The use of sparse sampling frequencies aims to strike a balance between these two aspects. More formal approaches to correct for the effects of bid-ask bounce and other types of microstructure noise also exist. Among the most popular bias-correction methods is the two time scales RV (TSRV) estimator of Zhang et al. (2005). In this approach the variance of the difference between the observed transaction prices and the latent efficient prices is estimated using the highest sampling frequency available and this is then subtracted from each of the intraday squared returns.

Martens and Van Dijk (2007) and Christensen and Podolskij (2007) propose the realized range (RR) estimator as a more efficient measure of ex-post volatility. The RR estimator replaces the squared intraday returns in the RV estimator by squared intraday ranges. The results of Martens and Van Dijk (2007) illustrate that in a frictionless market the RR estimator is indeed more efficient than the RV estimator when comparing similar sampling frequencies. These results continue to hold in settings where market microstructure noise,

in particular bid-ask bounce, is present.

The use of intraday ranges for volatility measurement is further complicated by a different source of market microstructure noise, namely infrequent trading. Trading does not occur continuously, that is, in practice we observe transactions at irregularly spaced points in time, see e.g. Engle (2000) or Griffin and Oomen (2008). For the RV estimator, non-trading increases the variance but does not cause a bias. In contrast, infrequent trading introduces a downward bias in RR estimators as the observed intraday high and low prices are likely to be below and above their 'true' values, respectively.[1] Christensen and Podolskij (2007) propose an adjustment of the standard RR estimator to account for the effects of non-trading.

Returning to the issue of bid-ask bounce, Christensen et al. (2009) propose a 'two time scales' RR (TSRR) estimator that aims to correct the upward bias due to bid-ask bounce along the same lines as the TSRV estimator of Zhang et al. (2005). The two time scales RR is implemented by estimating the bid-ask spread using the highest sampling frequency available and subtracting this quantity from each of the intraday ranges.

In this paper we extend the bias-adjustment for the realized range presented in Christensen et al. (2009) by relaxing their assumption that the observed high (low) price in each intraday interval originates from a transaction taking place at the ask (bid) quote. While this may be the most likely situation, in practice the high (low) price may also be observed as a transaction at the bid (ask) quote, such that an intraday range is not necessarily upward biased. Intuitively, the likelihood of an intraday range being upward biased decreases when the noise-to-volatility ratio becomes smaller or when the trading intensity of the asset becomes lower. We propose a heuristic adjustment of the RR that utilizes simulation-based estimates of the probabilities of an intraday range being upward biased, downward biased or unbiased. For the heuristic adjustment we need three inputs that are readily available from a sample path of tick data for a full trading day for which one wants to estimate the daily volatility. These inputs are estimates of the following quantities: (*i*) the daily range that is unaffected by noise, (*ii*) the non-trading probability and (*iii*) the half-spread. Using these inputs we simulate a geometric Brownian motion with variance (*i*) and implement noise with settings (*ii*) and (*iii*). For the simulated geometric Brownian motions we keep count of how many intraday ranges are upward biased,

---

[1]Note that a possible advantage of the 'standard' realized range estimator is that the positive bias due to bid-ask bounce and the downward bias due to non-trading offset each other to a certain extent.

2

unbiased or downward biased. By averaging over simulation runs we estimate probabilities for the three cases that can be attached to the ranks of sorted intraday ranges. We apply these probability ranks to the sorted vector of initial high-low ranges for which we are now able to indicate whether an intraday range is expected to be upward biased, unbiased or downward biased.

We study the proposed heuristic bias-adjustment for the realized range estimator in a simulation setting with plausible levels of bid-ask bounce and non-trading. Using Monte Carlo simulations with several different stochastic volatility models as data generating process we find that the heuristically adjusted realized range estimator TSRRh provides volatility estimates that compare favorably, in terms of bias and variance, with the (TS)RV and (TS)RR estimators studied in Christensen et al. (2009) and the (TS)RV estimators in Aït-Sahalia and Mancini (2008). In an empirical forecasting application for the relatively liquid IBM stock and Zimmer Holdings (ZMH), a relatively illiquid constituent of the S&P500 belonging to the health care sector, we also find encouraging results. For IBM the heuristically adjusted RR volatility estimator provides more efficient one-step ahead forecasts. For ZMH the TSRRh outperforms (TS)RV and TSRR and competes with the RR estimator.

Our paper is related to several recent articles examining the relative performance of different realized measures in terms of measuring and forecasting the daily integrated variance. Among the studies that focus on out-of-sample predictive ability, Liu et al. (2012) recently consider the model confidence set approach to test for 350 assets, selected from several asset classes, whether alternative volatility forecasts can beat RV forecasts. They conclude that there are better forecasts but that it is difficult to significantly improve upon the RV forecasts. Their study includes the realized range which is implemented in the form proposed by Christensen and Podolskij (2007), which takes non-trading into account but is not unadjusted for other forms of microstructure noise. They find that the realized range forecasts compare favorably, especially for interest rate futures. Aït-Sahalia and Mancini (2008) put forward forecasting results for TSRV and RV measures in the presence of jumps, noise correlated with the efficient price, autocorrelated noise, long-memory in volatility and leverage effects in volatility. In addition they compare TSRV and RV forecasts for the relatively liquid DJIA stocks. They find that TSRV forecasts are more efficient than RV forecasts. Andersen et al. (2011) evaluate out-of-sample volatility forecasts in a simulation setting that uses stochastic volatility diffusions.

The resulting efficient price processes are contaminated with microstructure noise. Their analysis is extended in several dimensions such as an implementation where the noise is serially correlated. They find that a combination of the TSRV and a RV estimator constructed by weighting different sampling frequencies performs best. Ghysels and Sinko (2011) evaluate volatility forecasts in the Mixed Data Sampling (MIDAS) framework and include results for iid-distributed noise and dependent noise. Consistent with Aït-Sahalia and Mancini (2008) they find that at high sampling frequencies TSRV forecasts achieve the highest efficiency. Christensen et al. (2009) compare (TS)RV and (TS)RR estimators and find that in the presence of bid-ask bounce TSRR and TSRV compete in terms of statistical efficiency and that TSRR is more efficient when more than 300 observations are available. In an empirical application Christensen et al. (2009) estimate the volatility of two highly liquid IT stocks, Microsoft and INTEL, and find that (TS)RV estimators have a smaller variance than RR. The TSRR they propose, however, has a smaller variance than the (TS)RV estimators.

The remainder of this paper is structured as follows. In Section 2 we develop the heuristic bias-adjustment for the RR estimator and discuss the (two time scales) realized volatility and (two time scales) realized range estimators. The simulation results are discussed in Section 3. Empirical forecasting results are presented in Section 4. We conclude in Section 5.

