The University of Sydney Business School
The University of Sydney

BUSINESS ANALYTICS WORKING PAPER SERIES

Generalized Variance: A Robust Estimator of Stock Price Volatility

M. Sutton, A. Vasnev and R. Gerlach
The University of Sydney

April 2015

Abstract

This paper proposes an ex-post volatility estimator, called generalized variance, that uses high frequency data to provide measurements robust to the idiosyncratic noise of stock markets caused by market microstructures. The new volatility estimator is analyzed theoretically, examined in a simulation study and evaluated empirically against the two currently dominant measures of daily volatility: realized volatility and realized range. The main finding is that generalized variance is robust to the presence of microstructures while delivering accuracy superior to realized volatility and realized range in several circumstances. The empirical study features Australian stocks from the ASX 20.

Key words: Volatility; Robust estimator.

BA Working Paper No: BAWP-2015-02
http://sydney.edu.au/business/business_analytics/research/working_papers

Generalized Variance: A Robust Estimator of Stock Price Volatility

M. Sutton, A. Vasnev* and R. Gerlach
The University of Sydney

April 2015

Abstract

This paper proposes an ex-post volatility estimator, called generalized variance, that uses high frequency data to provide measurements robust to the idiosyncratic noise of stock markets caused by market microstructures. The new volatility estimator is analyzed theoretically, examined in a simulation study and evaluated empirically against the two currently dominant measures of daily volatility: realized volatility and realized range. The main finding is that generalized variance is robust to the presence of microstructures while delivering accuracy superior to realized volatility and realized range in several circumstances. The empirical study features Australian stocks from the ASX 20.

Key words: Volatility; Robust estimator.

1 Introduction

Volatility measures variation in asset price movements over time and is an essential component in option pricing, portfolio optimization and financial risk management practices. The availability of high frequency, intraday price observations has enabled the construction of ex-post estimators of daily volatility, e.g., realized volatility (RV), formalized in Andersen et al. (2001b), and realized range (RR), introduced concurrently by Martens and van Dijk (2007) and Christensen and Podolskij (2007).

Andersen et al. (2001b) suggested employing returns over smaller and smaller time intervals, whereby RV would accurately measure the true volatility, a result noted for daily returns by Merton (1980). Subsequently, the properties of RV under ideal conditions (continuous and frictionless prices) have been well-studied: Andersen et al. (2003) and

*Corresponding e-mail address: andrey.vasnev@sydney.edu.au

Barndorff-Nielsen and Shephard (2002) showed that as the data frequency increases, RV converges to the true volatility; the asymptotic distribution of RV was shown to be multivariate Gaussian by Bandi and Russell (2008).

However, as the frequency of observation increases, so-called market microstructures cause prices and returns to deviate from the assumed stochastic processes. These may include the bid-ask bounce and non-synchronous trading. The effects of microstructures include bias and inefficiency in the estimation of volatility; see McAleer and Medeiros (2008) for a review of their impact on realized volatility. There have been many attempts to adjust RV to minimize market microstructure effects: Andersen et al. (2001b), Aït-Sahalia et al. (2005) and Patton (2011) found the optimal frequency for sampling observations was between 15 seconds and 5 minutes; Aït-Sahalia et al. (2005) found that modelling market microstructures improved RV estimation; Zhang et al. (2005) constructed a two-time scales RV (TTSRV) estimator using sub-sampling of observations, dramatically reducing bias in RV; Zhou (1996), Hansen and Lunde (2006), Barndorff-Nielsen et al. (2008) and Barndorff-Nielsen et al. (2009) used kernels to account for the time dependency structure imposed on the time series by market noise; Martens and van Dijk (2007) suggested scaling realized volatility by the close-close estimator across a past sample of daily prices; finally, Barndorff-Nielsen and Shephard (2004) introduced bi-power variation, which takes the absolute value of the multiplication of sequential pairs of returns.

The realized range (RR) is proposed by Martens and van Dijk (2007) and Christensen and Podolskij (2007), in which the highest and lowest prices observed within subintervals of a day are employed. Christensen and Podolskij (2007) examined the theoretical properties of the RR estimator under ideal conditions, first showing consistency and then indicating that the asymptotic distribution of RR is multivariate Gaussian and subsequently that the variance of RR, under ideal conditions, is approximately five times smaller than the variance of RV. Comparing the explanatory power of each estimator, Martens and van Dijk (2007) found that RR was superior to RV when applied to the S&P 500 futures index as well as individual stocks in the S&P 100. Papavassiliou (2012) provided confirmation of the usefulness of RR for actively traded Greek market stocks.

Via simulation, Martens and van Dijk (2007) examined the accuracy of RV and RR when the price process was affected by the bid-ask bounce and non-synchronous trading, finding that RR was more accurate at higher data sampling frequencies. Martens and van Dijk (2007) found that scaling realized range by a weight based on the Parkinson (1980) daily range estimator over the past 66 trading days improved the accuracy of realized range in the presence of market microstructures. Todorova (2012) found that this scaling improved the accuracy of realized range measures for stocks with low liquidity.

While realized volatility and realized range have advanced the field of ex-post volatility estimation, they still have serious shortcomings when addressing market microstructures. This paper proposes an estimator based on squared returns; however, instead of using returns over sequential intervals, the estimator calculates returns over expanding intervals. If the initial price used in the return is fixed as the starting price of the day, the effect of a temporary price shock may only be counted once. RV and RR would count the effect of a shock twice: once for the initial price jump and once for the reversal as depicted in Figure 1. The advantage of this approach is that the effect of some market microstructures on the estimation of volatility may be diminished. As a result, this simple modification to RV, named generalized variance (GV), is robust to noise created by market microstructures.

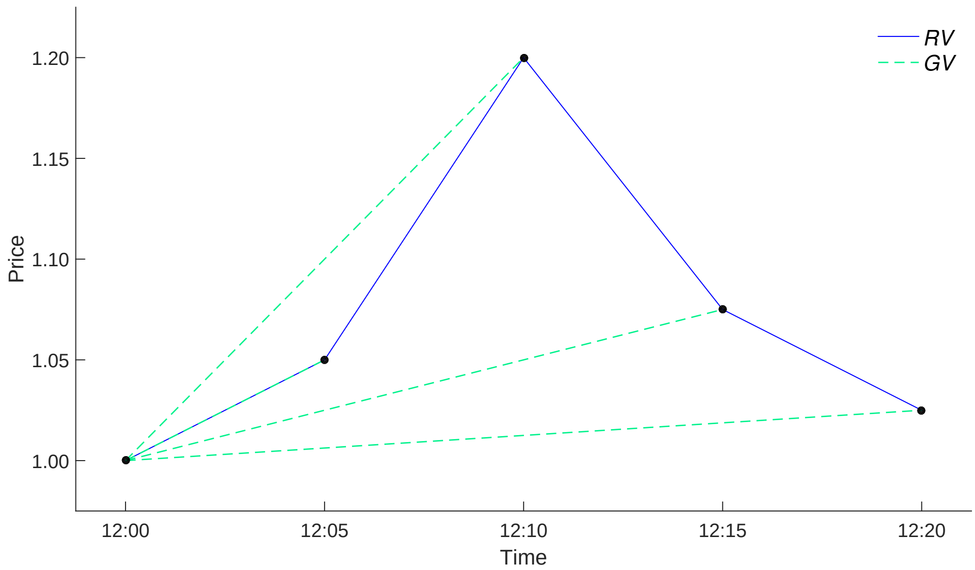


Figure 1: Stylised picture of returns used in realized volatility and generalized variance computations. Dots represent the observed stock prices. For each line, the starting point indicates the initial price, and the end point represents the end price in the calculation of a return. The dark blue solid lines represent the returns used to calculate realized volatility, while the light green dashed lines represent those for generalized variance. The jump at 12:10 causes two large returns for realized volatility, while only one large return is observed for generalized variance.

The paper is structured in four sections. Section 2 examines the theoretical properties of each estimator in ideal conditions and under the bid-ask bounce market microstructure. Section 3 details a simulation study conducted to test the results found in Section 2, also considering non-synchronous trading. The empirical study is presented in Section 4, describing the methodology applied, as well as the distributional properties and a comparison of realized volatility, realized range and generalized variance. Finally, Section 5 summarizes the core conclusions reached in this paper.

2 Theoretical Analysis

2.1 Price Process

The notation used in this paper is similar to Bandi and Russell (2008). This paper considers D days, each segmented into M equal intervals. The length of a day is normalized to 1, so that the length of each interval is $\delta = 1/M$. Continuous trading is assumed, such that the opening price on the i-th day is equal to the closing price of the $(i-1)$-th day, $P_{(i-1)}$. The price at the end of the j-th period on the i-th day is denoted as $P_{(i-1)+j\delta}$. $H_{(i-1)+j\delta}$ will represent the highest price during the j-th period on the i-th day; $H_{(i-1)+j\delta} = \max\limits_{(j-1)\leq t \leq j} P_{(i-1)+t\delta}$. The lowest price during the day is similarly defined as $L_{(i-1)+j\delta} = \min\limits_{(j-1)\leq t \leq j} P_{(i-1)+t\delta}$.

This paper considers a Brownian Motion process for logarithmic prices, in which the drift is zero and the volatility process is constant:

$$\log(P_t) - \log(P_0) = \int_0^t \sigma \, dW_s \, . \tag{1}$$

The object of interest for this paper is the variance of the price process, σ^2, where σ is from Eq. (1). The focus of this paper will be on the effect of market frictions.

To facilitate discussion of market microstructures, the observed price will be denoted as $\widetilde{P}_t$, with observed highest and lowest prices: $\widetilde{H}_t$ and $\widetilde{L}_t$. In the ideal case, the observed price is the true equilibrium price, $\widetilde{P}_t = P_t$. When there are market frictions, the observed price deviates from the true equilibrium price; in this case, $\widetilde{P}_t$ is not necessarily equal to P_t. When there are market frictions, the realized volatility, realized range and generalized variance estimators will be denoted by $\widetilde{RV}_i$, $\widetilde{RR}_i$, $\widetilde{GV}_i$, respectively.

2.2 Estimators

Realized volatility is defined as the sum of intraday squared returns. It is characterized by the following equation

$$\widehat{RV}_i = \sum_{j=1}^{M} \left[\log\left(\frac{\widetilde{P}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}} \right) \right]^2 \tag{2}$$

Realized range is defined as the sum of the normalized squared intraday ranges. The proceeding representation follows from Christensen and Podolskij (2007) and Martens and van Dijk (2007), who focus on an intraday version of the range estimator introduced by Parkinson (1980).

$$\widehat{RR}_i = \frac{1}{4\log 2} \sum_{j=1}^{M} \left[\log\left(\frac{\widetilde{H}_{(i-1)+j\delta}}{\widetilde{L}_{(i-1)+j\delta}} \right) \right]^2 \tag{3}$$

This paper introduces and examines a new estimator that considers intraday squared returns, calculated over an expanding time period. While realized volatility considers the squared return from period 1 to period 2 and then period 2 to period 3, this paper suggests considering the return from period 1 to period 3 or period 4 or period 5, etc. A general form for this estimator could be written as

$$\widehat{V}_{g,i}^2 = \sum_{j=1}^{M} \frac{\left[\log\left(\frac{\widetilde{P}_{(i-1)+(b_j)\delta}}{\widetilde{P}_{(i-1)+(a_j)\delta}}\right)\right]^2}{f(a_j, b_j)},$$

where $a_j, b_j \in \{0, 1, 2, ..., M\}$, $a_j < b_j$ and $f(a_j, b_j)$ is a function to correct for taking different sized intervals. The problem with this estimator is when prices are observed every second, there might be as many as 86,400 one-second intervals per day. Considering all possible sizes of intervals, there would be over 3.7 billion combinations for each day. This reduces the tractability of the problem.

To reduce the number of combinations, this paper proposes to fix $a = 0$, such that each "return" is the return on investment of buying an asset at the start of day i and selling it at the end of the j-th period. This paper will call this estimator generalized variance. It can be represented by

$$\widehat{GV}_i = \frac{2}{M+1} \sum_{j=1}^{M} \left[\log\left(\frac{\widetilde{P}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)}}\right)\right]^2 \tag{4}$$

The factor of $\frac{2}{M+1}$ is needed to account for differently sized intervals. As the first price is fixed as the opening price, this estimator is sensitive to changes in the opening price, introducing a source of error; changes in the opening price will affect all terms in the summation in Eq. (4). To reduce this impact, this paper suggests that a second estimate can be formed by fixing $b = M$ and allowing a to vary, a mirror of the generalized variance estimator. Then, an average of this estimate and $\widehat{GV}_i$ can be taken to reduce this sensitivity, forming the corrected generalized variance estimator,

$$\widehat{CGV}_i = \frac{2}{M+1} \sum_{j=1}^{M} \frac{1}{2} \left\{ \left[\log\left(\frac{\widetilde{P}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)}}\right)\right]^2 + \frac{1}{2} \left[\log\left(\frac{\widetilde{P}_{(i-1)+M\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right)\right]^2 \right\} \tag{5}$$

Due to their symmetry, generalized variance and its mirror have the same expectation. The combination $\widehat{CGV}_i$, however, has a lower variance due to the advantages from averaging estimates. It is noteworthy that $\widehat{V}_{g,i}^2$ could be constructed by combining generalized variance estimators from Eq. (4), in which the starting price is incremented forward by one observation each iteration and the previous starting price is excluded.

2.2.1 Additional Estimators

The following estimators are not examined theoretically; however, they are considered in the simulation study. The first is a modification to realized volatility, the realized kernel.

The form used by Barndorff-Nielsen et al. (2009) is

$$\widehat{RK}_i = \sum_{K=-\eta}^{\eta} k\left(\frac{K}{\eta+1}\right)\gamma_K\,,$$

$$\gamma_K = \sum_{j=|K|+1}^{M} \log\left(\frac{\widetilde{P}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right)\log\left(\frac{\widetilde{P}_{(i-1)+(j-K)\delta}}{\widetilde{P}_{(i-1)+(j-K-1)\delta}}\right)\,,$$

$$k(x) = \begin{cases} 1-6x^2+6x^3 & 0 \leq x \leq 1/2 \\ 2\left(1-x\right)^3 & 1/2 < x \leq 1 \\ 0 & x > 1 \end{cases}$$

where $k(x)$ is the Parzel kernel function and η is the bandwidth. The bandwidth for the Parzel kernel is chosen by $\eta = 3.5134\,\xi^{4/5}M^{3/5}$. ξ is estimated using the methods described in Barndorff-Nielsen et al. (2009).

The second additional estimator is the Garman and Klass (1980) form for range,

$$\widehat{RR}_{gk,a} = \frac{1}{2}\sum_{j=1}^{M}\left\{0.511\left[\log\left(\frac{\widetilde{H}_{(i-1)+j\delta}}{\widetilde{L}_{(i-1)+j\delta}}\right)\right]^2 - 0.019\log\left(\frac{\widetilde{P}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right)\log\left(\frac{\widetilde{H}_{(i-1)+j\delta}\widetilde{L}_{(i-1)+j\delta}}{\widetilde{P}^2_{(i-1)+(j-1)\delta}}\right) - 2\log\left(\frac{\widetilde{H}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right)\log\left(\frac{\widetilde{L}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right)\right\}\,.$$

Garman and Klass (1980) also introduced a more practical estimator, which has been adapted to be a realized measure in the following way

$$\widehat{RR}_{gk,b} = \frac{1}{2}\sum_{j=1}^{M}\left[\log\left(\frac{\widetilde{H}_{(i-1)+j\delta}}{\widetilde{L}_{(i-1)+j\delta}}\right)\right]^2 + \left(2\log(2)-1\right)\sum_{j=1}^{M}\left[\log\left(\frac{\widetilde{P}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right)\right]^2\,.$$

Finally, Rogers and Satchell (1991) introduced a range-based estimator that used the highest, lowest, opening and closing prices and was independent of drift. The form employed by this paper was constructed as

$$\widehat{RR}_{rs} = \sum_{j=1}^{M}\left[\log\left(\frac{\widetilde{H}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+j\delta}}\right)\log\left(\frac{\widetilde{H}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right) + \log\left(\frac{\widetilde{L}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+j\delta}}\right)\log\left(\frac{\widetilde{L}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)+(j-1)\delta}}\right)\right]\,.$$

2.3 Properties of the estimators under ideal conditions

Without market microstructures, the properties of the considered estimators are given by the following theorem.

Theorem 1. *Assume that p_t follows a Brownian Motion process and $\widetilde{p}_t = p_t$. Then,*

(a) the expectation, variance and asymptotic distribution of realized volatility are

$$\begin{aligned}\mathbb{E}\left[\widehat{RV}_i\right] &= \sigma^2\,,\\ \operatorname{Var}\left(\widehat{RV}_i\right) &= \frac{2\sigma^4}{M}\,,\\ \sqrt{M}\left(\widehat{RV}_i - \sigma^2\right) &\xrightarrow[n\to\infty]{d} \mathcal{N}\left(0,\, 2\sigma^4\right)\,.\end{aligned}$$

(b) the expectation, variance and asymptotic distribution of realized range are

$$\begin{aligned}\mathbb{E}\left[\widehat{RR}_i\right] &= \sigma^2\,,\\ \operatorname{Var}\left(\widehat{RR}_i\right) &= \frac{\lambda\sigma^4}{M}\,,\\ \sqrt{M}\left(\widehat{RR}_i - \sigma^2\right) &\xrightarrow[n\to\infty]{d} \mathcal{N}\left(0,\, \lambda\,\sigma^4\right)\,.\end{aligned}$$

(c) the expectation and variance of generalized variance are

$$\begin{aligned}\mathbb{E}\left[\widehat{GV}_i\right] &= \sigma^2\,,\\ \operatorname{Var}\left(\widehat{GV}_i\right) &= \frac{4\,(M^2+M+1)}{3M\,(M+1)}\sigma^4\,,\\ \lim_{M\to\infty}\operatorname{Var}\left(\widehat{GV}_i\right) &= \frac{4\sigma^4}{3}\,.\end{aligned}$$

(d) the expectation and variance of corrected generalized variance are

$$\begin{aligned}\mathbb{E}\left[\widehat{CGV}_i\right] &= \sigma^2\,,\\ \operatorname{Var}\left(\widehat{CGV}_i\right) &= \frac{2\,(M^2+M+1)\,\sigma^4}{3M\,(M+1)} + \frac{(M+2)\,\sigma^4}{3M}\,,\\ \lim_{M\to\infty}\operatorname{Var}\left(\widehat{CGV}_i\right) &= \sigma^4\,,\end{aligned}$$

where $\lambda = \frac{9\,\zeta(3)-\left(4\log(2)\right)^2}{\left(4\log(2)\right)^2} \approx 0.4$.