2 Volatility estimators, noise and bias-corrections

2.1 Volatility estimators

We assume that the logarithmic asset price P_t follows a driftless diffusion

$$dP_t = \sigma_t dW_t, \tag{1}$$

where σ is a strictly positive stochastic volatility process and W_t is a Wiener process. The daily interval is standardized to unity, such that the daily integrated variance (IV) is given by

$$IV_t = \int_{t-1}^{t} \sigma_s^2 ds. \tag{2}$$

Let $r_{t,j}^\Delta = \log P_{t+j\Delta} - \log P_{t+(j-1)\Delta}$ denote the log-return over the j-th intra-day interval of length Δ on day t, for a given interval length $0 < \Delta < 1$ such that we have $J = 1/\Delta$

intervals in a given day.[2] The realized variance estimator is calculated by summing squared intraday returns that are sampled from non-overlapping intervals of length Δ,

$$RV_t^\Delta = \sum_{j=1}^{J} r_{t,j}^2. \tag{3}$$

The realized range replaces the squared returns in RV by squared intraday ranges,

$$RR_t^\Delta = \frac{1}{4\log 2} \sum_{j=1}^{J} (\log H_{t,j} - \log L_{t,j})^2, \tag{4}$$

where $H_{t,j} = \sup_{(j-1)\Delta \leq i \leq j\Delta} P_{t+i}$ and $L_{t,j} = \inf_{(j-1)\Delta \leq i \leq j\Delta} P_{t+i}$ denote the high and low prices during the j-th interval on day t. In a frictionless market environment with continuous trading, both RV_t and RR_t are consistent estimates of the integrated variance IV_t when the sampling frequency $J \to \infty$. In the constant volatility case $\sigma_t = \sigma$ the variance of RV is $2\sigma^4\Delta^2$ and the variance of RR is approximately[3] $0.407\sigma^4\Delta^2$, which renders the RR about 5 times more efficient.

2.2 Market microstructure noise

Market microstructure noise refers to imperfections in the trading process of financial assets causing observed prices to deviate from the underlying 'true' price process. Microstructure noise generally implies that realized volatility and realized range measures are inconsistent estimators for the integrated variance, with the impact becoming more pronounced as the sampling frequency increases. We focus on bid-ask bounce and non-trading since these are the two most relevant sources of noise that affect range-based volatility estimates based on high-frequency intra-day transaction prices.

Bid-ask bounce

Observed transactions take place at bid and ask quotes causing negative autocorrelation in high-frequency returns as the observed price jumps transiently from ask to bid and vice versa, see e.g. Roll (1984). Hence, at the micro level bid-ask bounce introduces volatility in the observed price process that is unrelated to the volatility of the 'true' price process. For this reason bid-ask bounce causes an upward bias in high-frequency volatility estimates.

---

[2]For convenience we assume that Δ is such that J is an integer.

[3]The exact variance of the RR is $(\frac{9\zeta(3)}{(4\log 2)^2} - 1)\sigma^4\Delta^2$ where $\zeta(x) = \sum_{m=1}^{\infty} 1/m^x$ is Riemann's zeta function.

A general representation of bid-ask bounce and the relationship between the 'efficient' price P_t and the 'noisy' transaction price P_t^* is given by:

$$P_t^* = P_t + \omega_t, \tag{5}$$

where bid-ask bounce is represented by ω_t which follows an i.i.d. distribution with support on $+\omega$ and $-\omega$, such that ω represents the half-spread.

Infrequent trading

Strictly speaking, non-trading does not fall under the heading of microstructure noise as defined above, in the sense that observed transaction prices are (or can be) equal to the efficient price. As the price process is not observed continuously though, non-trading does affect the RR estimator. As the observed high and low prices in a given intraday interval are likely to be below and above their 'true' values, respectively, infrequent trading introduces a downward bias in the 'standard' RR estimator in (4). Effectively, in the presence of non-trading the scaling parameter $4 \log 2$, which is the variance of a continuously observed Brownian motion, is not appropriate. Following Christensen and Podolskij (2007), we therefore use

$$RR_t^\Delta = \frac{1}{\lambda_m} \sum_{j=1}^{J} (\log H_{t,j} - \log L_{t,j})^2, \tag{6}$$

where m is the number of observations in an intraday range. The appropriate scaling parameter $\lambda_m = \mathsf{E}[\max_{0 \leq s,t \leq m} (W_{t/m} - W_{s/m})^2]$ is determined through simulating an infrequently observed Brownian motion W and estimating the second moment of its range. Note that this adjustment destroys the possibility that the upward bias due to bid-ask bounce and the downward bias due to infrequent trading (partly) offset each other, necessitating a further adjustment of (6) to account for the effects of microstructure noise.

2.3 Correcting for bid-ask bounce

Subsampling aims at at improving the accuracy of realized measures by using multiple intraday sample paths through shifting the point at which a sample starts. Assuming one has access to 1-minute price observations at 9:30, 9:31, 9:32, etc. the standard approach to estimate RV using, for example, 5-minute returns is to use transaction prices at 9:30, 9:35, 9:40 etc. A way to exploit more of the available data is to use a 5-minute price sample consisting of observations 9:31, 9:36, 9:41 etc. This approach provides five different samples

giving rise to five different RV estimates. These can be averaged such that more data is used. The number of subsamples one can compute depends on the 'intended' sampling frequency and on the highest sampling frequency available. Assuming that there are S subsamples, the subsampled RV^S estimator is defined as:

$$RV_t^{\Delta,S} = \frac{1}{S}\sum_{s=1}^{S} RV_{t,s}^{\Delta}. \tag{7}$$

The two time scales estimator introduced in Zhang et al. (2005) combines the subsampled RV^S estimator at a 'sparse' frequency, e.g. 5-minutes, with an ultra-high-frequency estimator that is used to estimate the noise component. At the ultra-high-frequency RV is estimated using all of the $n+1$ observed price ticks in a trading day and is denoted RV^N. This 'all returns' estimator produces a consistent estimate of the quantity $2n\mathsf{E}(\omega^2)$ such that $\mathsf{E}(\omega^2) = RV^N/2n$. Combining the sparsely subsampled RV^S estimator and the 'all returns' estimate to remove the noise results in a consistent estimator of the integrated variance, the so-called two-time-scales realized variance (TSRV) estimator:

$$TSRV_t^{\Delta} = RV_t^{\Delta,S} - \frac{\bar{n}}{n}RV^N, \tag{8}$$

where $\bar{n} = n/S$. A small sample adjustment is applied to adjust for the fact that the number of returns in each of the sub-grids may not be equal:

$$TSRV_t^{\Delta,adj} = \frac{1}{1-\frac{\bar{n}}{n}}TSRV_t^{\Delta}. \tag{9}$$

For sufficiently large samples the correction term converges to unity. The TSRV estimator uses all available intraday price observations to estimate the noise component. For the RV subsampler at sparse frequencies, however, TSRV does not necessarily use all of the available data. Range-based volatility estimators by construction use all of the available data to calculate the highs and lows in an interval, and hence, make more efficient use of the high-frequency data to estimate volatility.

Similar to the TSRV estimator, Christensen et al. (2009) propose the use of a bias-correction for the realized range estimator based on two time scales. The bias-correction is derived under the assumption that the noise is represented by bid-ask bounce[4], i.e. an iid-noise distribution centered around zero with support on only two points, see also Equation

---

[4]It is hard, if not impossible, to derive a bias-adjustment for the RR estimator under noise distributions with unlimited support. Christensen et al. (2009) provide extensions to other microstructure noise distributions with bounded support such as a uniform noise distribution and rounding errors. The focus in their study, however, is also mainly on bid-ask bounce.

(5) for a general representation. The highest frequency time scale is used to estimate the impact of bid-ask bounce. Specifically, a consistent estimate of the half-spread is obtained using $\hat{\omega} = \sqrt{RV^{all}/2n}$. This quantity is then used to filter out the bid-ask spread ω in each interval of the sparsely sampled realized range estimator:

$$TSRR_t^{\Delta} = \frac{1}{\tilde{\lambda}_m} \sum_{j=1}^{J} (\log H_{t,j} - \log L_{t,j} - \gamma\hat{\omega})^2, \qquad (10)$$

where Christensen et al. (2009) use $\gamma = 2$ which is based on the implicit assumption that $H_{t,j}$ is always at the ask-quote and $L_{t,j}$ is always at the bid-quote. The scaling parameter $\tilde{\lambda}_m = \mathsf{E}[\max_{s:\omega_{s/m}=-\omega,t:\omega_{t/m}=\omega} (W_{t/m} - W_{s/m})^2]$ is determined through estimating the variance of the range of a discretely observed Brownian motion that is contaminated with noise.

The TSRR proposed in Christensen et al. (2009) takes into account that observed prices are contaminated by bid-ask bounce and that prices are observed infrequently. The latter is done through the multiplicative scaling parameters λ_m and $\tilde{\lambda}_m$ which take on different values for RR and TSRR due to microstructure noise. Underlying the additive part of the bid-ask correction where $\gamma = 2$, is the implicit assumption that the high is always an ask price and the low is always a bid price. In the presence of plausible levels of bid-ask bounce and non-trading, however, the probabilities of an intraday range being unbiased or downward biased are non-zero. The assumption of all intraday ranges being upward biased only holds when an asset trades very frequently throughout the day and a sufficiently large number of transactions is recorded in each of the intraday sampling intervals. In addition, the noise-to-volatility ratio should be sufficiently large. For illiquid assets such as stocks that are traded infrequently this assumption may not always hold. This can be exemplified by analyzing an artificial price path where in some specific intraday interval the high and low are equal, i.e. this interval should not contribute to the daily volatility. For the RV and RR estimators this is the case, as both the intraday return and range are zero for this interval and do not contribute to the daily volatility estimates. This specific interval will, however, introduce an upward bias in TSRR of $\frac{4\hat{\omega}^2}{\tilde{\lambda}}$. This upward bias for a specific interval also occurs when the high and low are non-equal but both were recorded at the bid quote (ask quote). For these reasons we relax the assumption that the observed high (low) price always originates from a transaction executed at the ask (bid) quote. Specifically, we use simulation-based estimates of the probabilities that a specific intraday range is unbiased or even downward biased. The underlying idea is that if one would sort all the observed intraday highs (lows), then the highest high (lowest low) is more likely to be at

8

the ask-quote (bid-quote) than is the case for the lowest (highest) observed high (low).

In more detail, we propose the following bias adjustment procedure that is based on simulation and sorting. Given a trading day of tick data that is contaminated by noise and infrequent trading:

1. Estimate the non-trading probability using the number of observed transactions on day t.

2. Use Parkinson (1980)'s daily high-low range estimator to obtain an initial estimate of the volatility for day t.[5]

3. Estimate bid-ask bounce, i.e. $\hat{\omega} = \sqrt{\mathsf{E}(\hat{\omega}^2)} = \sqrt{RV^N/2n}$.

4. Simulate intraday sample-paths based on a geometric Brownian motion with inputs being the estimated non-trading probability, the initial volatility estimate and the estimated bid-ask spread.

5. Using the bid-ask and non-trading contaminated simulated sample paths, estimate the probability of observing (a) no bias, (b) upward bias and (c) downward bias in the intraday range.[6]

6. Sort the empirical intraday high-low's. Based on the estimated probabilities from the previous step, calculate how many of the intraday ranges are expected to be (a) unbiased, (b) upward biased or (c) downward biased. Use Equation 11 and apply (a) $\gamma_j = 0$, (b) $\gamma_j = 2$ and (c) $\gamma_j = -2$ to adjust for (b) upward bias and (c) downward bias.

Hence, our estimator has the same form as the estimator proposed in Christensen and Podolskij (2007) with the difference being that we do not use $\gamma = 2$ to correct each of the intra-day ranges. Instead we propose to use

$$TSRRh_t^\Delta = \frac{1}{\tilde{\lambda}} \sum_{j=1}^{J} (\log H_{t,j} - \log L_{t,j} - \gamma_j \hat{\omega})^2, \tag{11}$$

---

[5]It is important that this estimator is (almost) not affected by microstructure noise (we will use the daily range, alternatively one can use another (almost) bias-free measure, e.g. the TSRV or the daily squared return).