See proofs in Appendix A.2.

The realized volatility estimator is unbiased in the ideal case, as observed in Theorem 1. In addition, the limit of the unconditional variance for RV, as frequency M increases to infinity, is zero. Consequently, in ideal conditions, RV is consistent. The asymptotic distribution of the realized volatility estimator is Gaussian (McAleer and Medeiros, 2008). The expectation and variance found in Theorem 1 are compatible with this result from McAleer and Medeiros (2008).

Theorem 1 shows that the realized range estimator is also unbiased, when there are no market microstructures. As in the case of RV, under these circumstances, RR is consistent. The unconditional variance of RR, however, is approximately five times smaller than the unconditional variance of RV; consequently, it is five times more efficient. This result was found by Parkinson (1980) for daily data and was rederived for realized range by Martens and van Dijk (2007) and Christensen and Podolskij (2007). Christensen and Podolskij (2007) also found that the asymptotic distribution of RR was Gaussian.

Under ideal conditions, Theorem 1 shows that the generalized variance estimator is unbiased. The unconditional variance has a curious form, quite different from those observed for RV and RR; this is a consequence of using expanding returns. As observed in Theorem 1, the limit as M approaches infinity of the unconditional variance of GV is $\frac{4\sigma^4}{3}$. This convergence is fast, however, as when $M = 144$ (the number of 10-minute intervals within a 24-hour trading day), $\frac{M^2+M+1}{M(M+1)} = 1.00005$. A repercussion of this non-zero lower bound on variance is that the generalized variance estimator is not as efficient as realized volatility or realized range. In addition, GV is convergent in distribution, rather than consistent, and hence, the asymptotic distributions of the estimators are not examined in this paper. As a result of these efficiency and consistency results, GV will not achieve the same accuracy as RV or RR in ideal conditions.

The results for corrected generalized variance under ideal conditions are largely similar to those for generalized variance in Theorem 1. The correction maintains the unbiased property, and the unconditional variance does not vanish as the frequency goes to infinity. The asymptotic unconditional variance for CGV of σ^4, however, is 25% smaller than what was found for GV. Consequently, under ideal conditions, corrected generalized variance will be more efficient than generalized variance.

The accuracy expected in the simulation study can be deduced by considering mean square error (MSE). MSE is a measure of accuracy that captures the bias and variance of an estimator, $MSE = Bias^2 + Variance$. In the ideal case, we would theoretically expect realized range to perform the best due to the efficiency of RR, closely followed by RV. The performance of both of these estimators should improve with the frequency of observations. GV and CGV are not expected to produce the same accuracy as the two other estimators in the ideal case and are not expected to improve as the frequency of observation increases.

2.4 Properties of the estimators under the Bid-Ask Bounce

The bid-ask bounce occurs when the price bounces between the bid price and the ask price and is described by Tsay (2010). In the ideal case, the assumption was that the observed price was always equal to the true equilibrium price. To examine the effect of the bid-ask bounce on the performance of the estimators, this paper assumed that the observed price could be observed as the true equilibrium price, $\widetilde{P}_t = P_t$, the bid price, $\widetilde{P}_t = (1-s)\,P_t$, or the ask price, $\widetilde{P}_t = (1+s)\,P_t$, where s is half the percentage spread. The sample space of observed prices has been explored in Table 1.

$\widetilde{P}_{i+a\delta}$	$\widetilde{P}_{i+b\delta}$	Squared returns
mid	mid	$\log\left(\frac{P_{i+b\delta}}{P_{i+a\delta}}\right)^2$
mid	bid	$\log\left(\frac{(1-s)P_{i+b\delta}}{P_{i+a\delta}}\right)^2$
mid	ask	$\log\left(\frac{(1+s)P_{i+b\delta}}{P_{i+a\delta}}\right)^2$
bid	mid	$\log\left(\frac{P_{i+b\delta}}{(1-s)P_{i+a\delta}}\right)^2$
bid	bid	$\log\left(\frac{(1-s)P_{i+b\delta}}{(1-s)P_{i+a\delta}}\right)^2 = \log\left(\frac{P_{i+b\delta}}{P_{i+a\delta}}\right)^2$
bid	ask	$\log\left(\frac{(1+s)P_{i+b\delta}}{(1-s)P_{i+a\delta}}\right)^2$
ask	mid	$\log\left(\frac{P_{i+b\delta}}{(1+s)P_{i+a\delta}}\right)^2$
ask	bid	$\log\left(\frac{(1-s)P_{i+b\delta}}{(1+s)P_{i+a\delta}}\right)^2$
ask	ask	$\log\left(\frac{(1+s)P_{i+b\delta}}{(1+s)P_{i+a\delta}}\right)^2 = \log\left(\frac{P_{i+b\delta}}{P_{i+a\delta}}\right)^2$

Table 1: Sample space of two observed prices, $\widetilde{P}_{i+a\delta}$ and $\widetilde{P}_{i+b\delta}$, with the associated "squared returns" when the observed price can be the true equilibrium price (mid), the bid price (bid) or the ask price (ask).

The case can be made simpler for realized range. Consider that for realized range, when prices are observed continuously, if the highest price during an interval is observed as the bid price or mid price, a price in an arbitrarily small neighborhood around the time of observation will be observed as an ask price. Thus, the highest price observed during this interval can be approximated by $(1+s)\,H_{(i-1)+j\delta}$, and the lowest observed price can be similarly approximated by $(1-s)\,L_{(i-1)+j\delta}$. This spread artificially inflates the range, causing an overestimate of the true realized range and consequently a biased estimate (Vortelinos, 2014). As a result, it is expected that in this scenario, RR will be biased.

The properties of the estimators under the bid-ask bounce are given by the following theorem.

Theorem 2. *Assume that* $\log(P_t)$ *follows a Brownian Motion process and* P_t *can be observed as the mid price, the bid price or the ask price with equal probability. Then,*

(a) the expectation and variance of the realized volatility estimator are

$$\mathbb{E}\left[\widetilde{RV}_i\right] = \sigma^2 + \frac{4Mb_1}{9},$$

$$\mathrm{Var}\left(\widetilde{RV}_i\right) = \frac{2\sigma^4}{M} + \frac{16b_1\sigma^2}{9} + \frac{4\,(6M-1)\,b_1^2}{81},$$

(b) the expectation and variance of the realized range estimator are

$$\mathbb{E}\left[\widetilde{RR}_i\right] = \sigma^2 + \sqrt{\frac{2M}{\pi}}\frac{b_2}{\log 2}\sigma + \frac{Mb_2^2}{4\log 2},$$

$$\mathrm{Var}\left(\widetilde{RR}_i\right) = \frac{\lambda\sigma^4}{M} + \frac{8\sqrt{2}b_2}{\sqrt{M}}\left(\frac{\pi^{3/2}}{3} - \frac{4\log 2}{\sqrt{\pi}}\right)\sigma^3 + 4b_2^2\left(4\log 2 - \frac{8}{\pi}\right)\sigma^2,$$

(c) the expectation and variance of the generalized variance estimator are

$$\mathbb{E}\left[\widetilde{GV}_i\right] = \sigma^2 + \frac{8Mb_1}{9(M+1)},$$

$$\mathrm{Var}\left(\widetilde{GV}_i\right) = \frac{4(M^2+M+1)}{3M(M+1)}\sigma^4 + \frac{32b_1}{9(M+1)}\sigma^2 + \frac{80Mb_1^2}{81(M+1)^2},$$

$$\lim_{M\to\infty}\mathrm{Var}\left(\widetilde{GV}_i\right) = \frac{4\sigma^4}{3},$$

$$= \lim_{M\to\infty}\mathrm{Var}\left(\widehat{GV}_i\right),$$

(d) the expectation and variance of the corrected generalized variance estimator are

$$\mathbb{E}\left[\widetilde{CGV}_i\right] = \sigma^2 + \frac{8Mb_1}{9(M+1)},$$

$$\mathrm{Var}\left(\widetilde{CGV}_i\right) = \left[\frac{2(M^2+M+1)}{3M(M+1)} + \frac{M+2}{3M}\right]\sigma^4 + \frac{16b_1\sigma^2}{9(M+1)} + \frac{4(11M-1)b_1^2}{81(M+1)^2},$$

$$\lim_{M\to\infty}\mathrm{Var}\left(\widetilde{CGV}_i\right) = \sigma^4,$$

$$= \lim_{M\to\infty}\mathrm{Var}\left(\widehat{CGV}_i\right),$$

where $\lambda = \frac{9\zeta(3) - (4\log(2))^2}{(4\log(2))^2} \approx 0.4$*,* $b_1 = \log(1+s)^2 + \log(1-s)^2 - \log(1+s)\log(1-s) > 0$ *and* $b_2 = \log(1+s) - \log(1-s) > 0$.

See proofs in Appendix A.2.

When the price is affected by the bid-ask bounce, the realized volatility estimator becomes biased. The factor $\frac{4b_1}{9}$ is positive and the bias grows linearly with the frequency of observations. A similar result was found by Bandi and Russell (2008); in the presence of market microstructures, $\widetilde{RV}_i \to \infty$ as $M \to \infty$. As realized volatility is biased in these circumstances and the bias grows with frequency, the realized volatility estimator is no longer consistent. Simultaneously, the variance of realized volatility is also affected by the bid-ask bounce; as frequency increases to infinity, the variance increases to infinity. This means that at higher frequencies, the realized volatility estimator will perform quite poorly. It is noteworthy that these results reduce to the ideal case when $s = 0$ and thus $b_1 = 0$.

The bid-ask bounce induces a bias in RR as a consequence of the spread artificially inflating the range, causing overestimation of the true realized range (Vortelinos, 2014). As prices are not observed continuously in practice, during any given interval, the observed highest (lowest) price may be lower (higher) than the highest (lowest) true price. The effect of this non-synchronous trading can bias range-based estimators downward (Beckers, 1983;

Martens and van Dijk, 2007; Vipul, 2008; Vortelinos, 2014). As in the case of realized volatility, the realized range estimator exhibits positive bias that grows linearly with frequency; consequently, realized range is inconsistent. The coefficients of both σ^3 and σ^2 are positive in the unconditional variance of realized range, so the bid-ask bounce also increases the variance of the realized range estimator. In contrast to realized volatility, however, as $M \to \infty$, the variance achieves a lower bound of

$$\lim_{M\to\infty} \operatorname{Var}\left(\widetilde{RR}_i\right) = 4b_2^2 \left(4\log(2) - \frac{8}{\pi}\right)\sigma^2 .$$

Theorem 2 shows that generalized variance is also biased when prices are affected by the bid-ask bounce. In contrast to realized volatility and realized range, however, this positive bias has an upper bound of $\frac{8b_1}{9}$. The variance of the generalized variance estimator is also inflated due to the bid-ask bounce. This inflation, however, decreases with frequency, such that as $M \to \infty$, the variance returns to its value in ideal conditions. Consequently, while generalized variance is not consistent, its behavior is more desirable than realized volatility and realized range when there are market microstructure effects. This shows that the generalized variance estimator is more robust to the effects of the bid-ask bounce.

Due to its construction, CGV has the same expectation as GV, resulting in a biased estimator under the bid-ask bounce. Similar to ideal conditions, however, corrected generalized variance has a lower unconditional variance. This is most apparent as the frequency of observation approaches infinity, where the unconditional variance again is 25% smaller than that found for generalized variance and achieves the same value as its variance under ideal conditions. As a result, CGV should provide better performance than generalized variance while retaining its robustness to market microstructures. Consequently, the rest of this paper will focus on corrected generalized variance.

Considering the accuracy of the estimators, all the estimators are biased when prices oscillate between the bid price, the mid price and the ask price. The bias of corrected generalized variance is bounded, while the bias of both realized range and realized volatility increase indefinitely with frequency. Comparisons of the variance of the estimators are moot when two of the three competitors are so immensely affected by bias. Corrected generalized variance would therefore be theoretically expected to exhibit the best performance when this market microstructure is present, especially at higher frequencies of observation.

While more general forms of market frictions have been explored in the literature, the theoretical effect of the bid-ask bounce has not been examined specifically. Different market microstructures have different impacts on the price processes and thus the behavior of estimators. This analysis of the bid-ask bounce provides information about the performance of estimators when applied to empirical prices for which processes are dominated by the bid-ask bounce. These theoretical results can be directly tested using a simulation study.

3 Simulation Study

3.1 Simulation Study Design

This simulation extended the study conducted by Martens and van Dijk (2007). Prices were simulated 24 hours a day for 1000 days, and the starting price was set to \$1, which was not reset each day. Estimates of the daily volatility were calculated at several frequencies: 1 second, 5 seconds, 10 seconds, 20 seconds, 1 minute, 5 minutes and 10 minutes.

In their simulation study, Martens and van Dijk (2007) examined Geometric Brownian Motion (GBM), setting the total annualized standard deviation of the GBM process, $\sigma_A = 0.21$. Thus, the total daily standard deviation is $\sigma_D = \sigma_A/\sqrt{252}$, as it was assumed that there were 252 trading days in each year. The prices were simulated using

$$dP_t = \mu P_t \, dt + \sigma_{gbm} P_t \, dW_t \,,$$

where μ is the drift parameter, σ_{gbm} is the diffusion parameter and W_t is a Wiener process. As prices were simulated 100 times per second, 86,400 seconds per day, the volatility was scaled: $\sigma_{gbm} = \sigma_D/N$, where $N = 8,640,000$ and μ was set to 0, following Martens and van Dijk (2007).

This paper adapted Martens and van Dijk (2007) by considering additional processes to simulate prices: the Merton Jump Diffusion model, GARCH(1,1) model and Heston Stochastic Volatility model. The results for these three remaining processes were very similar to those found for Geometric Brownian Motion, and so they are presented in Appendix B. This provides evidence that microstructures impact the behavior of volatility estimators more than allowing volatility to vary. The discussion presented here will focus on Geometric Brownian Motion.

3.2 Market Microstructures

Market microstructures are frictions due to the structure of a market and trades that occur within that market that cause prices to deviate from the ideal case. This paper followed Martens and van Dijk (2007) in considering two such frictions within the simulation study, the bid-ask bounce and non-synchronous trading.

As described in Section 2.4, the bid-ask bounce occurs when the price is observed jumping between the bid price, the mid price and the ask price. Martens and van Dijk (2007) simulated the bid-ask bounce by setting the spread to a fixed dollar value and assumed that prices were observed at either the bid or the ask price, with equal probability of occurrence. This study modified Martens and van Dijk (2007) by also including the mid price; the price process was simulated such that the price was equally likely to be observed as the bid, ask or mid price. In addition, the spread used was a fixed percentage, 0.05%.

Aitken and Frino (1996) found that the percentage spread of a stock was related to its price, volatility and level of trading activity. The following relationship was identified by regression

$$\log(S) = -3.075 - 0.177\log(P) - 0.576\log(V) + 0.720\log(\sigma) \ ,$$

where S is the percentage bid-ask spread, P is the stock price, V is the level of trading activity, as measured by the average number of trades per half hour, and σ is the annualized stock price volatility. In this paper, the trading volume V was set to 180; this is equivalent to a trade every 10 seconds. This relationship produced percentage spreads varying by approximately 0.1%. This second form of the bid-ask bounce will be referred to as the A&F bid-ask bounce.

A transaction is a discrete event in time; consequently, prices are observed at discrete times. In addition, assets often have periods of high activity, during which trades occur with high frequency, and periods of low activity, during which the volume of trades is much lower. The result is that the price observed at a certain time only reflects the last trade and can deviate from the assumed continuous price. This market friction is referred to as non-synchronous trading.

Following Martens and van Dijk (2007), non-synchronous trading was modelled by randomly selecting the time of each trade. The simulation was constructed so that on average, the price was only observed every 10 seconds. This is representative of the empirical data; for ASX 20 stocks between 2010 and 2013, the average time between trades was 9.25 seconds. This non-synchronous behavior was simulated by constructing a Poisson random variable, which averaged 8,640 arrivals or price observations per day. In the case in which two arrivals occurred at the same time (a rare occurrence), it was considered to be only one arrival.

3.3 Performance Evaluation

For Geometric Brownian Motion and the Merton Jump Diffusion process, the true daily volatility is known. For GARCH(1,1) and the Heston model, the true daily volatility was formed by aggregating, via summation, the true volatilities at each time increment. As there are approximately 8.6 million observations, this should be sufficiently accurate.

The measure of fit used to evaluate the estimates was mean absolute percentage error,

$$MAPE = \frac{1}{D}\sum_{i=1}^{D}\frac{|\sigma_i^2 - \hat{\sigma}_i^2|}{\sigma_i^2} \ ,$$

where D is the number of days, which in this case is 1000. This has a few advantages over other measures. Due to the already small value of variances, squaring the error would produce an extremely small value, which is avoided by this method. In addition, it corrects for the scale of the variance, which is especially useful for the stochastic volatility models.

3.4 Results

This section assesses the accuracy of each estimator for prices following Geometric Brownian Motion under the market microstructure conditions described in Section 3.2. In this section, generalized variance refers to corrected generalized variance, as only corrected generalized variance is considered.

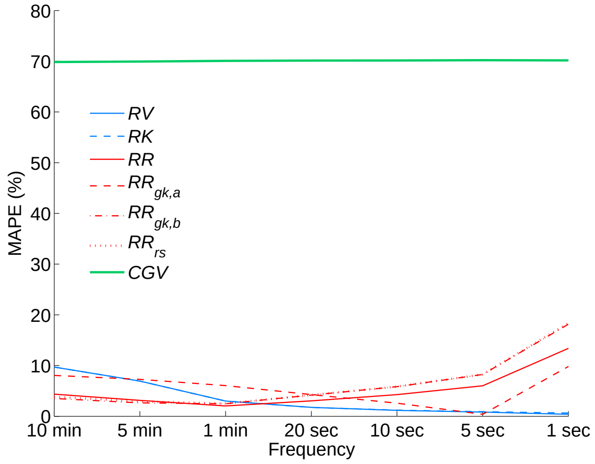
The mean squared error of an unbiased estimator is equal to the estimator's variance. As MAPE is also a measure of accuracy, we would expect to see a similar relationship between the variance of an estimator and the estimator's MAPE. The results for the Geometric Brownian Motion process are presented in Figure 2 and Table 2. The results for the three remaining processes are presented in Appendix B: the Merton Jump Diffusion process in Figure 4 and Table 5, the GARCH(1,1) process in Figure 5 and Table 6, and the Heston Stochastic Volatility process in Figure 6 and Table 7.

The accuracy of realized volatility clearly increases with the frequency in the ideal case. It was shown in Theorem 1 that the variance of the realized volatility estimator is $2\sigma^4/M$. As the frequency increases, the variance approaches zero, reflecting a decrease in uncertainty and thus a more accurate estimate. Table 2 shows that the MAPE for realized volatility decreases from 9.73% at 10 minutes to 0.38% at 1 second. It is important to note that in ideal conditions, the realized kernel estimator is indistinguishable from realized volatility; it was introduced in the literature as a means of improving realized volatility when the data exhibited market microstructures.

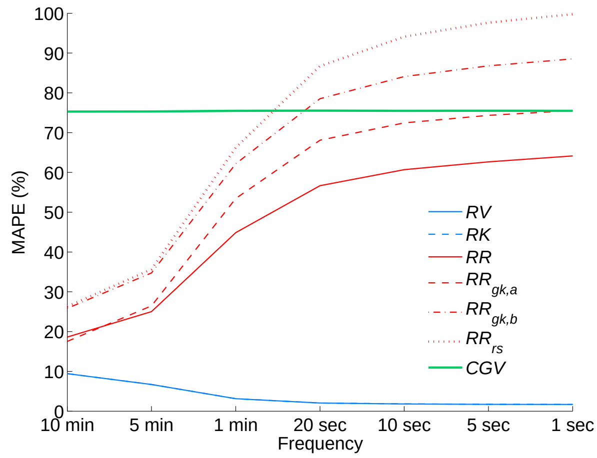
For frequencies ranging from 10 minutes to 1 minute, the Parkinson realized range estimator displays the behavior that is expected given its properties from Theorem 1 in Section 2.3; it is more efficient and thus more accurate than realized volatility, and its error decreases with frequency. Beyond 1 minute, this realized range estimator departs from expectation; the accuracy diminishes as frequency increases. Table 2 shows that the MAPE of realized range improves from 4.38% at 10 minutes to its minimum of 2.03% at 1 minute and deteriorates to 13.38% when observations are recorded each second. The highest frequency considered by Martens and van Dijk (2007) was 1 minute, and thus, their analysis does not indicate this behavior. The other realized range estimators perform similarly, with initial improvement in accuracy with frequency and subsequent decline.

Generalized variance does not achieve the same accuracy as the benchmarks under ideal conditions. Table 2 lists the MAPE of generalized volatility as approximately 70%. Figure 2(a) emphasizes the lack of relationship between the accuracy of generalized volatility and frequency. From Theorem 1 in Section 2.3, we know that

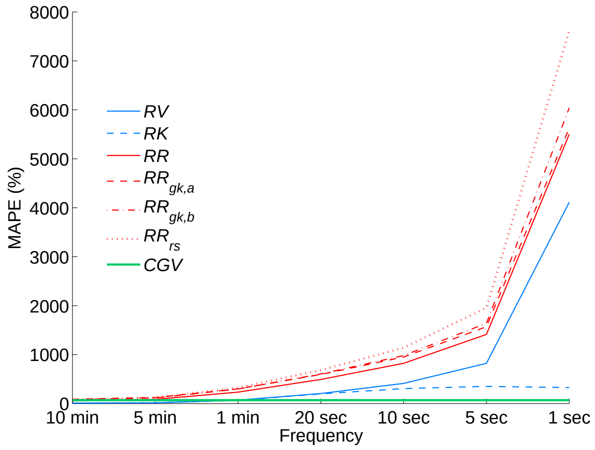
$$\begin{aligned}\operatorname{Var}\left(\widehat{CGV}_i\right) &= \frac{2\left(M^2+M+1\right)\sigma^4}{3M\left(M+1\right)} + \frac{\left(M+2\right)\sigma^4}{3M}\,,\\ \lim_{M\to\infty}\operatorname{Var}\left(\widehat{CGV}_i\right) &= \sigma^4\,.\end{aligned}$$



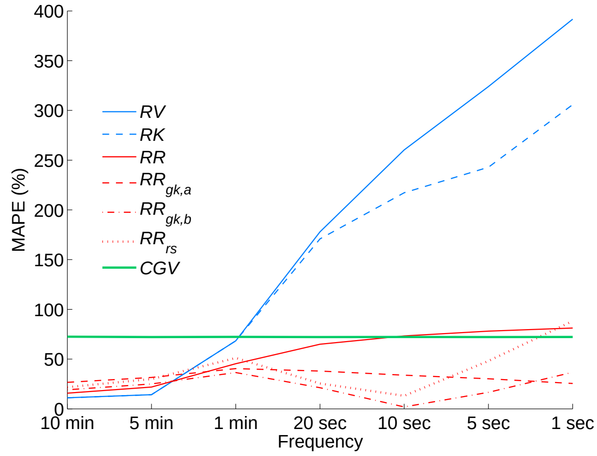
(a) No Market Microstructures



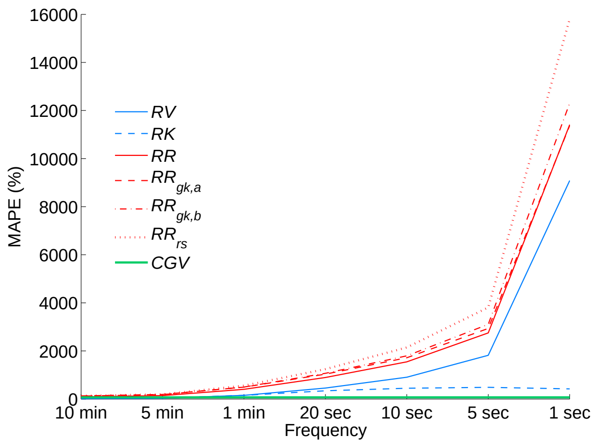
(b) Non-Synchronous Trading



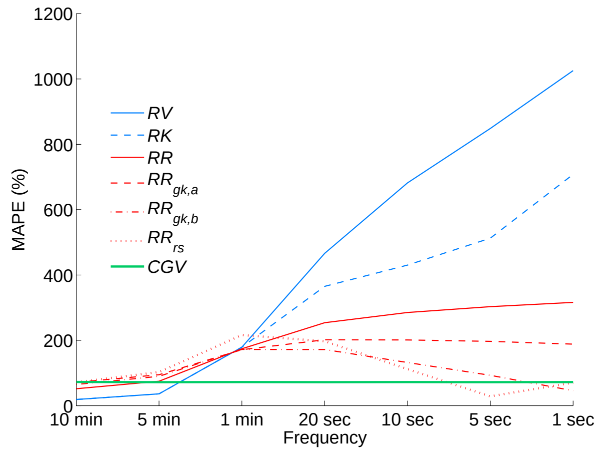
(c) Bid-Ask Bounce



(d) Bid-Ask Bounce and Non-Synchronous Trading



(e) A&F Bid-Ask Bounce



(f) A&F Bid-Ask Bounce and Non-Synchronous Trading

Figure 2: Accuracy in simulation under Geometric Brownian Motion.
The performance of the generalized volatility estimator is not significantly affected by market microstructures, additionally appearing independent of frequency. This suggests that generalized volatility is a robust estimator of the true volatility; in the presence of noise, it is often the best estimator at high frequencies.

	(a) No Market Microstructures							(b) Non-Synchronous Trading						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	**0.38**	0.63	13.38	9.82	18.16	18.34	70.20	**1.68**	1.69	64.15	75.54	88.54	99.77	75.42
5 sec	0.81	0.89	6.02	**0.39**	8.24	8.27	70.24	**1.72**	1.73	62.65	74.33	86.78	97.64	75.43
10 sec	**1.17**	1.20	4.27	2.58	5.84	5.84	70.20	**1.81**	1.83	60.69	72.45	84.09	94.10	75.41
20 sec	**1.75**	**1.75**	3.04	4.27	4.18	4.19	70.18	**2.05**	**2.05**	56.66	68.11	78.52	86.71	75.44
1 min	2.99	2.99	**2.03**	6.06	2.46	2.51	70.11	**3.14**	**3.14**	44.86	53.46	62.19	66.24	75.41
5 min	6.93	6.93	3.12	7.31	**2.64**	2.92	70.00	**6.70**	**6.70**	25.02	26.48	34.74	35.68	75.25
10 min	9.73	9.73	4.38	8.08	**3.50**	3.82	69.91	**9.43**	**9.43**	18.64	17.55	25.86	26.33	75.24

	(c) Bid-Ask Bounce							(d) Bid-Ask Bounce and Non-Synchronous Trading						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	4114.41	326.58	5504.59	5631.74	6041.75	7622.68	**68.88**	391.69	305.73	81.29	**25.49**	36.79	88.23	72.33
5 sec	822.82	351.14	1412.40	1569.27	1640.08	1957.20	**68.92**	323.98	242.77	78.14	30.21	**16.56**	48.08	72.21
10 sec	411.24	304.92	822.45	951.81	981.38	1140.23	**68.86**	260.32	217.23	73.28	33.80	**1.90**	13.15	72.27
20 sec	205.98	197.22	493.42	595.32	604.51	683.85	**68.92**	177.99	170.84	64.93	37.99	**21.29**	25.51	72.25
1 min	69.08	69.08	232.94	300.10	296.31	322.65	**68.94**	68.35	68.35	45.39	40.49	**36.52**	51.14	72.44
5 min	**14.25**	**14.25**	86.88	123.18	115.05	120.36	68.78	**14.20**	**14.20**	21.90	31.53	25.04	29.99	72.23
10 min	**10.98**	**10.98**	58.64	87.17	78.57	81.29	68.67	**11.14**	**11.14**	15.77	26.60	19.21	21.85	72.60

	(e) A&F Bid-Ask Bounce							(f) A&F Bid-Ask Bounce and Non-Synchronous Trading						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	9070.09	418.60	11389.60	11357.20	12285.49	15776.24	**68.85**	1023.20	706.33	315.05	187.81	**45.65**	70.57	72.28
5 sec	1813.78	486.40	2743.47	2930.13	3102.34	3801.82	**68.91**	846.28	511.37	302.17	196.19	92.57	**28.21**	72.08
10 sec	906.59	446.82	1545.49	1712.82	1792.39	2142.65	**68.84**	680.06	429.30	284.39	200.62	131.73	111.41	**72.19**
20 sec	453.85	343.65	895.57	1032.75	1066.25	1241.23	**68.93**	465.06	364.36	253.18	201.01	171.37	196.54	**72.15**
1 min	151.85	151.85	400.84	493.39	497.09	555.22	**68.93**	178.59	178.59	173.67	170.75	171.78	215.68	**72.43**
5 min	**30.39**	**30.39**	140.27	189.56	182.67	194.38	68.74	**35.97**	**35.97**	74.76	94.47	89.75	102.78	72.10
10 min	**16.43**	**16.43**	92.76	130.76	122.70	128.66	68.68	**18.92**	**18.92**	51.62	71.24	64.69	71.56	72.56

Table 2: Accuracy in Simulation under Geometric Brownian Motion Process. The estimators are labelled as in Section 2.2. This table gives the MAPE (%) values observed in Figure 2. The most accurate estimator for each frequency is in bold.

When $M = 144$ and the number of observations per day is at the 10-minute frequency, $\mathrm{Var}\left(\widehat{CGV}_i\right) = 1.005 \times \sigma^4$. As M increases, this factor becomes closer and closer to unity; however, even for the lowest frequency observed, it is already very close to 1. Consequently, the improvement in the variance from a frequency of 10 minutes to 1 second is not significant. The higher MAPE of the generalized variance estimator is due to the size of its variance relative to that of realized volatility and realized range.

Realized volatility is largely unaffected by non-synchronous trading (Vortelinos, 2014). This is confirmed in Figure 2(b); the shape of the realized volatility accuracy is identical to that found in the ideal case. Table 2, however, shows that at the higher frequencies, the accuracy of realized volatility has decreased slightly, from 0.81% to 1.72% at the 5-second frequency. Similar to the ideal case, as realized volatility is not affected, the realized kernel estimator's performance is very similar to that observed for realized volatility.

As discussed in Section 1, non-synchronous trading induces a downwards bias on range-based estimators; evidence for this is seen in Figure 2(b). If an estimator is biased downward for a true parameter that must be greater than or equal to zero, the largest percentage error the estimator can produce is 100%, which occurs when the estimate is 0. The shape of the realized range estimators is indicative of lines approaching a limit asymptotically. Rogers and Satchell's estimator best displays this, with MAPE in Table 2 commencing at 26.33% when prices are observed every 10 minutes; however, at the 1-second frequency, the MAPE is 99.77% as the estimator approaches zero.

The presence of non-synchronous trading does have an impact on the performance of generalized variance, as observed in Table 2, with MAPE increasing to 75.44% from 70.18% in ideal conditions when prices are observed every 20 seconds. The behavior of the accuracy of generalized variance, observed in Figure 2(b), is the same as in ideal conditions. While the effects of market microstructures on prices increase with frequency, these frictions do not diminish the performance of generalized variance.

Figure 2(c) shows the performance of the estimators when the price is affected by a bid-ask bounce with a 0.05% spread. Theorem 2 states that in this case, realized volatility is biased and its variance grows with frequency. Table 2 shows that when observing prices every 10 minutes, realized volatility is actually the best estimator; however, at the 1-minute frequency, the effects of the bid-ask bounce increase the MAPE of realized volatility above that of generalized variance to 69.08%. As opposed to the previous circumstances, the realized kernel estimator alleviates most of the impact of the bid-ask bounce at higher frequencies; it is the second best estimator for frequencies from 20 seconds to 1 second. By taking into account neighboring prices, the kernel estimator appears to be able to alleviate most of the effects of the bid-ask bounce.

In the presence of the bid-ask-bounce, realized range-based estimators perform the worst. The relationship between frequency and accuracy can be explained by Theorem 2; the realized range based on Parkinson is biased under the bid-ask bounce, and the bias is proportional to frequency. This simulation result shows that this also holds for the other forms of realized range considered.

The scale of Figure 2(c) obscures the performance of the generalized variance estimator. This is because its performance is again robust to the effects of a market microstructure; in this case, the bid-ask bounce. Table 2 shows that the MAPE of generalized volatility is approximately 68.8% at all frequencies. This results in its best-in-class performance for prices affected by this bid-ask bounce at frequencies above 1 minute.

When the previous form of the bid-ask bounce is combined with non-synchronous trading, realized volatility is the equal best estimator with the realized kernel estimator, with a MAPE of 11.14% and 14.20% at the 10- and 5-minute frequencies in Table 2. However, the combination of these two market microstructures produces too much noise at higher frequencies, with realized volatility recording the worst performance of the examined estimators when prices are observed more frequently than every 20 seconds. The accuracy of the realized kernel estimator departs from that of realized variance for frequencies above 20 seconds. Nevertheless, the realized kernel estimator still exhibits growth of error with frequency, resulting in only a mild improvement, with a MAPE of 217.23% compared to 260.32% for realized volatility at a frequency of 10 seconds. This result is surprising, as the realized kernel estimator is one of the best estimators under the bid-ask bounce and also under non-synchronous trading; under the combination, however, it is the second worst.

In contrast, while realized range-based estimators are inaccurate under the bid-ask bounce and under non-synchronous trading due to bias, under the combined market frictions, the upward bias and downward bias appear to cancel out. This results in a superior performance for realized range-based estimators for frequencies greater than 1 minute, as observed in Figure 2(d). It is interesting to note that there is no single best performing realized range estimator. For example, at the 20-second frequency, the Garman and Klass practical range-based estimator is the best, with a MAPE of 21.29% from Table 2; however, at the 1-second frequency, the original version of Garman and Klass's range performs best with a MAPE of 25.49%. The consequence of this is that it becomes difficult to select the form of range to employ in this situation, as the best estimator differs between frequencies.

The performance of generalized variance is largely unchanged even under combined market noises, recording a mean absolute percentage error of approximately 72.2% for all frequencies. The superior performance under the bid-ask bounce, however, does not completely carry over to this scenario; generalized variance outperforms realized volatility and realized kernel for frequencies above 1 minute; however, the realized range estimators are generally better. It is only at the 1-second frequency that the Parkinson-based and Rogers and Satchell-based estimators perform worse than generalized variance.

The A&F bid-ask bounce considers the relationship between the bid-ask spread and the price and volatility of the asset as well as the average trading volume. A key difference is also that the spread is approximately twice as high, with values that are often around 0.1%. This is clearly noticeable as the scale of the vertical axis on Figure 2(e) is twice that of Figure 2(c).

The examination of the estimators under the A&F bid-ask bounce in Figure 2(e) is identical to that conducted for the behavior observed in Figure 2(c). The impact of this different bid-ask bounce is found in Figure 2(f). The combination of the A&F bid-ask bounce and non-synchronous trading shows similar behavior for both realized volatility and the realized kernel estimator; the realized volatility estimator's accuracy deteriorates as frequency increases, and after the 1-minute frequency, the realized kernel estimator is able to reduce some of the effects of market microstructures. For both of these estimators, however, the accuracy has decreased overall when compared to the original form for the bid-ask bounce; the MAPE has increased from 69.08% to 151.85% when prices are observed every minute (from Table 2).

A similar outcome is observed for realized range, in which the shape of the curves in Figure 2(f) is very similar to that found in Figure 2(d), but accuracy has decreased overall. At a frequency of 1 minute, the MAPE has increased from 45.39% to 173.67% for Parkinson's form (Table 2). It is noteworthy that the Rogers and Satchell form for realized range is one of the best estimators in this circumstance, given that its accuracy is one of the worst for the range estimators in Figure 2(d). This again affirms the notion that while realized range estimators generally have superior performance, the choice of range form depends on the price series analyzed.