[6]Case (a) occurs when in an intra-day interval the observed high and low are both executed at a bid price (or both being an ask), (b) occurs when the observed high is an ask-price and the observed low is a bid-price (c) occurs when the high is a bid-price and the low is an ask price.

where we use $\gamma_j = 2$ if after sorting and using the simulated probabilities an intraday range is expected to be biased upward (b). Assuming that the J intraday ranges are sorted in a descending manner and the estimated probability of intraday ranges being biased upward is q, then the first Jq intraday ranges are expected to be biased upward. Similarly, assuming that the probability of an intraday being unbiased is estimated to be v, we use $\gamma_j = 0$ (a) for the subsequent Jv intraday ranges and for the remaining $J(1 - q - v)$ intraday ranges $\gamma_j = -2$ (c) is used.[7]

3 Monte Carlo Simulation

In the following Monte Carlo simulation experiments we compare ex-post volatility estimates using the (TS)RV and (TS)RR estimators with the newly proposed TSRRh estimator. The estimators are compared in terms of bias, variance and efficiency. We simulate the integrated variance using several stochastic volatility diffusions that were also used in Aït-Sahalia and Mancini (2008), among others. Returns and integrated volatilities are simulated from a Heston Jump-Diffusion, a Fractional Ornstein-Uhlenbeck process and a discrete-time log-volatility model. We simulate 1,000 trading days of 6.5 hours, i.e. 23,401 prices are simulated per day to match a time step of 1 second. Subsequently non-trading is implemented by assuming a trade is observed with probability 0.10 such that on average 2,340 'clean' prices are observed during the day. Microstructure noise is implemented by contaminating the prices with a half-spread of $\omega = 0.025\%$ on the asset price. Bid and ask prices are assumed to occur equally likely. In all experiments we use 100 sub-sample grids to calculate TSRV. For each daily TSRRh estimate 500 simulations are used to estimate the impact of bid-ask bounce for rank-sorted intraday ranges in order to implement the proposed bias-adjustment as in Equation (11).

3.1 Heston stochastic volatility jump-diffusion

The data generating process for returns and volatility under the Heston (1993) stochastic volatility jump-diffusion model is specified by

$$dP_t = (\mu - \sigma_t^2/2)dt + \sigma_t dW_{1,t},$$
$$d\sigma_t^2 = -\kappa(\sigma_t^2 - \alpha)dt + \gamma\sigma_t dW_{2,t} + J_t dq_t,$$

---

[7] Assuming Jq and Jv are integer.

10

with drift parameter $\mu = 5\%$, a long term average volatility $\alpha = 3.5\%$, and mean reversion parameter $\kappa = 5$. The volatility of volatility parameter $\gamma = 0.5$ facilitates leverage effects as the two Brownian motions are negatively correlated with $\rho = -0.5$. The occurrence of jumps in the volatility process has distribution $q_t \sim Poi(\phi)$ and the jump magnitude follows an exponential distribution $J_t \sim Exp(\zeta)$. Following Aït-Sahalia and Mancini (2008) we set $\lambda = 1/2$, $\zeta = 0.0007$. Empirical stylized facts are taken into account by the inclusion of jumps in the volatility process and a leverage effect to allow for the empirically plausible negative relation between returns and volatility shifts.

3.2 Fractional Ornstein-Uhlenbeck process

Following Aït-Sahalia and Mancini (2008) we simulate IV using a fractional Brownian motion,

$$dP_t = (\mu - \sigma_t^2/2)dt + \sigma_t dW_t,$$

$$d\sigma = -\kappa(\sigma_t - \alpha)dt + \gamma dW_{H,t},$$

where dW_t is a Wiener process and $dW_{H,t}$ is a fractional Brownian motion with Hurst index $H \in (0,1)$. A fractional Brownian motion is a continuous mean zero Gaussian process with stationary increments and covariance $E(W_{H,t}W_{H,s}) = \frac{1}{2}(s^{2H} + t^{2H} - |s - t|^{2H})$. The covariance structure illustrates that the increments are positively correlated when $\frac{1}{2} < H < 1$ and exhibit long-memory, for $H = \frac{1}{2}$ the increments are independent and correspond to a standard Brownian motion. To simulate the fractional Brownian motion we use the Davies and Harte (1987) algorithm with Hurst effect $H = 0.7$.

3.3 Discrete-time log-volatility model

In many applications the logarithm of volatility is used because the logarithm of (realized) volatility is empirically found to be closer to a Gaussian distribution (see e.g. Figure 1 in Andersen et al. (2001)). The discrete time model we use is the model employed in Andersen et al. (2003) and Aït-Sahalia and Mancini (2008). The daily integrated volatility l_t follows an AR(5) process

$$l_t = \frac{1}{2}\log(IV_t) = \phi_0 + \sum_{i=1}^{5}\phi_i l_{t-id} + e_t, \tag{12}$$

where IV_t is the daily integrated variance and e_t is white noise. Intraday efficient returns are obtained using $r_t = \sqrt{IV_t}z_t$ with $z_t \sim NID(0,1)$. For the parameters we use those

reported by Aït-Sahalia and Mancini (2008), $\phi_0 = -0.0161$, $\phi_1 = 0.35$, $\phi_2 = 0.25$, $\phi_3 = 0.20$, $\phi_4 = 0.10$, $\phi_5 = 0.09$ and $\sigma_e = 0.02$.

3.4 Monte Carlo results

Volatility estimation results using Monte Carlo simulations for the three stochastic volatility models[8] discussed above are summarized in Table 1. The microstructure noise settings used are a probability equal to 0.10 of observing a trade[9], which results in 2,340 observations per day on average and a half-spread of 0.025% of the asset price.

Under the Heston jump-diffusion the bias for the RV estimator (0.093) is somewhat smaller than would be expected based on using a half-spread of 0.025% of the asset price $(0.0975 = 2*390/5*0.025\%^2)$[10]. This is due to the quadratic variation being larger because of jumps in the volatility process. The variance of all the volatility estimators considered under the Heston jump-diffusion models is considerably larger than in models that do not incorporate jumps since the volatility estimators discussed here are not designed to be jump-robust. Theoretically the RR estimator is expected to have a substantially smaller variance than the (TS)RV estimators. It is interesting to compare the competing estimators in the presence of noise, non-trading and jumps in the volatility process. Indeed we find that at the 5-minute sampling frequency the variance of RR (0.071) is still more than 3 times smaller than the variance of RV (0.253) and less than half the variance of TSRV (0.170). In terms of RMSE the RR (0.460) performs better than RV (0.511) but in turn it is outperformed by the TSRV (0.414) because the latter is approximately unbiased (−0.038). The bias of the RR (0.374) estimator is substantially larger than the bias in the RV estimator. Bias-correcting the realized range as proposed by Christensen et al. (2009) successfully reduces the bias from 0.374 to −0.263 at the cost of an increase in variance from 0.071 for RR to 0.108 for TSRR. Despite the reduced bias, the TSRR (0.421) still does not improve upon TSRV (0.414). Taking into account that not all intraday ranges are upward biased and that the largest intraday ranges in a day are more likely to be upward biased

---

[8]Results for a Brownian motion with constant volatility are similar in the sense that TSRRh improves upon (TS)RV because of having a smaller variance leading to a smaller RMSE. The TSRRh also improves upon (TS)RR because of a smaller bias that comes at the cost of a modest increase in variance. This bias-variance trade-off results in TSRRh having a smaller RMSE than (TS)RR as well. Results are available upon request.