In Figure 2(f), the robustness of generalized variance to these forms of market noise results in the best estimation of true volatility for prices observed every 1 minute, 20 seconds and 10 seconds, as observed in Table 2. In addition, the fact that there is no optimal choice for the form of the range in the class of realized range estimators suggests that generalized variance should be the preferred estimator, as its performance is robust and provides optimal performance without requiring consideration of the properties of the price series examined.

It is important to note, however, that the most accurate estimate in four of the six examples is produced by realized volatility when prices are observed every 10 minutes. This questions the need for increasing the frequency and employing generalized variance. Andersen et al. (2001a) motivate the use of realized volatility because in the ideal case, as frequency increases, it will converge to the true volatility. Andersen et al. (2001a), however, are forced to suggest halting this increase in frequency at five minutes because of the effect of noise on the realized volatility estimator. Another argument in favor of generalized variance is that while the focus is currently on the estimation of daily volatility, the benefits of high frequency data should allow volatility to be measured over even shorter time scales. This has been obstructed previously because of the effect of market frictions on realized volatility and realized range. Generalized variance, however, provides robust estimates even in these circumstances and can enable estimation of volatility over shorter time intervals than a day.

4 Empirical Study

4.1 Data

The ASX 20 is the set of 20 stocks on the Australian Stock Exchange (ASX) with the highest market capitalization. The stocks that were listed on the ASX 20 as of 13/05/14 were examined and are presented in Table 8. The data were sourced from SIRCA's Australian Equities database. Three of the stocks did not have the required data, and thus, the analysis was conducted on the 17 remaining stocks.

The date range selected for this empirical report was 01/01/10 to 31/12/13. This date range was chosen because there are approximately 1000 working or trading days within this period, reflecting the sample size used in the simulation study. This paper's analysis focuses on the "normal" trading hours[1] to ensure consistency in the level of trading for observations of each day. An additional complication is that the ASX spreads the start time of the stocks on its exchange. Consequently, this analysis uses data during the period 10:10am - 4:00pm for each trading day.

There were entire days within the dataset for which there were no prices recorded. These days were removed after importation; however, there were no missing observations for days that had observations. No further data cleaning was performed.

4.2 Empirical Methodology

The methodology used for this study follows Martens and van Dijk (2007). First, the unconditional distribution of the estimators was examined as well as the distribution of daily returns when scaled by each of the estimators. Subsequently, the explanatory power of each estimator as a proxy for the true volatility was explored to determine the best estimator of volatility. This paper adapted this methodology by employing the Realized GARCH model from Hansen et al. (2012). The Realized GARCH model was specifically constructed to take advantage of realized volatility and realized range and allows for extensive inquiry into the applicability of each estimator when used to fit price series.

This empirical study focuses on the realized volatility estimator, the realized range estimator based on Parkinson (1980) and generalized variance. Other estimators were not considered due to the limitation that only two estimators could be compared at a time in this methodology. Inclusion of more estimators would dramatically increase the number of comparisons required and was not tractable. The results for the 1-second, 10-second,

[1]ASX definition of "normal" trading hours is available at www.asx.com.au/about/trading-hours.htm

1-minute and 5-minute frequencies are presented in Appendix C. Distributional results for all estimators are available upon request, as are results for the 5-second, 20-second and 10-minute frequencies.

The Diebold and Mariano (1995) test compares the ability of two estimators, $\widehat{V}_{\eta,i}$ and $\widehat{V}_{\xi,i}$, in forecasting a proxy, $\widehat{V}_{proxy,i}$, by examining the test statistic, DM. Due to the large number of companies, frequencies, proxies and competing estimators, the number of "wins" an estimate records was used to judge the performance of the estimators over the entire dataset, instead of individually for each stock. For the R^2, MSE and MAD measures, a "win" was defined as follows:

$$\begin{cases} \widehat{V}_{\xi,i-1} \text{ wins}, & \text{if } \Phi(DM) < \alpha/2\,, \\ \text{neither wins}, & \text{if } \alpha/2 \leq \Phi(DM) \leq 1-\alpha/2\,, \\ \widehat{V}_{\eta,i-1} \text{ wins}, & \text{if } \Phi(DM) > 1-\alpha/2\,, \end{cases}$$

where $\Phi(x)$ is the cumulative distribution function for a standard normal distribution and α is the significance level, chosen as 0.05. For the encompassing regressions test, the loss differentials were defined as

$$\begin{aligned} d_{\xi,i} &= \left(\widehat{V}_{\eta,i-1} - \widehat{V}_{\xi,i-1}\right)\left(\widehat{V}_{proxy,i} - \widehat{V}_{\xi,i-1}\right)\,, \\ d_{\eta,i} &= \left(\widehat{V}_{\xi,i-1} - \widehat{V}_{\eta,i-1}\right)\left(\widehat{V}_{proxy,i} - \widehat{V}_{\eta,i-1}\right)\,. \end{aligned}$$

The DM statistics, DM_ξ and DM_η, were calculated for their respective loss differentials, $d_{\xi,i}$ and $d_{\eta,i}$. A "win" was defined by

$$\begin{cases} \widehat{V}_{\xi,i-1} \text{ wins}, & \text{if } \Phi(DM_\xi) < \alpha/2, & \Phi(DM_\eta) > \alpha/2\,, \\ \widehat{V}_{\xi,i-1} \text{ wins}, & \text{if } \Phi(DM_\xi) > 1-\alpha/2, & \Phi(DM_\eta) < 1-\alpha/2\,, \\ \widehat{V}_{\eta,i-1} \text{ wins}, & \text{if } \Phi(DM_\eta) < \alpha/2, & \Phi(DM_\xi) > \alpha/2\,, \\ \widehat{V}_{\eta,i-1} \text{ wins}, & \text{if } \Phi(DM_\eta) > 1-\alpha/2, & \Phi(DM_\xi) < 1-\alpha/2\,, \\ \text{neither wins}, & \text{otherwise}\,. \end{cases}$$

The Martens and van Dijk (2007) methodology considered exponentially smoothing volatility estimates. This paper updates this by employing a more recent model introduced by Hansen et al. (2012), the Realized GARCH model. The following log-linear specification of the Realized GARCH model was employed.

$$\begin{aligned} r_t &= \sigma_t z_t\,, & z_t &\sim \text{i.i.d.}\,(0,\,1)\,, \\ \log\sigma_t^2 &= \omega + \beta\log\sigma_{t-1}^2 + \gamma\log V_{t-1}\,, \\ \log V_t &= \xi + \phi\log\sigma_t^2 + \tau_1 z_t + \tau_2\left(z_t^2 - 1\right) + u_t\,, & u_t &\sim \text{i.i.d.}\,\mathcal{N}\big(0,\,\sigma_u^2\big)\,, \end{aligned} \tag{6}$$

where σ_t^2 is the conditional variance of returns and V_{t-1} is one of the measures of volatility. The parameters $\theta = \{\omega, \beta, \gamma, \xi, \phi, \tau_1, \tau_2, \sigma_u^2\}$ were estimated using maximum likelihood. The log-likelihood function for this specification is reported by Hansen et al. (2012) as

$$l(r,V) = -\frac{1}{2}\sum_{i=1}^{d}\left[\log(2\pi) + \log\big(\sigma_i^2\big) + \frac{r_i^2}{\sigma_i^2}\right] - \frac{1}{2}\sum_{i=1}^{d}\left[\log(2\pi) + \log\big(\sigma_u^2\big) + \frac{u_i^2}{\sigma_u^2}\right]\,,$$

where d is the number of days over which the model is fitted. The first 10% of the data was used as a training sample to fit the Realized GARCH model, while the remaining observations were used as the evaluation sample. The Realized GARCH model readily produces a one-step-ahead forecast using the GARCH equation, Eq. (6). The data were then updated with the next observation and fit again to produce the proxy value. The Martens and van Dijk (2007) methodology for forecast comparison was repeated for the series of conditional variances estimated. The Realized GARCH model also allows for the comparison of predictive likelihoods for each estimator. The predictive likelihood assesses how accurate the predictive density of returns is for each model; the most accurate forecast of σ_t^2 will give the highest log predictive likelihood. For the presented formulation, the cumulative log predictive likelihood for the evaluation sample is given by

$$\sum_{i=\lfloor \lambda D \rfloor}^{D} \log p(r_i | r_1, \ldots r_{i-1}; \theta) = -\frac{1}{2} \sum_{i=\lfloor \lambda D \rfloor}^{D} \left[\log\left(2\pi\sigma_i^2\right) + \frac{r_i^2}{\sigma_i^2} \right]$$

where D is the total number of days in the sample, $\lambda = 0.1$ to reflect the size of the training sample and $\lfloor \cdot \rfloor$ is the integer part of its argument.

4.3 Results

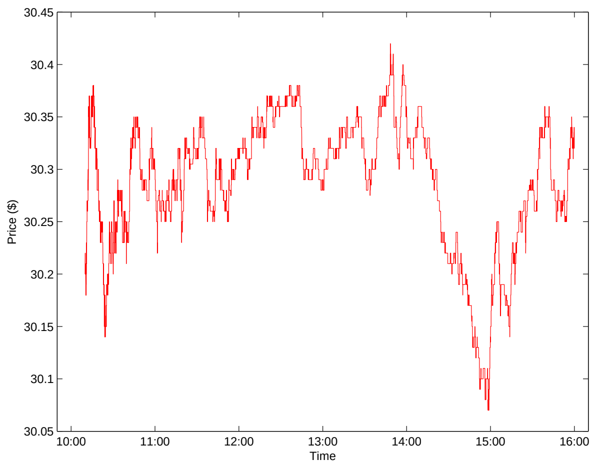
4.3.1 Unconditional Distributions of Estimators

The distributions of the annualized percentage volatility estimates are presented in Tables 9, 10 and 11 in Appendix C. The most interesting observations are discussed below.

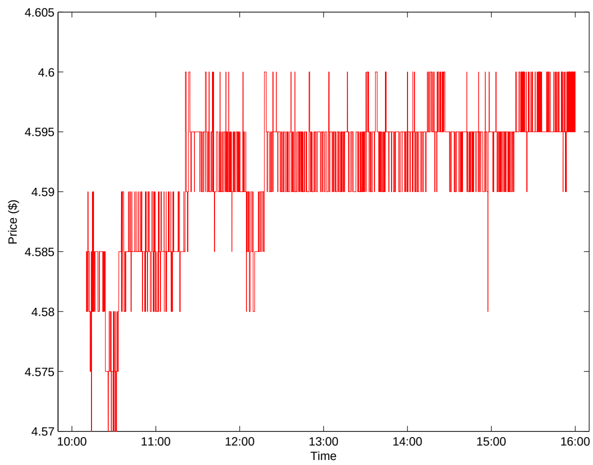
The mean of the realized volatility increases with frequency; this is due to the upward bias that affects the estimator when the price series is affected by market microstructures. The mean of the realized range estimator also exhibits a curious behavior; the mean estimate at the 1-second frequency is always significantly lower than that observed at the 10-second or 1-minute frequency. For the 17 stocks observed, the time between trades averages 9.25 seconds; so estimating volatility at a frequency higher than this results in a downward bias for realized range due to non-synchronous trading. The mean annualized volatility estimated by the generalized variance estimator, however, only changes slightly as frequency increases. This further suggests that generalized variance is robust to the effects of market microstructures.

Similar observations can be made of the variances of each estimator. It should be noted, however, that the variance of the generalized variance estimator is generally larger than that observed for the other estimators. This is because generalized variance does not achieve the same efficiency as the other estimators.

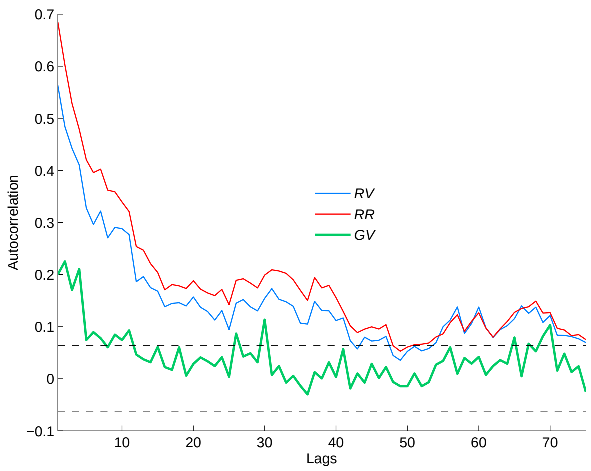
Figure 3 highlights the effect of noise on a time series; it juxtaposes the price of the Westpac (WBC) stock against that of Telstra (TLS). The price series for Westpac is quite reminiscent of Brownian Motion. The price of the Telstra stock observed in Figure 3(b) is extremely noisy in comparison; this is because the movement of TLS is amplified by the effect of tick size. According to ASX rules, the minimum price movement for Telstra is \$0.01. The Telstra price commonly moves less than \$0.05 in a day, which means that price fluctuations of \$0.01 cover a significant proportion of the movement within a day. This results in inaccurate estimates of volatility by realized volatility and realized range. The effect is particularly pronounced in the variance of the realized volatility estimator, which reaches 10071.33 in Table 11. For this time series, the generalized variance estimator appears to ignore the market microstructures, while realized volatility and realized range are overwhelmed by the effect of the frictions.



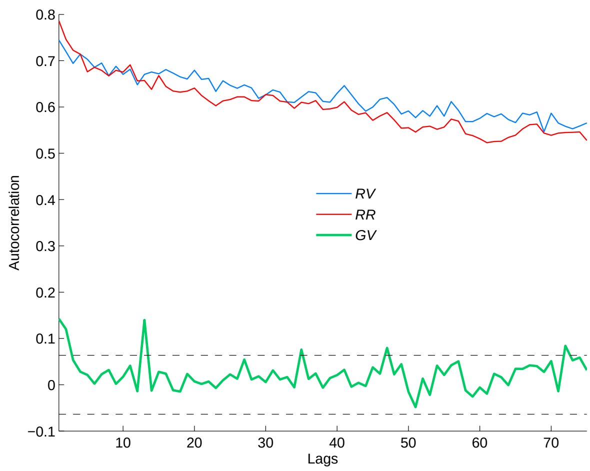
(a) WBC stock price



(b) TLS stock price



(c) Autocorrelation functions of WBC volatility measures



(d) Autocorrelation functions of TLS volatility measures

Figure 3: Comparison of Stock Prices with Low and High Levels of Market Frictions. Panel (a) shows a stock price with low levels of market frictions, Westpac Banking Corporation (WBC), and panel (b) displays a stock price that exhibits a high level of noise due to market microstructures, Telstra Corporation (TLS). Both series are for the trading day on 25/02/13. Panels (c) and (d) display autocorrelation functions of the volatility measures computed for WBC and TLS, respectively.

4.3.2 Distributions of Standardized Daily Returns

Geometric Brownian Motion theory expects that returns standardized by the standard deviation of prices are Gaussian. Consequently, Jarque and Bera (1987) tests for normality were applied to the distributions of daily returns standardized by the square root of each volatility estimate. This is the methodology employed by Martens and van Dijk (2007). The summary statistics and Jarque-Bera p-values for each of the distributions are listed in Tables 12, 13 and 14 in Appendix C; p-values greater than 0.05 represent non-rejection of the null hypothesis and are bold in the tables.

The variance of standardized returns shows the opposite pattern to that observed for the unconditional distributions; the variance decreases with frequency for realized volatility, and at high frequencies, it is quite high for realized range. Daily returns standardized by generalized variance have habitually high variances. The standardization routinely reduces the skewness of returns.[2] Realized volatility and range appear to reduce the excess kurtosis of returns, attaining values close to three, the kurtosis of a normal distribution. In comparison, the kurtosis of daily returns standardized by generalized variance is often amplified, taking returns further from normality.

Although it has been reported since its establishment by Andersen et al. (2001b) that returns standardized by realized volatility are approximately Gaussian, this paper shows that this is not a general statement; only six out of 17 stocks considered exhibit Gaussian returns when standardized by realized volatility. Martens and van Dijk (2007) reported that S&P 500 returns standardized by realized range were Gaussian, while the realized volatility standardization resulted in a non-normal series; however, only eight out of 17 stocks do not reject the null hypothesis of the Jarque and Bera (1987) test when applied to daily returns standardized by realized range. While generalized variance does not achieve Gaussian returns, it is common for realized volatility and realized range to also generate non-normal standardized returns. It is clear that the efficacy of standardization depends on the asset considered.

4.3.3 Comparison of Predictive Ability

As stated in Section 4.2, Diebold and Mariano (1995) tests were conducted, and the winner was determined in a round robin tournament of estimators. It should be stated that the highest possible number of wins in any cell of the table is 34 (observed for realized volatility when forecasting the realized volatility proxy and assessed by the MAD metric at the 1-second frequency in Table 3(a)). This is because each estimator competes twice for each of the stocks. In addition, the maximum possible number of wins in a row for a metric is 51, as there are three competitions for each of the 17 stocks considered. This maximum is attained by the MAD metric at the 1-second frequency with realized volatility

[2] ANZ is an exception.

Table 3: Number of Wins from Diebold and Mariano (1995) tests. R^2, MSE, MAD, ENC as defined in Section 4.2. The maximum possible number of wins in any cell of the table is 34. The highest possible number of wins in any row of the table is 51. These maximums may not be achieved. Higher is better and represents better predictive ability. Panel (a) compares lagged volatility estimates as forecasts, while panel (b) compares the forecasts from Realized GARCH models. Clearly realized volatility and realized range are capable of forecasting themselves. Realized range is judged to be superior as it records more wins when forecasting realized volatility than realized volatility does when forecasting realized range. Generalized variance appears to contain information that is different from the other estimators, as it does not forecast the other estimators well but is equally not forecasted well by them.

(a) Raw lagged volatility estimates as forecasts

	$\widehat{RV}_i$				$\widehat{RR}_i$				$\widehat{GV}_i$			
	R^2	MSE	MAD	ENC	R^2	MSE	MAD	ENC	R^2	MSE	MAD	ENC
Proxy: $\widehat{RV}_i$												
1 sec	33	33	34	19	14	2	2	2	0	8	15	12
10 sec	26	31	34	21	17	14	17	13	0	0	0	0
1 min	18	19	30	13	17	14	17	16	0	0	0	0
5 min	12	12	18	15	27	22	32	23	0	0	0	0
Proxy: $\widehat{RR}_i$												
1 sec	18	0	0	5	15	33	34	26	0	12	17	10
10 sec	20	3	4	8	17	33	34	26	0	1	11	0
1 min	16	12	17	19	19	32	34	22	0	0	0	0
5 min	16	12	17	19	31	25	34	22	0	0	0	0
Proxy: $\widehat{GV}_i$												
1 sec	4	2	0	6	0	17	32	0	0	13	16	7
10 sec	3	2	3	8	1	22	31	20	0	6	15	3
1 min	5	6	5	15	1	23	31	19	0	4	8	4
5 min	4	10	13	14	4	16	29	19	0	3	4	0

(b) Realized GARCH forecasts

	RV-RG$_{i-1}$				RR-RG$_{i-1}$				GV-RG$_{i-1}$			
	R^2	MSE	MAD	ENC	R^2	MSE	MAD	ENC	R^2	MSE	MAD	ENC
Proxy: RV-RG$_i$												
1 sec	26	32	34	18	16	17	17	11	0	0	0	0
10 sec	25	28	29	6	17	17	17	12	0	0	0	0
1 min	28	28	34	10	17	17	17	7	0	0	0	0
5 min	25	25	34	15	16	16	15	9	0	0	0	1
Proxy: RR-RG$_i$												
1 sec	14	7	11	1	33	34	34	21	0	5	5	0
10 sec	17	12	12	4	31	32	34	15	0	4	5	0
1 min	17	15	16	8	31	31	34	17	0	0	1	0
5 min	14	12	17	11	30	28	34	16	0	0	0	0
Proxy: GV-RG$_i$												
1 sec	3	2	5	6	4	10	11	5	32	31	34	26
10 sec	2	0	0	5	6	9	11	9	33	31	34	23
1 min	3	3	4	6	1	4	6	4	32	30	34	21
5 min	1	3	4	5	5	3	8	3	29	28	34	16

as the proxy; 34 wins are awarded to realized volatility, two to realized range and 15 to generalized variance. As there is the possibility that in a competition, neither estimator is significantly better, these maximums are not always attained.

The number of wins when considering raw lagged volatility estimates as forecasts is displayed in Table 3(a). It is clear that realized volatility is quite adept at forecasting itself one period ahead, especially as frequency increases; this is shown by wins close to the maximum of 34. It is interesting, however, that at the 5-minute frequency, the realized range estimator is better at forecasting realized volatility one period ahead than realized volatility itself. The realized range estimator wins 27 competitions when assessed by R^2, compared to only 12 won by realized volatility. Generalized variance does not capture the behavior of realized volatility at low frequencies. When prices are observed every second, however, generalized variance can be assessed as better than realized range at forecasting realized volatility.

Realized range is extremely good at forecasting itself one period ahead, with the number of wins even higher than those observed for realized volatility. Assessed by the MAD metric, the realized range estimator is significantly better at forecasting itself for every single stock at every frequency when compared to its competitors. At low frequencies, realized volatility is better than generalized variance at forecasting realized range; however, the opposite holds true at the 1-second frequency.

In contrast to realized volatility and realized range, generalized volatility does not excel at forecasting itself one period ahead. The highest number of wins recorded is 16 when assessed by the MAD metric at the 1-second frequency. This is because realized volatility and realized range display volatility clustering. This results in high first-order autocorrelation, as observed in Figures 3(c) and 3(d), which is a strong indicator of performance when assessed using these one-step-ahead forecast comparisons. Because generalized variance is convergent in distribution and the variance of its unconditional distribution is higher, its realizations have a weaker time dependence structure than realized range and realized volatility. This results in lower first-order autocorrelation, and thus it does not perform as well as the other estimators when assessed by this methodology.

Martens and van Dijk (2007) concluded that realized range was superior to realized volatility. This conclusion was reached because while both were excellent at forecasting themselves one period ahead, it was assessed that realized range was better at forecasting realized volatility than vice versa. This paper confirms these results when the methodology is applied to ASX 20 stocks. The low first-order autocorrelation hampers the performance of generalized variance, such that realized range is the best at forecasting generalized variance one step ahead. As a result, even when the new estimator, generalized variance, is considered, the realized range estimator is the best estimator of stock volatility when assessed using Diebold and Mariano (1995) tests on raw lagged volatility estimates.

Table 3(b) displays the number of wins given to each estimator from Diebold and

Mariano (1995) tests conducted on forecasts from the Realized GARCH model. The realized volatility Realized GARCH model will be denoted as RV-RG. The other models will be similarly denoted as RR-RG for realized range and GV-RG for generalized variance. The R^2, MSE and MAD metrics all clearly show that the RV-RG model is excellent at forecasting its own one-step-ahead conditional variance. It is generally significantly better than the RR-RG and GV-RG models. When the RR-RG and GV-RG models compete in forecasting the RV-RG model one step ahead, the GV-RG model only records one win, and the RR-RG model is judged to be significantly better than the GV-RG model for most stocks, frequencies and metrics.

A similar story is told for the RR-RG model; it is very capable at forecasting itself one period ahead, winning the vast majority of Diebold and Mariano (1995) tests in which the model competed. The RV-RG model is better than the GV-RG model when forecasting the RR-RG model. A comparison of realized volatility and realized range suggests that realized range is superior, as the RR-RG model is better at forecasting the RV-RG model than the RV-RG model is at forecasting the RR-RG model. This reinforces the result from Table 3(a).

It is clear that the GV-RG model still does not capture the behavior of realized volatility and realized Range-based Realized GARCH models. The RV-RG and RR-RG models, however, also fail to capture the patterns in the GV-RG one-step-ahead forecasts. The maximum number of wins by either the RV-RG or the RR-RG model in Table 3(b) is 11 for the MAD metric when prices are observed each second. The GARCH equation in the the Realized GARCH model compensates for the low first-order autocorrelation of generalized variance. This allows the GV-RG model to be equally as good at forecasting itself as the RV-RG and RR-RG models are at forecasting themselves.

As the GV-RG model does not forecast the other models well and the other models similarly do not excel at forecasting the GV-RG model, the conclusion may be reached that generalized variance captures different information than realized volatility or realized range. This is due to its construction; it does not consider ranges, and the returns that it employs in its calculation are quite different from those used to calculate realized volatility.

Table 4 contains the cumulative log predictive likelihoods for each of the models for each of the stocks. The model that has the highest predictive likelihood is the model that best describes the density of returns and thus is the model with the best fit. For each stock and frequency, the highest cumulative log predictive likelihood is shown in bold.

There is no clear pattern or consensus about the estimator that provides the best fit; it appears to be dependent on the properties of the price series considered. For example, the GV-RG model has the best fit for three of the frequencies for AMP, but it has the worst fit for the price of ANZ stock. Similarly, the RV-RG model is the best for QBE stock at all frequencies but performs poorly when fitted to the price of Telstra shares. ANZ, NAB, WBC and WOW all select the RR-GV model as the best fit at all frequencies; however,

Table 4: Cumulative predictive log likelihoods for ASX 20 Stocks. This is a measure of how well each Realized GARCH model utilizing realized volatility (RV-RG), realized range (RR-RG) or generalized variance (GV-RG) fits the price series of each stock. Higher is better, and the highest for each frequency and for each stock is in bold. There is no estimator that is the best at fitting all stocks. While the realized volatility and realized range exhibit good levels of fit for some stocks, they also poorly fit other stocks. The result of this is that generalized variance is a safe choice due to its robustness, providing dependable results.

	RV-RG	RR-RG	GV-RG	RV-RG	RR-RG	GV-RG	RV-RG	RR-RG	GV-RG
		AMP			ANZ			BHP	
1 sec	2100.78	2477.61	**2555.16**	2648.36	**2675.71**	2664.45	**2599.01**	2595.81	2587.62
10 sec	2170.48	2399.13	**2553.81**	2678.54	**2684.63**	2664.37	**2603.34**	2603.23	2587.57
1 min	2412.26	2464.51	**2554.57**	2678.72	**2684.89**	2667.17	2602.48	**2604.47**	2586.71
5 min	**2560.77**	2548.48	2558.59	2680.25	**2684.11**	2670.79	2596.41	**2603.28**	2585.11
		BXB			CBA			CSL	
1 sec	2294.22	2450.45	**2507.31**	**2842.28**	2827.20	2819.70	**2628.34**	2605.61	2607.49
10 sec	2370.99	2490.86	**2505.34**	2839.50	**2840.75**	2825.31	2625.55	**2628.22**	2601.59
1 min	2484.08	2501.08	**2505.31**	2841.13	**2841.70**	2822.52	2628.26	**2628.82**	2606.71
5 min	2502.94	**2512.03**	2506.94	**2844.20**	2843.87	2826.88	2625.22	**2629.73**	2605.60
		MQG			NAB			NCM	
1 sec	**2319.51**	2304.45	2319.14	2565.66	**2570.68**	2554.93	**2135.57**	2134.44	2105.60
10 sec	2316.96	2315.71	**2319.06**	2574.64	**2582.09**	2562.13	**2135.35**	2131.79	2101.89
1 min	**2322.29**	2319.65	2315.62	2572.60	**2576.26**	2561.24	**2129.46**	2128.27	2101.98
5 min	**2325.63**	2321.96	2316.21	2575.17	**2575.78**	2555.55	2128.33	**2128.74**	2102.13
		ORG			QBE			RIO	
1 sec	2403.34	2437.00	**2437.19**	**2121.18**	2110.58	2089.33	**2392.98**	2380.85	2368.81
10 sec	2432.86	**2446.75**	2435.41	**2115.58**	2109.00	2092.68	**2391.43**	2390.83	2368.25
1 min	**2454.43**	2447.50	2438.21	**2139.60**	2124.03	2083.70	2389.47	**2389.63**	2371.06
5 min	**2449.09**	2444.57	2436.84	**2143.34**	2139.26	2119.35	2386.09	**2390.63**	2365.64
		STO			TLS			WBC	
1 sec	2313.19	**2324.93**	2298.68	1705.16	2527.62	**2716.34**	2573.90	**2597.90**	2593.89
10 sec	2334.77	**2335.09**	2313.72	1889.74	2154.79	**2721.49**	2593.82	**2602.66**	2596.63
1 min	**2335.74**	2334.61	2301.33	2336.90	2281.74	**2717.97**	2603.83	**2603.93**	2596.58
5 min	**2338.97**	2334.40	2297.93	2708.89	2641.12	**2730.12**	2603.70	**2603.96**	2585.66
		WOW			WPL				
1 sec	2788.48	**2874.87**	2854.51	2513.77	2501.77	**2517.73**			
10 sec	2841.73	**2877.88**	2842.09	**2516.04**	2512.13	2515.58			
1 min	2877.96	**2878.96**	2845.94	2514.50	**2518.90**	2515.56			
5 min	2869.88	**2878.22**	2859.53	2511.62	**2517.89**	2517.73			

for stocks such as AMP and MQG, the realized range is not the best fit at any frequency.

It is noteworthy that the predictive likelihood of generalized variance does not change greatly with frequency, providing further evidence of its robustness to the effects of market microstructures. In addition, Table 4 shows that there are circumstances in which generalized variance provides an improvement over both of the benchmark estimators. Consequently, generalized variance provides a safe and dependable estimator of volatility and its behavior and performance are not affected by the microstructure noise inherent in financial markets.

5 Conclusion

This paper proposed a new estimator of volatility that employed the same data as realized volatility and realized range but attempted to provide superior volatility estimation when prices do not follow the ideal theoretical case of Geometric Brownian Motion. Named generalized variance, this estimator considered a simple modification to realized volatility; instead of calculating returns over sequential periods of equal length, generalized variance considers returns over an expanding interval. This simple change resulted in several interesting properties that were theoretically analyzed, investigated in a simulation study and tested on an empirical dataset.

The theoretical analysis showed that in ideal conditions, realized range is the optimal choice to estimate the volatility of a price process. In the presence of the bid-ask bounce, however, it was found that realized volatility, realized range and generalized variance are all upwardly biased. The biases of realized volatility and realized range grow linearly with the frequency of observations. The bias of generalized variance, however, is bounded. Consequently, generalized variance can provide substantial improvements in volatility estimation in the presence of market noise, especially at high frequencies.

The simulation study tested the robustness of generalized variance to two market microstructures: bid-ask bounce and non-synchronous trading. The simulation showed that while the performance of generalized variance does not match the performance of the other estimators in the ideal case, generalized variance provides more robust estimates of volatility in the presence of market microstructures. In many circumstances, while other estimators performed better at certain frequencies, generalized variance delivered solid performance for all frequencies.

The empirical study showed that realized volatility is quite susceptible to noise in empirical data, as the mean and variance of its realizations often increased with frequency. This was consistent with the theoretical and simulation results. A comparison of forecast accuracy found that realized range was the best overall proxy for the true volatility. When considering the fit of the Realized GARCH model, the empirical analysis found that for several stocks, generalized variance was superior to both realized volatility and realized range.

References

Aitken, M. & Frino, A. (1996). The determinants of market bid ask spreads on the australian stock exchange: Cross-sectional analysis. *Accounting & Finance*, *36*(1), 51–63.

Aït-Sahalia, Y., Mykland, P. A. & Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies*, *18*(2), 351–416.

Andersen, T. G., Bollerslev, T. P., Diebold, F. X. & Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*(1), 43–76.

Andersen, T. G., Bollerslev, T. P., Diebold, F. X. & Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, *96*(453), 42–55.

Andersen, T. G., Bollerslev, T. P., Diebold, F. X. & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, *71*(2), 579–625.

Bandi, F. M. & Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies*, *75*(2), 339–369.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2008). Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, *76*(6), 1481–1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. & Shephard, N. (2009). Realised kernels in practice: Trades and quotes. *The Econometrics Journal*, *12*(3), C1–C32.

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

Barndorff-Nielsen, O. E. & Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37.

Beckers, S. (1983). Variances of security price returns based on high, low, and closing prices. *The Journal of Business*, *56*(1), 97–112.

Christensen, K. & Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics*, *141*(2), 323–349.

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*(2), 385–407.

Diebold, F. X. & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253–263.

Garman, M. B. & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *The Journal of Business*, *53*(1), 67–78.

Hansen, P. R., Huang, Z. & Shek, H. H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, *27*(6), 877–906.

Hansen, P. R. & Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, *24*(2), 127–161.

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

Jarque, C. M. & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, *55*(2), 163–172.

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

McAleer, M. & Medeiros, M. C. (2008). Realized volatility: A review. *Econometric Reviews*, *27*(1-3), 10–45.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1-2), 125–144.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, *8*(4), 323–361.

Papavassiliou, V. G. (2012). The efficiency of the realized range measure of daily volatility: Evidence from Greece. *Economic Notes*, *41*(3), 173–182.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, *53*(1), 61–65.

Patton, A. J. (2011). Data-based ranking of realised volatility estimators. *Journal of Econometrics*, *161*(2), 284–303.

Revuz, D. & Yor, M. (1999). *Continuous Martingales and Brownian Motion.* Springer-Verlag, Berlin.

Rogers, L. C. G. & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, *1*(4), 504–512.

Todorova, N. (2012). Volatility estimators based on daily price ranges versus the realized range. *Applied Financial Economics*, *22*(3), 215–229.

Tsay, R. (2010). *Analysis of Financial Time Series* (3rd ed.). John Wiley & Sons, Inc., Hoboken, New Jersey.

Vipul, J. J. (2008). Estimation and forecasting of stock volatility with range-based estimators. *Journal of Futures Markets*, *28*(6), 561–581.

Vortelinos, D. I. (2014). Optimally sampled realized range-based volatility estimators. *Research in International Business and Finance*, *30*, 34–50.

Wilmott, P. (2007). *Paul Wilmott Introduces Quantitative Finance* (2nd ed.). Wiley-Interscience, New York, NY, USA.

Zhang, L., Mykland, P. A. & 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*(472), 1394–1411.

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

A Theoretical Appendix

A.1 Necessary Definitions and Lemmas

Definition A.1 (Wilmott (2007)). *The stochastic integral of a function $f(\tau)$ from $\tau = 0$ to $\tau = t$, is given by*

$$\int_0^t f(\tau)\; dX(\tau) = \lim_{n\to\infty} \sum_{k=1}^{n} f(t_{k-1}) \Big[X(t_k) - X(t_{k-1}) \Big] \,,$$

where $t_k = kt/n$ and $X(\tau)$ is a stochastic process.

Definition A.2 (Revuz and Yor (1999)). *Brownian motion $\{W(t)\}$ is a stochastic process with the following three properties:*

1. *(Independence of increments): For all $0 = t_0 < t_1 < \ldots < t_m$, the increments $W(t_1) - W(t_0)\,,\, W(t_2) - W(t_1)\,, \ldots, W(t_m) - W(t_{m-1})$ are independent.*
2. *(Stationary normal increments): $W(t) - W(s) \sim \mathcal{N}(0,\, t-s)$.*
3. *(Continuity of paths): $W(t)\,,\, t \geq 0$ are continuous functions of t.*

Definition A.3. *A logarithmic price process is defined in this chapter by,*

$$\log(P_t) = \log(P_0) + \int_0^t \sigma\, dW_s\,,$$

or equivalently, using $\log(P_t) = p_t$,

$$p_t = p_0 + \int_0^t \sigma\, dW_s\,.$$

Lemma A.1. *Let W_s be a Brownian Motion. Then,*

$$\mathbb{E}\left[\left(\int_a^{t+a} \sigma\, dW_s \right)^r \right] = \mathbb{E}\left[\left(\int_0^{t} \sigma\, dW_s \right)^r \right] .$$

Proof. True by stationarity of Brownian Motion increments from Definition A.2.