[9]The trading probability is in line with the results presented in Table 1 in Hansen and Lunde (2006).

[10]Errors are multiplied with 10^4 to improve readability.

than the smallest intraday ranges is exemplified by TSRRh (-0.239) having a smaller bias than TSRR (-0.263). As a result the RMSE of TSRRh (0.407) is also smaller than the RMSE (0.414) of the unbiased TRSV estimator. At the 30-minute sampling frequency the impact of noise is substantially smaller as expected[11] and for this reason it is optimal to use the RR without bias-correction.

Across models we find that using 5-minute intervals to estimate daily volatility outperforms the lower 30-minute and daily sampling frequencies in terms of variance and statistical efficiency. Under the fractional Brownian motion model the TSRV estimator minimizes the bias (-0.023) at the 5-minute sampling frequency as was the case under the Heston model. Again the realized range-based estimators achieve a smaller variance than (TS)RV. However, it is also the most biased estimator and for this reason the least efficient with a RMSE of 0.422. The TSRV (-0.023) successfully reduces the bias of RV (0.098) and achieves a RMSE of 0.219. Similarly the TSRR is very successful in reducing the bias of RR (0.395) to -0.138 and also has a smaller RMSE (0.208) than the (TS)RV estimators. By using the informational content contained in the size of the intraday ranges through implementing the TSRRh the bias is further reduced from -0.138 for TSRR, down to -0.120 for TSRRh which results in TSRRh having the smallest RMSE (0.199).

For the discrete-time log-volatility model we find similar results in the sense that at the 5-minute sampling frequency the TSRV estimator minimizes the bias (-0.037) but has a variance (0.048) that is inferior to that of the RR (0.022), TSRR (0.024) and TSRRh (0.025) estimators. The TSRRh (-0.134) is less biased than the TSRR (-0.151) which in turn is less biased than RR (0.380). The result is that, similar to the results under the Heston Jump-Diffusion and the fractional Brownian motion model, the TSRRh at the 5-minute sampling frequency achieves the smallest RMSE in the discrete-time log-volatility model.

---

[11]For instance the expected RV bias is now only 0.01625.

Table 1: Monte Carlo Results for Volatility Estimators under Stochastic Volatility Diffusions

DGP	Heston Jump Diffusion			Fractional Brownian Motion			Discrete Log-volatility		
Estimator	Bias	Variance	RMSE	Bias	Variance	RMSE	Bias	Variance	RMSE
RV-5min.	0.093	0.253	0.511	0.098	0.075	0.292	0.084	0.075	0.287
TSRV-5min.	-0.038	0.170	0.414	-0.023	0.047	0.219	-0.037	0.048	0.222
RR-5min.	0.374	0.071	0.460	0.395	0.022	0.422	0.380	0.022	0.408
TSRR-5min.	-0.263	0.108	0.421	-0.138	0.024	0.208	-0.151	0.024	0.217
TSRRh-5min.	-0.239	0.109	0.407	-0.120	0.025	0.199	-0.134	0.025	0.208
RV-30min.	0.011	1.430	1.196	0.017	0.400	0.632	0.000	0.413	0.642
TSRV-30min.	-0.202	0.875	0.957	-0.123	0.243	0.509	-0.139	0.249	0.518
RR-30min.	0.153	0.340	0.603	0.159	0.097	0.349	0.145	0.100	0.348
TSRR-30min.	-0.181	0.344	0.614	-0.100	0.091	0.318	-0.114	0.095	0.328
TSRRh-30min.	-0.181	0.343	0.613	-0.098	0.091	0.317	-0.113	0.095	0.327
RV-390min.	0.038	18.666	4.321	0.009	5.362	2.316	0.006	5.494	2.344
TSRV-390min.	0.035	18.697	4.324	0.007	5.371	2.317	0.004	5.502	2.346
RR-390min.	0.054	3.887	1.972	0.043	1.123	1.061	0.037	1.146	1.071
TSRR-390min.	-0.086	3.714	1.929	-0.056	1.061	1.032	-0.062	1.083	1.043
TSRRh-390min.	-0.097	3.662	1.916	-0.061	1.048	1.025	-0.068	1.070	1.037

Note: The table summarizes the Monte Carlo estimation results for (TS)RV, (TS)RR and TSRRh as ex-post volatility estimators. Volatility and returns are simulated from the Heston Jump-Diffusion, a fractional Brownian Motion and a discrete-time Log-Volatility Model. The bias, variance and RMSE statistics are based on 1,000 daily sample paths with 23,401 prices per day. For each of the 1,000 sample paths the calculation of the TSRRh volatility estimator is based on 500 Monte Carlo simulations that are used for estimating the probabilities of observing bid-ask noise in the intraday ranges. Non-trading and bid-ask bounce are implemented by setting the probability of observing a trade to 0.10 and using a half-spread of 0.025% on the asset price. Results are reported for sampling frequencies of 5, 30 and 390 minutes.

14

4 Empirical application

For a relatively liquid (IBM) and illiquid (Zimmer Holdings, ZMH) stock we obtain intra-day transaction prices and quotes from the TAQ database for the 1/1/2006 – 12/31/2008 period. The data are cleaned following the procedures documented in Barndorff-Nielsen et al. (2009) with the exception that we do not use moving-average rules to judge the adequacy of observed transactions.[12] Using the cleaned data we estimate the bid-ask spreads, following Roll (1984), to be 2.13 basis points (bps) for IBM and 4.93 bps for ZMH. The daily and intra-daily variation in bid-ask spreads through our sample period is, however, quite substantial. This particularly applies to the financial market turmoil in 2008. The trading probabilities are estimated to be 0.084 for ZMH and 0.201 for IBM on a 1-second time-grid.[13]

We evaluate the out-of-sample forecasting performance of the heuristically bias-adjusted RR, (TS)RV and (TS)RR estimators. For each realized measure we use an AR(1) model (with intercept) to construct one day ahead volatility forecasts, using a rolling window of one year to estimate the AR(1) coefficients. The out-of-sample period is 1/1/2007–12/31/2008. We compare volatility forecasts using the commonly used 5-minute sampling frequency. This choice is motivated by the Monte-Carlo results described in Section 3. We report Mincer-Zarnowitz and encompassing regression results to evaluate the predictive accuracy. In the Monte Carlo simulation we illustrated that for several stochastic volatility models the TSRRh is a highly efficient volatility estimator in the presence of noise and non-trading. Since for empirical data the integrated variance is unknown we compare the volatility forecasts using forecast comparison regressions rather than bias, variance and RMSE.