Lemma A.2. *Let W_s be a Brownian Motion. Then for $0 \leq t_a \leq t_b$,*

$$\int_{t_a}^{t_b} \sigma\, dW_s = \Big[W(t_b) - W(t_a) \Big]\, \sigma\,.$$

Proof.

$$\begin{aligned}
\int_0^t \sigma\, dW_s &= \lim_{n\to\infty} \sum_{k=1}^{n} \sigma \Big[W(t_k) - W(t_{k-1})\Big] \\
&= \sigma \lim_{n\to\infty} \sum_{k=1}^{n} \left[W\left(\frac{kt}{n}\right) - W\left(\frac{(k-1)\,t}{n}\right)\right] \\
&= \sigma \lim_{n\to\infty} \left[W(t) - W\left(\frac{(n-1)\,t}{n}\right) + W\left(\frac{(n-1)\,t}{n}\right)\right. \\
&\qquad \left. - \ldots - W\left(\frac{t}{n}\right) + W\left(\frac{t}{n}\right) - W(0)\right] \\
&= \sigma \lim_{n\to\infty} \Big[W(t) - W(0)\Big] \\
&= \Big[W(t) - W(0)\Big]\,\sigma\,. \\
\int_{t_a}^{t_b} \sigma\, dW_s &= \int_0^{t_b} \sigma\, dW_s - \int_0^{t_a} \sigma\, dW_s \\
&= \Big[W(t_b) - W(0)\Big]\,\sigma - \Big[W(t_a) - W(0)\Big]\,\sigma \\
&= \Big[W(t_b) - W(t_a)\Big]\,\sigma\,,
\end{aligned}$$

where we have used Definition A.1.

Lemma A.3. *Let W_s be a Brownian Motion. Then for $0 \leq t_a \leq t_b$,*

$$\mathbb{E}\left[\int_{t_a}^{t_b} \sigma\, dW_s\right] = 0\,.$$

Proof.

$$\begin{aligned}
\mathbb{E}\left[\int_{t_a}^{t_b} \sigma\, dW_s\right] &= \mathbb{E}\left[\sigma \Big\{W(t_b) - W(t_a)\Big\}\right] \\
&= \mathbb{E}\left[\Big\{W(t_b) - W(t_a)\Big\}\right]\sigma \\
&= 0\,,
\end{aligned}$$

where we have used Lemma A.2 and expectation of Brownian Motion increments from Definition A.2.

Lemma A.4. *Let W_s be a Brownian Motion. Then for non-overlapping increments:* $0 \leq t_a \leq t_b$, $0 \leq t_c \leq t_d$,

$$\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\right] = 0\,.$$

Proof.

$$\begin{aligned}\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\right] &= \mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)\right]\mathbb{E}\left[\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\right] \\ &= 0\,,\end{aligned}$$

where we have used independence of non-overlapping Brownian Motion increments from Definition A.2, $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$ for independent variables X and Y and Lemma A.3.

Lemma A.5. *Let W_s be a Brownian Motion. Then for $t_a \leq t_b$,*

$$\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2\right] = \sigma^2\,(t_b - t_a)\ .$$

Proof.

$$\begin{aligned}\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2\right] &= \mathbb{E}\left[\left\{\sigma\left[W(t_b) - W(t_a)\right]\right\}^2\right] \\ &= \mathbb{E}\left[\left\{W(t_b) - W(t_a)\right\}^2\right]\sigma^2 \\ &= \left[\mathrm{Var}\Big(W(t_b) - W(t_a)\Big) + \left(\mathbb{E}\left[W(t_b) - W(t_a)\right]\right)^2\right]\sigma^2 \\ &= (t_b - t_a)\,\sigma^2\,,\end{aligned}$$

where we have used Lemma A.2, linearity of expectation, $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \Big(\mathbb{E}[X]\Big)^2$ and expectation and variance of Brownian Motion increments from Definition A.2.

Lemma A.6. *Let W_s be a Brownian Motion. Then for $t_a \leq t_b$,*

$$\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^4\right] = 3\,(t_b - t_a)^2\,\sigma^4\,.$$

Proof.

$$\begin{aligned}\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^4\right] &= \mathbb{E}\left[\left\{\sigma\left[W(t_b) - W(t_a)\right]\right\}^4\right] \\ &= \mathbb{E}\left[\left\{W(t_b) - W(t_a)\right\}^4\right]\sigma^4 \\ &= 3\,(t_b - t_a)^2\,\sigma^4\,,\end{aligned}$$

where we have used Lemma A.2, linearity of expectation, the distribution of Brownian Motion increments from Definition A.2 and the fourth moment of $\mathcal{N}(0,\,v)$ is $3v^2$.

Lemma A.7. *Let W_s be a Brownian Motion. Then for non-overlapping increments:* $0 \le t_a \le t_b$, $0 \le t_c \le t_d$,

$$\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^3 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\right] = 0\,.$$

Proof.

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^3 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\right] \\
&\quad = \mathbb{E}\left[\left\{\sigma\left[W(t_b) - W(t_a)\right]\right\}^3 \left\{\sigma\left[W(t_d) - W(t_c)\right]\right\}\right] \\
&\quad = \mathbb{E}\left[\left\{W(t_b) - W(t_a)\right\}^3\right] \mathbb{E}\left[\left\{W(t_d) - W(t_c)\right\}\right] \sigma^4 \\
&\quad = 0\,,
\end{aligned}$$

where we have used Lemma A.2, linearity of expectation, independence and distribution of non-overlapping Brownian Motion increments from Definition A.2, $\mathbb{E}\left[X^3 Y\right] = \mathbb{E}\left[X^3\right]\mathbb{E}\left[Y\right]$ for independent X and Y, skewness of $\mathcal{N}(0,\, v)$ is 0 and Lemma A.3.

Lemma A.8. *Let W_s be a Brownian Motion. Then for non-overlapping increments:* $0 \le t_a \le t_b$, $0 \le t_c \le t_d$,

$$\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] = (t_b - t_a)\,(t_d - t_c)\,\sigma^4\,.$$

Proof.

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] \\
&\quad = \mathbb{E}\left[\left\{\sigma\left[W(t_b) - W(t_a)\right]\right\}^2 \left\{\sigma\left[W(t_d) - W(t_c)\right]\right\}^2\right] \\
&\quad = \mathbb{E}\left[\left\{W(t_b) - W(t_a)\right\}^2\right] \mathbb{E}\left[\left\{W(t_d) - W(t_c)\right\}^2\right] \sigma^4 \\
&\quad = (t_b - t_a)\,(t_d - t_c)\,\sigma^4\,,
\end{aligned}$$

where we have used Lemma A.2, linearity of expectation, independence and variance of non-overlapping Brownian Motion increments from Definition A.2 and $\mathbb{E}\left[X^2 Y^2\right] = \mathbb{E}\left[X^2\right]\mathbb{E}\left[Y^2\right]$ for independent X and Y.

Lemma A.9. *Let W_s be a Brownian Motion. Then for overlapping increments,* $t_a \le \boldsymbol{t_c} \le \boldsymbol{t_b} \le t_d$,

$$\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] = \Big[(t_c - t_a)\,(t_d - t_c) + (t_b - t_c)\,(2t_b - 3t_c + t_d)\Big]\,\sigma^4\,.$$

Proof.

$$
\begin{aligned}
\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] &= \mathbb{E}\left[\left(\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)^2 + \left(\int_{t_c}^{t_b} \sigma\, dW_s\right)\right)^2 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] \\
&= \mathbb{E}\left[\left(\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)^2 + 2\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)\left(\int_{t_c}^{t_b} \sigma\, dW_s\right) + \left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^2\right)\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] \\
&= \mathbb{E}\left[\left(\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)^2 + \left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^2\right)\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] \\
&= \mathbb{E}\left[\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)^2\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2 + \left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^2\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] \\
&= \mathbb{E}\left[\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)^2\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2 + \left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^2\left(\left(\int_{t_c}^{t_b} \sigma\, dW_s\right) + \left(\int_{t_b}^{t_d} \sigma\, dW_s\right)\right)^2\right] \\
&= \mathbb{E}\left[\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)^2\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2 + \left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^4 + \left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^2\left(\int_{t_b}^{t_d} \sigma\, dW_s\right)^2\right] \\
&= \mathbb{E}\left[\left(\int_{t_a}^{t_c} \sigma\, dW_s\right)^2\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)^2\right] + \mathbb{E}\left[\left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^4\right] + \mathbb{E}\left[\left(\int_{t_c}^{t_b} \sigma\, dW_s\right)^2\left(\int_{t_b}^{t_d} \sigma\, dW_s\right)^2\right] \\
&= (t_c - t_a)(t_d - t_c)\,\sigma^4 + 3\,(t_b - t_c)\,\sigma^4 + (t_b - t_c)(t_d - t_b)\,\sigma^4 \\
&= (t_c - t_a)(t_d - t_c)\,\sigma^4 + (t_b - t_c)(2t_b - 3t_c + t_d)\,\sigma^4
\end{aligned}
$$

where we have used linearity of expectation and Lemmas A.6 and A.8.

Lemma A.10. *Let W_s be a Brownian Motion. Then for non-overlapping increments:* $0 \le t_a \le t_b$, $0 \le t_c \le t_d$, $0 \le t_e \le t_f$,

$$
\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\left(\int_{t_e}^{t_f} \sigma\, dW_s\right)\right] = 0\,.
$$

Proof.

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)^2 \left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\left(\int_{t_e}^{t_f} \sigma\, dW_s\right)\right] \\
&\quad = \mathbb{E}\left[\left\{\sigma\left[W(t_b) - W(t_a)\right]\right\}^2 \left\{\sigma\left[W(t_d) - W(t_c)\right]\right\}\left\{\sigma\left[W(t_f) - W(t_e)\right]\right\}\right] \\
&\quad = \mathbb{E}\left[\left\{W(t_b) - W(t_a)\right\}^2\right]\mathbb{E}\left[\left\{W(t_d) - W(t_c)\right\}\left\{W(t_f) - W(t_e)\right\}\right]\sigma^4 \\
&\quad = 0\,,
\end{aligned}
$$

where we have used Lemmas A.2 and A.4, independence and variance of non-overlapping Brownian Motion increments from Definition A.2, and for independent X and Y, $\mathbb{E}\left[X^2 Y\right] = \mathbb{E}\left[X^2\right]\mathbb{E}\left[Y\right]$.

Lemma A.11. *Let W_s be a Brownian Motion. Then for non-overlapping increments:* $0 \le t_a \le t_b$, $0 \le t_c \le t_d$, $0 \le t_e \le t_f$, $0 \le t_g \le t_h$

$$
\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\left(\int_{t_e}^{t_f} \sigma\, dW_s\right)\left(\int_{t_g}^{t_h} \sigma\, dW_s\right)\right] = 0\,.
$$

Proof.

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_{t_a}^{t_b} \sigma\, dW_s\right)\left(\int_{t_c}^{t_d} \sigma\, dW_s\right)\left(\int_{t_e}^{t_f} \sigma\, dW_s\right)\left(\int_{t_g}^{t_h} \sigma\, dW_s\right)\right] \\
&= \mathbb{E}\Bigg[\left\{\sigma\left[W(t_b) - W(t_a)\right]\right\}\left\{\sigma\left[W(t_d) - W(t_c)\right]\right\} \\
&\qquad\qquad \times \left\{\sigma\left[W(t_f) - W(t_e)\right]\right\}\left\{\sigma\left[W(t_h) - W(t_g)\right]\right\}\Bigg] \\
&= \mathbb{E}\left[\left\{W(t_b) - W(t_a)\right\}\left\{W(t_d) - W(t_c)\right\}\right] \\
&\qquad\qquad \times \mathbb{E}\left[\left\{W(t_f) - W(t_e)\right\}\left\{W(t_h) - W(t_g)\right\}\right]\sigma^4 \\
&= 0\,,
\end{aligned}
$$

where we have used independence of non-overlapping Brownian Motion increments from Definition A.2, $\mathbb{E}[X_1X_2X_3X_4] = \mathbb{E}[X_1X_2]\,\mathbb{E}[X_3X_4]$ for independent variables, X_1, X_2, X_3 and X_4 and Lemma A.4.

Lemma A.12 (Parkinson (1980))**.** *Assume p_t follows a Brownian Motion process and r is real and ≥ 1. Then,*

$$
\begin{aligned}
\mathbb{E}[R^q] &= \lambda_q \left(\sigma^2\left[t_b - t_a\right]\right)^{q/2}\,, \\
\lambda_q &= \frac{4}{\sqrt{\pi}}\Gamma\left(\frac{q+1}{2}\right)\left(1 - \frac{4}{2^q}\right)\zeta(q-1)\,2^{q/2}\,,
\end{aligned}
$$

where $R = \max_{t_a \leq t \leq t_b}\{p_t\} - \min_{t_a \leq t \leq t_b}\{p_t\}$. Note that $\lambda_1 = 2\sqrt{\frac{2}{\pi}}$, $\lambda_2 = 4\log(2)$, $\lambda_3 = \frac{2}{3}\sqrt{2}\pi^{3/2}$ and $\lambda_4 = 9\zeta(3)$.

Proof. See Parkinson (1980).

A.2 Proof of Propositions

Theorem 1. *Assume that p_t follows a Brownian Motion process and $\widetilde{p}_t = p_t$. Then,*

(a) the expectation, variance and asymptotic distribution of realized volatility are

$$
\begin{aligned}
\mathbb{E}\left[\widehat{RV}_i\right] &= \sigma^2\,, \\
\mathrm{Var}\left(\widehat{RV}_i\right) &= \frac{2\sigma^4}{M}\,, \\
\sqrt{M}\left(\widehat{RV}_i - \sigma^2\right) &\xrightarrow[n\to\infty]{d} \mathcal{N}\left(0,\, 2\sigma^4\right)\,.
\end{aligned}
$$

(b) the expectation, variance and asymptotic distribution of realized range are

$$
\begin{aligned}
\mathbb{E}\left[\widehat{RR}_i\right] &= \sigma^2\,, \\
\mathrm{Var}\left(\widehat{RR}_i\right) &= \frac{\lambda\sigma^4}{M}\,, \\
\sqrt{M}\left(\widehat{RR}_i - \sigma^2\right) &\xrightarrow[n\to\infty]{d} \mathcal{N}\left(0,\, \lambda\,\sigma^4\right)\,.
\end{aligned}
$$

(c) the expectation and variance of generalized variance are

$$\mathbb{E}\left[\widehat{GV}_i\right] = \sigma^2 \,,$$
$$\mathrm{Var}\left(\widehat{GV}_i\right) = \frac{4\left(M^2+M+1\right)}{3M\left(M+1\right)}\sigma^4 \,,$$
$$\lim_{M\to\infty} \mathrm{Var}\left(\widehat{GV}_i\right) = \frac{4\sigma^4}{3} \,.$$

(d) the expectation and variance of corrected generalized variance are

$$\mathbb{E}\left[\widehat{CGV}_i\right] = \sigma^2 \,,$$
$$\mathrm{Var}\left(\widehat{CGV}_i\right) = \frac{2\left(M^2+M+1\right)\sigma^4}{3M\left(M+1\right)} + \frac{\left(M+2\right)\sigma^4}{3M} \,,$$
$$\lim_{M\to\infty} \mathrm{Var}\left(\widehat{CGV}_i\right) = \sigma^4 \,,$$

where $\lambda = \frac{9\,\zeta(3) - \left(4\log(2)\right)^2}{\left(4\log(2)\right)^2} \approx 0.4.$

Proof of Theorem 1(a). See Andersen et al. (2001b), Andersen et al. (2001a) or Barndorff-Nielsen and Shephard (2002).

Proof of Theorem 1(b). See Martens and van Dijk (2007) or Christensen and Podolskij (2007).

Proof of Theorem 1(c). Because of space constraints, the following notation is used;

$$\int_{a\delta}^{b\delta} \sigma\, dW_s = I_{a,b} \,. \tag{7}$$

We use the form of generalised variance in Eq. (4),

$$\widehat{GV}_i = \frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{\widetilde{P}_{(i-1)+j\delta}}{\widetilde{P}_{(i-1)0\delta}}\right)\right]^2 \,.$$

By Lemma A.1, we can drop the $(i-1)$ from the index;

$$\widehat{GV}_i = \frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{0\delta}}\right)\right]^2 \,.$$

Under the assumption that $\widetilde{p}_t = p_t$ and that p_t behaves as described in Definition A.3, then the generalised variance estimator can be written as,

$$
\begin{aligned}
\widehat{GV}_i &= \frac{2}{M+1}\sum_{j=1}^{M}\left(\int_0^{j\delta}\sigma\, dW_s\right)^2, \\
&= \frac{2}{M+1}\sum_{j=1}^{M}\left(\sum_{k=1}^{j}\int_{(k-1)\delta}^{k\delta}\sigma\, dW_s\right)^2, \\
&= \frac{2}{M+1}\sum_{j=1}^{M}\left(\sum_{k=1}^{j} I_{k-1,k}\right)^2, \qquad (8)\\
&= \frac{2}{M+1}\sum_{j=1}^{M}\left[\sum_{k=1}^{j} I_{k-1,k}^2 + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k}I_{l-1,l}\right], \qquad (9)
\end{aligned}
$$

where we have used Eq. (7). By Lemmas A.4 and A.5 and $\delta = \frac{1}{M}$,

$$
\mathbb{E}\left[I_{k-1,k}I_{n-1,n}\right] = \begin{cases} \frac{\sigma^2}{M} & \text{if } k = n\,, \\ 0 & \text{otherwise.} \end{cases} \qquad (10)
$$

We can now evaluate the expectation of generalised variance.