---

[12]Transactions and quotes are cleaned as follows: 1: Delete observations not originating from the NYSE 2: Delete all implausible data, e.g. negative quotes/prices those equal to 0, 0.01 or e.g. 999.9., observations associated with a negative spread (ask<bid) etc. 3: Delete observations with sale condition other than "E"/"F". 4: Delete observations with time stamps outside the 9:30–16:00 hours. 5: Delete all corrected observations ($corr \neq 0$) 6: When multiple transaction prices have the same time stamp use the median, do the same for bid-quotes and ask-quotes. 7: Delete transactions that traded more than a spread size outside the bid-ask spread.

[13]For IBM the number of observed transactions before data cleaning procedures is substantially larger with 29,923 observations per day. We follow the convention to limit ourselves to the 1-second time grid, as described in the footnote above, we take the median of those transactions and this dramatically reduces the resulting number of transactions that are used to estimate the volatility.

We run Mincer-Zarnowitz and encompassing regressions to evaluate the competing forecasts and following Aït-Sahalia and Mancini (2008) we use the two-time-scales realized variance TSRV as the ex-post volatility measure. Hence, the Mincer-Zarnowitz regressions are of the form

$$TSRV_t = \alpha + \beta x_{t|t-1} + \varepsilon_t, \tag{13}$$

where $x_{t|t-1}$ is the volatility forecast for day t conditional on the data available at day $t-1$. In the encompassing regressions the realizations are regressed on two competing forecasts (being, e.g., the realized range and realized volatility forecast),

$$TSRV_t = \alpha + \beta_1 x_{1,t|t-1} + \beta_2 x_{2,t|t-1} + \varepsilon_t. \tag{14}$$

For these regressions we report the coefficient estimates and their corresponding t-statistics based on Newey-West HAC robust standard errors (20 lags).

4.1 Empirical forecast results

Table 2 summarizes the Mincer-Zarnowitz regression results for volatility forecasts based on the (TS)RV, (TS)RR and TSRRh estimators. We find for both stocks that the differences in forecast accuracy are small due to the high correlation between volatility forecasts. For the relatively liquid IBM stock, we find that the realized variance forecasts have a Mincer-Zarnowitz R^2 of 49.6%. The two-time-scales realized volatility manages an R^2 of 50.8%. It slightly underperforms the unadjusted realized range forecasts which explain 50.9% of the variation in the ex-post TSRV estimates. This finding is quite remarkable, in the sense that the TSRV serves as proxy for the integrated variance in the Mincer-Zarnowitz regressions. Forecasts based on the bias-adjusted realized range proposed by Christensen et al. (2009) achieves an R^2 of 50%, hence the bias-adjusted realized range performs slightly worse compared to its unadjusted counterpart. Consistent with the volatility estimation results in the Monte Carlo simulations, the empirical forecasts based on the heuristically adjusted realized range outperform the forecasts based on other estimators as the TSRRh achieves an R^2 of 51.0%.

For the relatively illiquid stock, Zimmer Holdings (ZMH), we find that the R^2's are substantially lower than for IBM volatility forecasts. Interestingly, the advantage of a bias-correction almost vanishes. This may be due to the fact that most corrections, in contrast to TSRR(h), are derived under continuous-time assumptions that do not hold for illiquid stocks. For example, the standard realized volatility has a Mincer-Zarnowitz R^2

16

Table 2: Mincer-Zarnowitz Forecast Regressions

	Panel A: IBM 5m					Panel C: IBM 5m with outlier correction				
	RV	TSRV	RR	TSRR	TSRRh	RV	TSRV	RR	TSRR	TSRRh
α	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
tstat	1.311	1.471	1.531	1.523	1.398	1.494	1.642	1.668	1.660	1.539
β	1.231	1.222	1.010	1.196	0.914	1.164	1.147	0.946	1.123	0.858
tstat	23.801	19.853	18.282	19.448	17.546	21.980	16.745	11.509	12.612	11.182
R^2	0.496	0.508	0.509	0.500	0.510	0.685	0.687	0.686	0.681	0.689
	Panel B: ZMH 5m					Panel D: ZMH 5m with outlier correction				
α	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
tstat	-0.928	-0.879	-0.242	-1.022	-0.988	-1.076	-1.030	-0.372	-1.230	-1.150
β	1.736	1.820	1.562	2.054	1.250	1.691	1.779	1.522	2.003	1.220
tstat	9.124	7.295	10.921	9.995	8.812	10.019	7.740	12.244	11.040	9.517
R^2	0.331	0.328	0.335	0.328	0.333	0.433	0.432	0.437	0.430	0.436

Note: The table summarizes the results of Mincer-Zarnowitz forecast regressions with and without an outlier-correction applied to 10/10/2008. The (TS)RV, (TS)RR and TSRRh forecasts are generated using a AR(1) process that is dynamically re-estimated using a moving window with window length 100 days. The sampling-frequencies reported are 5-minutes and daily. The imperfect volatility proxy used is the TSRV at the 5-minute sampling frequency.

of 33.1%, being somewhat higher than that of the TSRV (32.8%). Again we expected the latter to actually have a small advantage since it is the ex-post quantity used to evaluate the forecasts. Unreported simulation results indicate that TSRV does not outperform the standard RV estimator due to the noise estimate $RV^N/2N$ being inaccurate when N is small in practice, whereas in the theory outlined by Zhang et al. (2005) it is assumed that $N \to \infty$. When N is large we can assume that the volatility signal in RV^N is dwarfed by the noise signal. It is easy to see, however, that when N is small the volatility signal in RV^N increases. For this reason it causes a downward bias due to overcorrecting for noise.[14] The (TS)RV and TSRR forecasts are outperformed by the unadjusted realized range ($R^2 = 33.5\%$) and the novel heuristic adjustment ($R^2 = 33.3\%$). The bid-ask adjustment of Christensen et al. (2009) is at par with the two-time-scales estimator (32.8%). The heuristic bias-adjustment for the realized range (33.3%) outperforms (TS)RV and TSRR. Hence, for the relatively illiquid ZMH stock we find that bias-adjustments do not pay-off

---

[14]See e.g. also Zhang et al. (2005) or Aït-Sahalia and Mancini (2008) who report a very small negative bias in TSRV in a setting where 23,401 transactions per day are observed, if we move to more realistic settings and the number of observations decreases, this negative bias becomes more pronounced. Of course, using a lower sampling frequency for TSRV could reduce the impact of non-trading.

in terms of forecasting performance, it is in this case better to just use the RR volatility estimator without applying a bias-correction and if we insist on using intraday data, then the TSRRh is preferred based on its forecast regression R^2.

Table 3 summarizes the results for encompassing forecast regressions. We find that for IBM the forecasts obtained from the TSRV estimator encompass those from the unadjusted RV estimator, as expected based on the results in Aït-Sahalia and Mancini (2008). The coefficient on TSRV (1.570) is statistically significant ($t = 3.500$) whereas the coefficient on RV is negative (-0.357) and statistically insignificant ($t = -0.846$). Similarly, the unadjusted realized range encompasses the unadjusted realized variance with coefficients being $1.909(1.692)$ and $-0.100(-0.134)$, respectively. Adding RR or TSRV forecasts to unadjusted RV forecasts results in the same R^2 of 50.9%. When we add the forecasts based on the TSRR estimator to RV forecasts we find that the R^2 shrinks to 50.1% and both coefficients are statistically insignificant. However, adding the forecasts based on the heuristic bias-adjustment for realized range (TSRRh) to unadjusted RV forecasts actually improves the R^2 to 51.0% with its coefficient being 0.992 (1.709) and the coefficient on RV being $-0.107(-0.145)$. In addition we report encompassing regression results for all other (bi-variate) forecast combinations and find that adding the unadjusted RR forecasts to the TSRV forecasts results in similar and statistically insignificant coefficients being $0.562(0.733)$ and $0.546(0.599)$, respectively, and an R^2 of 51.0%. Hence, combining RV and TSRRh forecasts results in the same R^2 as combining TSRV and RR. When we add the TSRR forecasts to TSRV forecasts we again find statistically insignificant coefficients being $1.633(1.053)$ on TSRV and $-0.408(-0.269)$ on TSRR. In contrast, we find that TSRV $6.143(4.891)$ and TSRRh $-6.139(-3.857)$ compete, having statistically significant coefficients of similar absolute size but opposite signs, due to a high correlation between the forecasts. Running an encompassing regression for TSRR and TSRRh forecasts results in both forecasts being statistically significant and opposite signs with TSRR having a coefficient of $-3.219(-2.332)$ and $3.346(3.334)$ for TSRRh. Looking at the 10 possible forecast combinations the optimal combination found for the IBM data is that of TSRV and TSRRh forecasts with an R^2 of 52.9%.

A similar analysis for the relatively illiquid stock (ZMH) illustrates that in contrast to the IBM results now RV $1.652(1.561)$ forecasts outperform TSRV $0.090(0.084)$ forecasts. The RV forecasts $0.254(0.327)$, however, are outperformed by the RR $1.336(2.321)$ fore-

casts, as expected. However, the RV forecasts 1.719(2.007) almost reduce the coefficient on TSRR 0.021(0.024) to zero. Hence, whereas the bias-adjustments worked for the relatively liquid IBM data this is not the case for the illiquid ZMH data. We find similar results when we add the TSRV or TSRR to RR forecasts, that is, the unadjusted RR forecasts are better than the TSRV and TSRR. If one uses high-frequency data combined with a bias correction, then the TSRRh is preferred over TSRR and TSRV.

Table 3: Empirical Encompassing Forecast Regressions

Panel A: IBM 5m

	RV+TSRV	RV+RR	RV+TSRR	RV+TSRRh	TSRV+RR	TSRV+TSRR	TSRV+TSRRh	RR+TSRR	RR+TSRRh	TSRR+TSRRh
α	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
tstat	1.538	1.909	1.802	1.683	1.725	1.624	1.542	1.547	1.281	1.207
β_1	-0.357	-0.100	0.349	-0.107	0.546	1.633	0.486	6.143	-0.588	-3.219
tstat	-0.846	-0.134	0.368	-0.145	0.599	1.053	0.569	4.891	-0.260	-2.332
β_2	1.570	1.090	0.861	0.992	0.562	-0.408	0.554	-6.139	1.446	3.346
tstat	3.500	1.692	0.914	1.709	0.733	-0.269	0.842	-3.857	0.700	3.334
R^2	0.509	0.509	0.501	0.510	0.510	0.509	0.511	0.529	0.510	0.521

Panel B: ZMH 5m

	RV+TSRV	RV+RR	RV+TSRR	RV+TSRRh	TSRV+RR	TSRV+TSRR	TSRV+TSRRh	RR+TSRR	RR+TSRRh	TSRR+TSRRh
α	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
tstat	-0.932	-0.347	-0.947	-1.009	-0.092	-1.034	-0.939	1.787	-0.496	-0.851
β_1	1.652	0.254	1.719	0.522	-0.259	0.910	-0.056	5.556	1.078	-0.634
tstat	1.561	0.327	2.007	0.622	-0.282	1.193	-0.073	7.458	1.059	-0.420
β_2	0.090	1.336	0.021	0.878	1.780	1.035	1.288	-5.319	0.391	1.630
tstat	0.084	2.321	0.024	1.524	2.424	1.351	2.358	-5.442	0.448	1.684
R^2	0.331	0.335	0.331	0.334	0.335	0.329	0.333	0.342	0.335	0.334

Panel C: IBM 5m with outlier correction

	RV+TSRV	RV+RR	RV+TSRR	RV+TSRRh	TSRV+RR	TSRV+TSRR	TSRV+TSRRh	RR+TSRR	RR+TSRRh	TSRR+TSRRh
α	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
tstat	1.709	2.067	1.954	1.861	1.878	1.868	1.675	1.661	1.434	1.358
I(20081010)	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
tstat	54.862	32.956	39.821	36.919	74.747	100.894	82.796	53.829	47.287	54.283
β_1	0.387	0.500	0.908	0.371	0.652	1.449	0.453	4.370	-2.669	-2.780
tstat	0.653	0.555	0.827	0.441	0.837	1.163	0.695	4.416	-0.992	-1.826
β_2	0.769	0.545	0.250	0.588	0.412	-0.300	0.522	-4.093	3.270	2.959
tstat	1.240	0.699	0.226	0.881	0.597	-0.238	0.973	-3.275	1.359	2.711
R^2	0.688	0.688	0.685	0.690	0.688	0.687	0.690	0.695	0.691	0.697

Panel D: ZMH 5m with outlier correction

	RV+TSRV	RV+RR	RV+TSRR	RV+TSRRh	TSRV+RR	TSRV+TSRR	TSRV+TSRRh	RR+TSRR	RR+TSRRh	TSRR+TSRRh
α	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
tstat	-1.110	-0.426	-1.139	-1.156	-0.357	-1.186	-1.121	1.734	-0.713	-0.969
I(20081010)	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
tstat	72.364	83.896	72.779	76.846	111.750	81.674	75.200	101.525	100.563	79.768
β_1	1.160	0.177	1.514	0.245	0.059	1.147	0.085	5.165	0.815	-0.815
tstat	1.572	0.256	2.133	0.333	0.075	1.741	0.124	6.646	0.796	-0.541
β_2	0.563	1.365	0.213	1.046	1.472	0.719	1.162	-4.850	0.570	1.709
tstat	0.711	2.632	0.303	1.954	2.495	1.202	2.431	-4.676	0.643	1.756
R^2	0.434	0.437	0.433	0.437	0.437	0.432	0.436	0.443	0.437	0.437

Note: The table summarizes encompassing forecast regression results with and without an outlier correction applied to 10/10/2008. The (TS)RV, (TS)RR and TSRRh forecasts are generated using a AR(1) process that is dynamically re-estimated using a moving window with window length 100 days. The sampling-frequencies reported are 5-minutes and daily. The imperfect volatility proxy used is the TSRV at the 5-minute sampling frequency.