$$
\begin{aligned}
\mathbb{E}\left[\widehat{GV}_i\right] &= \mathbb{E}\left[\frac{2}{M+1}\sum_{j=1}^{M}\left\{\sum_{k=1}^{j} I_{k-1,k}^2 + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k}I_{l-1,l}\right\}\right] \\
&= \frac{2}{M+1}\sum_{j=1}^{M}\left\{\sum_{k=1}^{j}\mathbb{E}\left[I_{k-1,k}^2\right] + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1}\mathbb{E}\left[I_{k-1,k}I_{l-1,l}\right]\right\} \\
&= \frac{2}{M+1}\sum_{j=1}^{M}\left\{\sum_{k=1}^{j}\frac{\sigma^2}{M}\right\} \\
&= \frac{2}{M+1}\frac{\sigma^2}{M}\sum_{j=1}^{M}\sum_{k=1}^{j} 1 \\
&= \frac{2}{M+1}\frac{\sigma^2}{M}\frac{M}{2}(M+1) \\
&= \sigma^2\,, \qquad (11)
\end{aligned}
$$

where we have used Eq. (9), linearity of expectation, Eq. (10) and $\delta = \frac{1}{M}$. To calculate the variance, we shall use the equation $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \left(\mathbb{E}\left[X\right]\right)^2$. We have already calculated $\mathbb{E}\left[X\right]$, so it is left to calculate $\mathbb{E}\left[X^2\right]$. From Eq. (8),

$$\left(\widehat{GV}_i\right)^2 = \left[\frac{2}{M+1}\sum_{j=1}^{M}\left(\sum_{k=1}^{j} I_{k-1,k}\right)^2\right]^2 ,$$

$$= \frac{4}{(M+1)^2}\left\{\sum_{j=1}^{M}\left[\sum_{k=1}^{j} I_{k-1,k}^2\right]^2 + 2\sum_{j=2}^{M}\sum_{l=1}^{j-1}\left(\sum_{k=1}^{j} I_{k-1,k}\right)^2\left(\sum_{m=1}^{l} I_{m-1,m}\right)^2\right\} ,$$

$$= \frac{4}{(M+1)^2}\left\{\underbrace{\sum_{j=1}^{M}\left[\sum_{k=1}^{j} I_{k-1,k}^2\right]^4}_{\text{(A)}} + \underbrace{2\sum_{j=2}^{M}\sum_{l=1}^{j-1}\left(\sum_{k=1}^{j} I_{k-1,k}\right)^2\left(\sum_{m=1}^{l} I_{m-1,m}\right)^2}_{\text{(B)}}\right\} . \tag{12}$$

We will first concentrate on the term (A) in Eq. (12). Using

$$\left(\sum_{j=1}^{j} I_{k-1,k}^2\right)^2 = \left[\sum_{k=1}^{j} I_{k-1,k}^2 + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k} I_{l-1,l}\right] ,$$

we can calculate (A).

$$\begin{aligned}
\text{(A)} &= \sum_{j=1}^{M}\left[\sum_{k=1}^{j} I_{k-1,k}^2 + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k} I_{l-1,l}\right]^2 \\
&= \sum_{j=1}^{M}\left\{\alpha + \beta + \gamma\right\} ,
\end{aligned} \tag{13}$$

where,

$$\begin{aligned}
\alpha &= \left[\sum_{k=1}^{j} I_{k-1,k}^2\right]^2 \\
&= \sum_{k=1}^{j} I_{k-1,k}^4 + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k}^2 I_{l-1,l}^2 ,
\end{aligned} \tag{14}$$

$$\begin{aligned}
\beta &= 4\left[\sum_{k=1} I_{m-1,m}^2\right]\left[\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k} I_{l-1,l}\right] \\
&= 4\sum_{m=1}^{j}\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{m-1,m}^2 I_{k-1,k} I_{l-1,l} ,
\end{aligned} \tag{15}$$

$$\begin{aligned}
\gamma &= 4\left[\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k} I_{l-1,l}\right]^2 \\
&= 4\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k}^2 I_{l-1,l}^2 + 8\sum_{k=3}^{j}\sum_{l=1}^{k}\sum_{m=2}^{k-1}\sum_{n=1}^{m-1} I_{k-1,k} I_{l-1,l} I_{m-1,m} I_{n-1,n} .
\end{aligned} \tag{16}$$

Now, collecting Lemmas A.6, A.7, A.8, A.10, A.11 and $\delta = \frac{1}{M}$, we can say that,

$$
\mathbb{E}\left[I_{k-1,k}I_{l-1,l}I_{q-1,q}I_{r-1,r}\right] = \begin{cases} \frac{3\sigma^4}{M} & \text{if } k=l=q=r\,, \\ \frac{\sigma^4}{M} & \text{if } k=l\neq q=r\,, \\ \frac{\sigma^4}{M} & \text{if } k=q\neq l=r\,, \\ \frac{\sigma^4}{M} & \text{if } k=r\neq l=q\,, \\ 0 & \text{otherwise.} \end{cases} \tag{17}
$$

Using Eq. (14), (15), (16), (17) and linearity of expectation, we can determine the expectations of α, β and γ.

$$
\begin{aligned}
\mathbb{E}\left[\alpha\right] &= \mathbb{E}\left[\sum_{k=1}^{j} I_{k-1,k}^4 + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k}^2 I_{l-1,l}^2\right] \\
&= \sum_{k=1}^{j} \mathbb{E}\left[I_{k-1,k}^4\right] + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} \mathbb{E}\left[I_{k-1,k}^2 I_{l-1,l}^2\right] \\
&= \sum_{k=1}^{j} \frac{3\sigma^4}{M} + 2\sum_{k=2}^{j}\sum_{l=1}^{k-1} \frac{\sigma^4}{M} && (18) \\
\mathbb{E}\left[\beta\right] &= \mathbb{E}\left[4\sum_{m=1}^{j}\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{m-1,m}^2 I_{k-1,k} I_{l-1,l}\right] \\
&= 4\sum_{m=1}^{j}\sum_{k=2}^{j}\sum_{l=1}^{k-1} \mathbb{E}\left[I_{m-1,m}^2 I_{k-1,k} I_{l-1,l}\right] \\
&= 0 && (19) \\
\mathbb{E}\left[\gamma\right] &= \mathbb{E}\left[4\sum_{k=2}^{j}\sum_{l=1}^{k-1} I_{k-1,k}^2 I_{l-1,l}^2 + 8\sum_{k=3}^{j}\sum_{l=1}^{k}\sum_{m=2}^{k-1}\sum_{n=1}^{m-1} I_{k-1,k}I_{l-1,l}I_{m-1,m}I_{n-1,n}\right] \\
&= 4\sum_{k=2}^{j}\sum_{l=1}^{k-1} \mathbb{E}\left[I_{k-1,k}^2 I_{l-1,l}^2\right] + 8\sum_{k=3}^{j}\sum_{l=1}^{k}\sum_{m=2}^{k-1}\sum_{n=1}^{m-1} \mathbb{E}\left[I_{k-1,k}I_{l-1,l}I_{m-1,m}I_{n-1,n}\right] \\
&= 4\sum_{k=2}^{j}\sum_{l=1}^{k-1} \frac{\sigma^4}{M} && (20)
\end{aligned}
$$

We can now calculate the expectation of (A).

$$
\begin{aligned}
\mathbb{E}\Big[(\mathrm{A})\Big] &= \mathbb{E}\left[\sum_{j=1}^{M}\left\{\alpha+\beta+\gamma\right\}\right] \\
&= \sum_{j=1}^{M}\Big\{\mathbb{E}\left[\alpha\right]+\mathbb{E}\left[\beta\right]+\mathbb{E}\left[\gamma\right]\Big\} \\
&= \sum_{j=1}^{M}\left[\sum_{k=1}^{j}\frac{3\sigma^4}{M^2} + 6\sum_{k=2}^{j}\sum_{l=1}^{k-1}\frac{\sigma^4}{M^2}\right] \\
&= \frac{3\sigma^4}{M^2}\sum_{j=1}^{M}\left[j + 2\frac{(j-1)\,j}{2}\right] \\
&= \frac{3\sigma^4}{M^2}\frac{1}{6}M\left(M+1\right)\left(2M+1\right) \\
&= \frac{\left(M+1\right)\left(2M+1\right)\sigma^4}{2M}\,, && (21)
\end{aligned}
$$

where we have used linearity of expectation and Eq. (18), (19) and (20). We now return to the term (B) from Eq. (12).

$$
\begin{aligned}
(\mathrm{B}) &= \sum_{j=2}^{M}\sum_{l=1}^{j-1}\left(\sum_{k=1}^{j} I_{k-1,k}\right)^2\left(\sum_{m=1}^{l} I_{m-1,m}\right)^2 \\
&= \sum_{j=2}^{M}\sum_{l=1}^{j-1}\left[\sum_{k=1}^{j} I_{k-1,k}^2 + 2\sum_{k=2}^{j}\sum_{n=1}^{k-1} I_{k-1,k}I_{n-1,n}\right]\left[\sum_{m=1}^{l} I_{m-1,m}^2 + 2\sum_{m=2}^{l}\sum_{p=1}^{m-1} I_{m-1,m}I_{p-1,p}\right] \\
&= \sum_{j=2}^{M}\sum_{l=1}^{j-1}\left[\pi+\mu+\nu+\rho\right] , \qquad (22)
\end{aligned}
$$

where,

$$
\pi = \left[\sum_{k=1}^{j} I_{k-1,k}^2\right]\left[\sum_{m=1}^{l} I_{m-1,m}^2\right] , \qquad (23)
$$

$$
\begin{aligned}
\mu &= 2\left[\sum_{k=1}^{j} I_{k-1,k}^2\right]\left[\sum_{m=2}^{l}\sum_{p=1}^{m-1} I_{m-1,m}I_{p-1,p}\right] \\
&= 2\sum_{k=1}^{j}\sum_{m=2}^{l}\sum_{p=1}^{m-1} I_{k-1,k}^2 I_{m-1,m}I_{p-1,p} , \qquad (24)
\end{aligned}
$$

$$
\begin{aligned}
\nu &= 2\left[\sum_{m=1}^{l} I_{m-1,m}^2\right]\left[\sum_{k=2}^{j}\sum_{n=1}^{k-1} I_{k-1,k}I_{n-1,n}\right] \\
&= 2\sum_{m=1}^{l}\sum_{k=2}^{j}\sum_{n=1}^{k-1} I_{m-1,m}^2 I_{k-1,k}I_{n-1,n} , \\
\rho &= 4\left[\sum_{k=2}^{j}\sum_{n=1}^{k-1} I_{k-1,k}I_{n-1,n}\right]\left[\sum_{m=2}^{l}\sum_{p=1}^{m-1} I_{m-1,m}I_{p-1,p}\right] \\
&= 4\sum_{k=2}^{j}\sum_{n=1}^{k-1}\sum_{m=2}^{l}\sum_{p=1}^{m-1} I_{k-1,k}I_{n-1,n}I_{m-1,m}I_{p-1,p} . \qquad (25)
\end{aligned}
$$

As we are concerned with the expectation of (B), we shall consider the expectation of π, μ, ν and ρ in turn. For π, note that the expansion and multiplication of the sums will generate jl terms. As $l < j$, there will only be l quartic terms (of the form $I_{k-1,k}^4$). The rest of the terms ($jl - l$ of them) will be of the form $I_{k-1,k}^2 I_{n-1,n}^2$. It remains to use Lemmas A.6 and A.8 to find the expectations of those terms, recalling $\delta = \frac{1}{M}$. Substitution of these results into Eq. (23) will produce,

$$
\mathbb{E}\left[\pi\right] = l\frac{3\sigma^4}{M^2} + (j-1)\,l\frac{\sigma^4}{M^2} . \qquad (26)
$$

The expectation of μ is zero. To see this, recognise that by construction of the sums, $m \neq p$. Thus the expansion of the sums in Eq. (24) will generate terms of the form $I_{k-1,k}^3 I_{n-1,n}$ and $I_{k-1,k}^2 I_{n-1,n}I_{m-1,m}$. Lemmas A.7 and A.10 show the expectation of these terms to be zero. Consequently,

$$
\mathbb{E}\left[\mu\right] = 0 . \qquad (27)
$$

and similarly,

$$
\mathbb{E}\left[\nu\right] = 0 . \qquad (28)
$$

Finally, we turn our attention to γ. By construction of the sums in Eq. (25), it is apparent that $n \neq k$ and $p \neq m$. This means that there are no quartic terms. In addition, the double quadratic terms, $I_{k-1,k}^2 I_{m-1,m}^2$, will only occur when $n = m$ and $p = k$. Recall that $l < j$; consequently, when $k \geq l$, $p \neq k$.

Therefore, the number of double quadratic terms only depends on l. We discard all other terms as we know from Eq. (17) that their expectations are zero.

When $l = 2$, there is only one term, as $(p, m) = (1, 2)$, recalling that we impose the constraint that $n = m$ and $k = p$. When $l = 3$, m can take values $\{2, 3\}$ and therefore p can take values $\{1, 2\}$. As $m \neq p$, we therefore have two extra terms,

$$(p, m) = \begin{cases} \{1, 2\} & l \geq 2 \\ \{1, 3\} & l \geq 3 \\ \{2, 3\} & l \geq 3 \end{cases} .$$

This pattern continues, such that the number of terms for a given l is

$$\sum_{i=1}^{l-1} i = \frac{(l-1)\,l}{2} .$$

Consequently,

$$\mathbb{E}\left[\gamma\right] = 4\frac{(l-1)\,l}{2}\frac{\sigma^4}{M^2} , \tag{29}$$

where we have used Lemma A.8 and $\delta = \frac{1}{M}$. Substituting Eq. (26), (27), (28), (29) into the expectation of (B) in Eq. (22), we find that

$$\begin{aligned}
\mathbb{E}\left[(\mathrm{B})\right] &= \mathbb{E}\left[\sum_{j=2}^{M}\sum_{l=1}^{j-1}\left[\pi + \mu + \nu + \rho\right]\right] , \\
&= \sum_{j=2}^{M}\sum_{l=1}^{j-1}\left\{\mathbb{E}\left[\pi\right] + \mathbb{E}\left[\mu\right] + \mathbb{E}\left[\nu\right] + \mathbb{E}\left[\rho\right]\right\} , \\
&= \sum_{j=2}^{M}\sum_{l=1}^{j-1}\left[\frac{3l\sigma^4}{M^2} + (j-1)\,l\frac{\sigma^4}{M^2} + 2\,(l-1)\,l\frac{\sigma^4}{M^2}\right] , \\
&= \frac{\sigma^4}{M^2}\sum_{j=2}^{M}\sum_{l=1}^{j-1}\left[3l + lj - l + 2l^2 - 2l\right] , \\
&= \frac{\sigma^4}{M^2}\sum_{j=2}^{M}\sum_{l=1}^{j-1}\left[lj + 2l^2\right] , \\
&= \frac{\sigma^4}{M^2}\frac{1}{24}\,(M-1)\,M\,(M+1)\,(7M+2) , \\
&= \frac{(M-1)\,(M+1)\,(7M+2)\,\sigma^4}{24M} ,
\end{aligned} \tag{30}$$

where we have used linearity of expectation and Eq. (22), (26), (27), (28) and (29). We can now evaluate the expectation of $\widehat{GV}_i^2$;

$$
\begin{aligned}
\mathbb{E}\left[\widehat{GV}_i^2\right] &= \mathbb{E}\left[\frac{4}{(M+1)^2}\Big((a)+2(b)\Big)\right],\\
&= \frac{4}{(M+1)^2}\left(\mathbb{E}\left[(a)\right]+2\mathbb{E}\left[(b)\right]\right),\\
&= \frac{4}{(M+1)^2}\left(\frac{(M+1)(2M+1)\sigma^4}{2M}+2\left[\frac{(M-1)(M+1)(7M+2)\sigma^4}{24M}\right]\right),\\
&= \frac{7M^2+7M+4}{3M(M+1)}\sigma^4, \qquad (31)
\end{aligned}
$$

where we have used Eq. (12), (21) and (30) and linearity of expectation. As stated before, the variance can be calculated using $\mathrm{Var}(X)=\mathbb{E}\left[X^2\right]-\Big(\mathbb{E}\left[X\right]\Big)^2$.

$$
\begin{aligned}
\mathrm{Var}\Big(\widehat{GV}_i\Big) &= \mathbb{E}\left[\widehat{GV}_i^2\right]-\Big(\mathbb{E}\left[\widehat{GV}_i\right]\Big)^2\\
&= \frac{7M^2+7M+4}{3M(M+1)}\sigma^4-\sigma^4\\
&= \frac{4\left(M^2+M+1\right)}{3M(M+1)}\sigma^4, \qquad (32)
\end{aligned}
$$

where we have used Eq. (11) and (31). Finally, the limit of the variance of $\widehat{GV}_i$ can be found using L'Hôpital's rule;

$$
\begin{aligned}
\lim_{M\to\infty}\mathrm{Var}\Big(\widehat{GV}_i\Big) &= \lim_{M\to\infty}\frac{4\left(M^2+M+1\right)}{3M(M+1)}\sigma^4\\
&= \frac{4}{3}\sigma^4 \qquad (33)
\end{aligned}
$$

Proof of Theorem 1(d). We use the form of corrected generalised variance in Eq. (5),

$$
\widehat{CGV}_i=\frac{1}{2}\left\{\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{(i-1)+j\delta}}{P_{(i-1)0\delta}}\right)\right]^2+\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{(i-1)+M\delta}}{P_{(i-1)+(j-1)\delta}}\right)\right]^2\right\}.
$$

By Lemma A.1, we can drop the $(i-1)$ from the index:

$$
\widehat{CGV}_i=\frac{1}{2}\left\{\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2+\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right\}.
$$

$$
\mathbb{E}\left[\widehat{CGV}_i\right]=\frac{1}{2}\left\{\mathbb{E}\left[\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right]+\mathbb{E}\left[\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right]\right\},
$$

where we have used linearity of expectation. By Theorem 1(c), we know that,

$$
\mathbb{E}\left[\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right]=\sigma^2.
$$

Similarly, as the result only depends on the number and length of the intervals considered,

$$\mathbb{E}\left[\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right] = \sigma^2 \, .$$

Thus,

$$\mathbb{E}\left[\widehat{CGV}_i\right] = \sigma^2 \, . \tag{34}$$

We now turn our attention to the variance of $\widehat{CGV}_i$. $\widehat{CGV}_i$ can be rewritten as,

$$\begin{aligned}
\widehat{CGV}_i &= \frac{1}{M+1}\left\{\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2 + \sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right\} . \\
\mathrm{Var}\left(\widehat{CGV}_i\right) &= \mathrm{Var}\left(\frac{1}{M+1}\left[\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2 + \sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right]\right) \\
&= \frac{1}{(M+1)^2}\mathrm{Var}\left(\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2 + \sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right) \\
&= \frac{1}{(M+1)^2}\left\{\mathrm{Var}\left(\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right) + \mathrm{Var}\left(\sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right)\right. \\
&\qquad \left. + \, 2\,\underbrace{\mathrm{Cov}\left(\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2, \sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right)}_{\text{(C)}}\right\}
\end{aligned} \tag{35}$$