4.2 Outlier correction

During our out-of-sample period, which contains the height of the recent financial crisis and the beginning of its aftermath, several trading days exhibited extremely high volatility and can be regarded as outliers. It is interesting to analyze how an outlier correction would influence the results. There is a vast literature on how to adjust for outliers, such as truncating values that are more than several standard deviations away from the (local) average of the volatility process or incorporating dummy variables etc. Because there are several ways to go and we do not want to alter the empirical data too much we will only incorporate a dummy for 10/10/2008 which was found to be an outlier using several approaches and analyze how this alters the results discussed above.[15]

The Newey-West t-statistic for the dummy variable on 10/10/2008 is larger than 70 for all estimators when using Mincer-Zarnowitz regressions and for the IBM data the t-stat is 145 for the dummy when using RV forecasts. Note the huge increase in the Mincer-Zarnowitz R^2's for IBM and ZMH by explicitly incorporating this outlier. For the IBM data the average R^2 shifts 18.1% in absolute terms and 35.8% in relative terms and for ZMH the shifts are 10.3% and 31.0%, respectively.

For the IBM data the conclusions do not change in the sense that if we rank the forecasts on the Mincer-Zarnowitz regression R^2 the TSRRh ($R^2 = 68.9\%$) forecasts still slightly outperform (TS)RV and (TS)RR forecasts. Similarly, for the ZMH data we again find that the unadjusted RR has the largest R^2 being 43.7% and if we insist on using a bias-adjusted estimator the TSRRh achieves the best result with $R^2 = 43.6\%$.

For the encompassing regressions we find that for the IBM forecasts the TSRRh are not rendered obsolete by the other forecasts. The forecast combination that has the highest R^2 is that of TSRR and TSRRh. The ZMH results illustrate that in the encompassing regressions with outlier correction the TSRRh performs satisfactorily as it outperforms the (TS)RV and TSRR and it competes with the unadjusted realized range. Hence, including an outlier dummy for the most severe outlier in our sample does not alter the main conclusions.

---

[15]For example, the RV on 10/10/2008 is more than 8 standard deviations away from the unconditional average.

5 Conclusion

We have proposed a novel heuristic bias-correction for realized range-based volatility estimates. For the heuristic adjustment we use three inputs that are easily and accurately estimated from high-frequency data. The needed inputs are estimates of the following quantities: (i) the daily range that is unaffected by noise, (ii) the non-trading probability and (iii) the half-spread. Using these inputs we simulate a geometric Brownian motion with variance (i) and implement noise with settings (ii) and (iii). For the simulated Brownian motions we keep count of how many intraday ranges are upward biased (most likely), unbiased or downward biased (least likely). By averaging over simulation runs we estimate probabilities for the three cases that can be attached to the ranks of sorted intraday ranges. We apply these probability ranks to the sorted vector of initial high-low ranges for which we are now able to indicate whether an intraday range is expected to be upward biased, unbiased or downward biased.

Using three stochastic volatility models for the integrated volatility, which can include jumps, leverage effects and dependence in the increments of a Brownian motion, we find that in the presence of bid-ask bounce and non-trading, volatility estimates based on the new heuristically bias-adjusted realized range estimator (TSRRh) are more efficient than estimates based on the realized variance, realized range and their two time scales adjusted counterparts.

In an empirical setting we evaluated out-of-sample volatility forecasts using Mincer-Zarnowitz and encompassing forecast regressions. For the relatively liquid IBM stock we find that the heuristically bias-adjusted realized range estimator (TSRRh) compares favorably to forecasts based on the (TS)RV and (TS)RR estimators. For the relatively illiquid Zimmer Holdings stock (ZMH), we find that TSRRh improves upon (TS)RV and TSRR forecasts and is on par with the RR estimator.

References

Aït-Sahalia, Y. and Mancini, L.: 2008, Out of sample forecasts of quadratic variation, *Journal of Econometrics* **147**, 17–33.

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

Andersen, T., Bollerslev, T., Diebold, F. and Labys, P.: 2003, Modeling and forecasting realized volatility, *Econometrica* **71**, 529–626.

Andersen, T., Bollerslev, T. and Meddahi, N.: 2011, Realized volatility forecasting and market microstructure noise, *Journal of Econometrics* **160**(1), 220–234.

Barndorff-Nielsen, O., Hansen, P., Lunde, A. and Shephard, N.: 2009, Realised kernels in practice: trades and quotes, *Econometrics Journal* **12**, 1–32.

Christensen, K. and Podolskij, M.: 2007, Realized range-based estimation of integrated variance, *Journal of Econometrics* **141**, 323–349.

Christensen, K., Podolskij, M. and Vetter, M.: 2009, Bias-correcting the realized range-based variance in the presence of market microstructure noise, *Finance and Stochastics* **13**, 239–268.

Davies, R. and Harte, D.: 1987, Tests for hurst effect, *Biometrika* **74**, 95–102.

Engle, R.: 2000, The econometrics of ultra high frequency data, *Econometrica* **68**, 1–22.

Ghysels, E. and Sinko, A.: 2011, Volatility forecasting and microstructure noise, *Journal of Econometrics* **160**(1), 257–271.

Griffin, J. and Oomen, R.: 2008, Sampling returns for realized variance calculations: tick time or transaction time?, *Econometric Reviews* **27**, 230–253.

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

Heston, S.: 1993, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* **6**(2), 327–343.

Liu, L., Patton, A. and Sheppard, K.: 2012, Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. Duke University, Working Paper.

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

McAleer, M. and Medeiros, M.: 2008, Realized volatility: a review, *Econometric Reviews* **27**, 10–45.

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

Roll, R.: 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* **39**(4), 1127–1139.

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