By Theorem 1(c), we know that,

$$\begin{aligned}
\mathrm{Var}\left(\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right) &= \mathrm{Var}\left(\frac{M+1}{2}\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right) , \\
&= \left(\frac{M+1}{2}\right)^2 \mathrm{Var}\left(\widehat{GV}_i\right) \, .
\end{aligned} \tag{36}$$

And similarly,

$$\mathrm{Var}\left(\sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right) = \left(\frac{M+1}{2}\right)^2 \mathrm{Var}\left(\widehat{GV}_i\right) \, . \tag{37}$$

We turn our attention to (C). Let

$$X = \sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2 \text{ and } Y = \sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2 .$$

Then,

$$\begin{aligned}
\text{(C)} &= \mathrm{Cov}(X, Y) \, , \\
&= \mathbb{E}\left[XY\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right] \, .
\end{aligned} \tag{38}$$

$$
\begin{aligned}
\mathbb{E}\left[XY\right] &= \mathbb{E}\left[\left\{\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right\}\left\{\sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right\}\right] \\
&= \mathbb{E}\left[\sum_{j=1}^{M-1}\sum_{k=j}^{M-1}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\left[\log\left(\frac{P_{M\delta}}{P_{k\delta}}\right)\right]^2+\sum_{j=1}^{M}\sum_{k=0}^{j-1}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\left[\log\left(\frac{P_{M\delta}}{P_{k\delta}}\right)\right]^2\right] \\
&= \sum_{j=1}^{M-1}\sum_{k=j}^{M-1}\mathbb{E}\left[\underbrace{\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\left[\log\left(\frac{P_{M\delta}}{P_{k\delta}}\right)\right]^2}_{j\leq k}\right]+\sum_{j=1}^{M}\sum_{k=0}^{j-1}\mathbb{E}\left[\underbrace{\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\left[\log\left(\frac{P_{M\delta}}{P_{k\delta}}\right)\right]^2}_{j>k}\right] \\
&= \frac{\sigma^4}{M^2}\left[\sum_{j=1}^{M-1}\sum_{k=j}^{M-1}j\left(M-k\right)+\sum_{j=1}^{M}\sum_{k=0}^{j-1}k\left(M-k\right)+\left(j-k\right)\left(2j-3k+M\right)\right] \\
&= \frac{\sigma^4}{M^2}\left[\frac{1}{24}\left(M-1\right)M\left(M+1\right)\left(M+2\right)+\frac{1}{24}M\left(M+1\right)\left(9M^2+17M+10\right)\right] \\
&= \frac{\left(M+1\right)^2\left(5M+4\right)\sigma^4}{12M}\,, && (39)
\end{aligned}
$$

where we have used linearity of expectation, Definition A.3, Lemmas A.8 and A.8 and $\delta = 1/M$.

$$
\begin{aligned}
\mathbb{E}\left[X\right] &= \mathbb{E}\left[\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right] \\
&= \mathbb{E}\left[\frac{M+1}{2}\frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)\right]^2\right] \\
&= \frac{M+1}{2}\sigma^2\,, && (40)
\end{aligned}
$$

where we have used Theorem 1(c). Similarly,

$$
\begin{aligned}
\mathbb{E}\left[Y\right] &= \mathbb{E}\left[\sum_{j=1}^{M}\left[\log\left(\frac{P_{M\delta}}{P_{(j-1)\delta}}\right)\right]^2\right]\,, \\
&= \frac{M+1}{2}\sigma^2\,. && (41)
\end{aligned}
$$

Substituting Eq. (39), (40) and (41) into Eq. (38),

$$
\begin{aligned}
\text{(C)} &= \frac{\left(M+1\right)^2\left(5M+4\right)\sigma^4}{12M}-\frac{\left(M+1\right)^2\sigma^4}{4}\,, \\
&= \frac{\left(M+1\right)^2\left(M+2\right)\sigma^4}{6M}\,. && (42) \\
\mathrm{Var}\left(\widehat{CGV}_i\right) &= \frac{1}{\left(M+1\right)^2}\left[2\left(\frac{M+1}{2}\right)^2\mathrm{Var}\left(\widehat{GV}_i\right)+2\frac{\left(M+1\right)^2\left(M+2\right)\sigma^4}{6M}\right] \\
&= \frac{2\left(M^2+M+1\right)\sigma^4}{3M\left(M+1\right)}+\frac{\left(M+2\right)\sigma^4}{3M}\,, && (43)
\end{aligned}
$$

where we have used Eq. (36), (37) and (42) and Theorem 1(c). By L'Hôpital's rule;

$$
\lim_{M\to\infty}\mathrm{Var}\left(\widehat{CGV}_i\right) = \sigma^4\,. \qquad (44)
$$

Theorem 2. *Assume that* $\log(P_t)$ *follows a Brownian Motion process and* P_t *can be observed as the mid price, the bid price or the ask price with equal probability. Then,*

(a) the expectation and variance of the realized volatility estimator are

$$\begin{aligned}\mathbb{E}\left[\widetilde{RV}_i\right] &= \sigma^2 + \frac{4Mb_1}{9}\,,\\ \operatorname{Var}\left(\widetilde{RV}_i\right) &= \frac{2\sigma^4}{M} + \frac{16b_1\sigma^2}{9} + \frac{4\left(6M-1\right)b_1^2}{81}\,,\end{aligned}$$

(b) the expectation and variance of the realized range estimator are

$$\begin{aligned}\mathbb{E}\left[\widetilde{RR}_i\right] &= \sigma^2 + \sqrt{\frac{2M}{\pi}}\frac{b_2}{\log 2}\sigma + \frac{Mb_2^2}{4\log 2}\,,\\ \operatorname{Var}\left(\widetilde{RR}_i\right) &= \frac{\lambda\sigma^4}{M} + \frac{8\sqrt{2}b_2}{\sqrt{M}}\left(\frac{\pi^{3/2}}{3} - \frac{4\log 2}{\sqrt{\pi}}\right)\sigma^3 + 4b_2^2\left(4\log 2 - \frac{8}{\pi}\right)\sigma^2\,,\end{aligned}$$

(c) the expectation and variance of the generalized variance estimator are

$$\begin{aligned}\mathbb{E}\left[\widetilde{GV}_i\right] &= \sigma^2 + \frac{8Mb_1}{9\left(M+1\right)}\,,\\ \operatorname{Var}\left(\widetilde{GV}_i\right) &= \frac{4\left(M^2+M+1\right)}{3M\left(M+1\right)}\sigma^4 + \frac{32b_1}{9\left(M+1\right)}\sigma^2 + \frac{80Mb_1^2}{81\left(M+1\right)^2}\,,\\ \lim_{M\to\infty}\operatorname{Var}\left(\widetilde{GV}_i\right) &= \frac{4\sigma^4}{3}\,,\\ &= \lim_{M\to\infty}\operatorname{Var}\left(\widehat{GV}_i\right)\,,\end{aligned}$$

(d) the expectation and variance of the corrected generalized variance estimator are

$$\begin{aligned}\mathbb{E}\left[\widetilde{CGV}_i\right] &= \sigma^2 + \frac{8Mb_1}{9\left(M+1\right)}\,,\\ \operatorname{Var}\left(\widetilde{CGV}_i\right) &= \left[\frac{2\left(M^2+M+1\right)}{3M\left(M+1\right)} + \frac{M+2}{3M}\right]\sigma^4 + \frac{16b_1\sigma^2}{9\left(M+1\right)} + \frac{4\left(11M-1\right)b_1^2}{81\left(M+1\right)^2}\,,\\ \lim_{M\to\infty}\operatorname{Var}\left(\widetilde{CGV}_i\right) &= \sigma^4\,,\\ &= \lim_{M\to\infty}\operatorname{Var}\left(\widehat{CGV}_i\right)\,,\end{aligned}$$

where $\lambda = \frac{9\,\zeta(3) - \left(4\log(2)\right)^2}{\left(4\log(2)\right)^2} \approx 0.4$*,* $b_1 = \log(1+s)^2 + \log(1-s)^2 - \log(1+s)\log(1-s) > 0$ *and* $b_2 = \log(1+s) - \log(1-s) > 0$*.*

Proof of Theorem 2(a).

From Table 1, there are 9 possible pairs of mid, bid and ask price. Using $\widetilde{RV}_i = \sum_{j=1}^{M} \log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^2$ from Eq. (2),

$$\begin{aligned}
\mathbb{E}\left[\widetilde{RV}_i\right] &= \sum_{j=1}^{M} \Bigg\{ \frac{1}{3}\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^2\right] + \frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right) + \log(1-s)\right\}^2\right] + \frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right) + \log(1+s)\right\}^2\right] \\
&\quad + \frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right) - \log(1-s)\right\}^2\right] + \frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right) + \log(1+s) - \log(1-s)\right\}^2\right] \\
&\quad + \frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right) - \log(1+s)\right\}^2\right] + \frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right) - \log(1+s) + \log(1-s)\right\}^2\right] \Bigg\}, \\
&= \sum_{j=1}^{M} \left\{ \mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^2\right] + \frac{4}{9}\left(\log(1+s)^2 + \log(1-s)^2 - \log(1+s)\log(1-s)\right) \right\}, \qquad (45) \\
&= \sum_{j=1}^{M} \mathbb{E}\left[\left(\int_{(j-1)\delta}^{j\delta} \sigma\, dW_s\right)^2\right] + \frac{4b_1}{9}, \\
&= \sum_{j=1}^{M} \frac{\sigma^2}{M} + \frac{4b_1}{9}, \\
&= \sigma^2 + \frac{4Mb_1}{9}, \qquad (46)
\end{aligned}$$

where $b_1 = \log(1+s)^2 + \log(1-s)^2 - \log(1+s)\log(1-s)$ and we have used the linearity of expectation, Definition A.3 and Lemma A.5 and $\delta = \frac{1}{M}$.

From Eq. (2),

$$\widetilde{RV}_i^2 = \left[\sum_{j=1}^{M} \log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^2\right]^2, \tag{47}$$

$$= \sum_{j=1}^{M} \log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^4 + 2\sum_{j=1}^{M}\sum_{k=1}^{j-1} \log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^2 \log\left(\frac{\widetilde{P}_{(j-k)\delta}}{\widetilde{P}_{(j-k-1)\delta}}\right)^2, \tag{48}$$

$$\mathbb{E}\left[\widetilde{RV}_i^2\right] = \underbrace{\mathbb{E}\left[\sum_{j=1}^{M} \log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^4\right]}_{\text{(D)}} + \underbrace{\mathbb{E}\left[2\sum_{j=3}^{M}\sum_{k=2}^{j-1} \log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^2 \log\left(\frac{\widetilde{P}_{(j-k)\delta}}{\widetilde{P}_{(j-k-1)\delta}}\right)^2\right]}_{\text{(E)}} + \underbrace{\mathbb{E}\left[2\sum_{j=2}^{M} \log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^2 \log\left(\frac{\widetilde{P}_{(j-1)\delta}}{\widetilde{P}_{(j-2)\delta}}\right)^2\right]}_{\text{(F)}}, \tag{49}$$

where we have used linearity of expectation. We split Eq. (49) and calculate the expectations separately.

$$\text{(E)} = 2\sum_{j=3}^{M}\sum_{k=2}^{j-1} \mathbb{E}\left[\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^2 \log\left(\frac{\widetilde{P}_{(j-k)\delta}}{\widetilde{P}_{(j-k-1)\delta}}\right)^2\right] \tag{50}$$

$$= 2\sum_{j=3}^{M}\sum_{k=2}^{j-1} \left\{\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^2 \log\left(\frac{P_{(j-k)\delta}}{P_{(j-k-1)\delta}}\right)^2\right] + \frac{36b_1}{81}\left(\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^2\right] + \mathbb{E}\left[\log\left(\frac{P_{(j-k)\delta}}{P_{(j-k-1)\delta}}\right)^2\right]\right) + \frac{16b_1^2}{81}\right\} \tag{51}$$

$$= 2\sum_{j=3}^{M}\sum_{k=2}^{j-1} \left\{\mathbb{E}\left[\left(\int_{(j-1)\delta}^{j\delta} \sigma\, dW_s\right)^2 \left(\int_{(j-k-1)\delta}^{(j-k)\delta} \sigma\, dW_s\right)^2\right] + \frac{72b_1}{81}\mathbb{E}\left[\left(\int_{(j-1)\delta}^{j\delta} \sigma\, dW_s\right)^2\right] + \frac{16b_1^2}{81}\right\}$$

$$= 2\sum_{j=3}^{M}\sum_{k=2}^{j-1} \left\{\frac{\sigma^4}{M^2} + \frac{72b_1}{81}\frac{\sigma^2}{M} + \frac{16b_1^2}{81}\right\}$$

$$= (M-2)(M-1)\left[\frac{\sigma^4}{M^2} + \frac{8b_1\sigma^2}{9M} + \frac{16b_1^2}{81}\right], \tag{52}$$

where we have used linearity of expectation, Definition A.3, Lemmas A.5 and A.8 and $\delta = \frac{1}{M}$.

$$
\begin{aligned}
(\mathrm{F}) &= 2\sum_{j=2}^{M}\mathbb{E}\left[\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{(j-1)\delta}}\right)^{2}\log\left(\frac{\widetilde{P}_{(j-1)\delta}}{\widetilde{P}_{(j-2)\delta}}\right)^{2}\right] && (53)\\
&= 2\sum_{j=2}\left\{\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^{2}\log\left(\frac{P_{(j-1)\delta}}{P_{(j-2)\delta}}\right)^{2}\right]+\frac{12b_1}{27}\left(\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^{2}\right]+\mathbb{E}\left[\log\left(\frac{P_{(j-1)\delta}}{P_{(j-2)\delta}}\right)^{2}\right]\right)+\frac{6b_1^2}{27}\right\} && (54)\\
&= 2\sum_{j=2}^{M}\left\{\frac{\sigma^4}{M^2}+\frac{24b_1\sigma^4}{27M}+\frac{6b_1^2}{27}\right\} && (55)\\
&= \frac{2\,(M-1)\,\sigma^4}{M^2}+\frac{16\,(M-1)\,b_1\sigma^2}{9M}+\frac{4\,(M-1)\,b_1^2}{9}\,, && (56)
\end{aligned}
$$

where we have used linearity of expectation, Definition A.3, Lemmas A.5 and A.8 and $\delta=\frac{1}{M}$. Focusing on (D) from Eq. (49) and recalling the possible combinations from Table 1,

$$
\begin{aligned}
(\mathrm{D}) &= \sum_{j=1}^{M}\Bigg\{\frac{1}{3}\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^{4}\right]+\frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)+\log(1+s)\right\}^{4}\right]+\frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)+\log(1-s)\right\}^{4}\right]\\
&\quad+\frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)-\log(1+s)+\log(1-s)\right\}^{4}\right]+\frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)+\log(1-s)\right\}^{4}\right]\\
&\quad+\frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)+\log(1+s)-\log(1-s)\right\}^{4}\right]+\frac{1}{9}\mathbb{E}\left[\left\{\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)-\log(1-s)\right\}^{4}\right]\Bigg\}\,,\\
&= \sum_{j=1}^{M}\left\{\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^{4}\right]+\frac{8b_1}{3}\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{(j-1)\delta}}\right)^{2}\right]+\frac{4b_1^2}{9}\right\}\,, && (57)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{j=1}^{M} \left\{ \mathbb{E}\left[\left(\int_{(j-1)\delta}^{j\delta} \sigma \, dW_s \right)^4 \right] + \frac{8b_1}{3} \mathbb{E}\left[\left(\int_{(j-1)\delta}^{j\delta} \sigma \, dW_s \right)^2 \right] + \frac{4b_1^2}{9} \right\} , \\
&= \sum_{j=1}^{M} \left\{ \frac{3\sigma^4}{M^2} + \frac{8b_1\sigma^2}{3M^2} + \frac{4b_1^2}{9} \right\} , \\
&= \frac{3\sigma^4}{M} + \frac{8b_1\sigma^2}{3} + \frac{4Mb_1^2}{9} , \qquad (58)
\end{aligned}
$$

where we have used linearity of expectation, Definition A.3 and Lemmas A.5 and A.6 and $\delta = \frac{1}{M}$. We can now calculate the expectation of $\widetilde{RV}_i^2$ and thus the variance of $\widetilde{RV}_i$.

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{RV}_i^2\right] &= (\mathrm{D}) + (\mathrm{E}) + (\mathrm{F}) \\
&= \left(\frac{2}{M} - 1 \right) \sigma^4 + \frac{8\,(M+2)\, b_1 \sigma^2}{9} + \frac{4\left(4M^2 + 6M - 1\right) b_1^2}{81} .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathrm{Var}\left(\widetilde{RV}_i^2\right) &= \mathbb{E}\left[\widetilde{RV}_i^2\right] - \left(\mathbb{E}\left[\widetilde{RV}_i\right]\right)^2 , \\
&= \left(\frac{2}{M} - 1 \right) \sigma^4 + \frac{8\,(M+2)\, b_1 \sigma^2}{9} + \frac{4\left(4M^2 + 6M - 1\right) b_1^2}{81} - \left(\sigma^4 + \frac{8Mb_1\sigma^2}{9} + \frac{16M^2 b_1^2}{81} \right) , \\
&= \frac{2\sigma^4}{M} + \frac{16 b_1 \sigma^2}{9} + \frac{4\,(6M-1)\, b_1^2}{81} ,
\end{aligned}
$$

where we have used Eq. (46).

Proof of Theorem 2(b). We use the form of realised range from Eq. (3).

$$
\begin{aligned}
\widetilde{RR}_i &= \frac{1}{4\log 2} \sum_{j=1}^{M} \log\left(\frac{\widetilde{H}_{j\delta}}{\widetilde{L}_{j\delta}} \right)^2 \\
&= \frac{1}{4\log 2} \sum_{j=1}^{M} \left[\log\left(\frac{H_{j\delta}}{L_{j\delta}} \right) + \log\left(\frac{1+s}{1-s} \right) \right]^2 \\
&= \frac{1}{4\log 2} \sum_{j=1}^{M} \left\{ \log\left(\frac{H_{j\delta}}{L_{j\delta}} \right)^2 + 2\log\left(\frac{H_{j\delta}}{L_{j\delta}} \right) b_2 + b_2^2 \right\} \qquad (59)
\end{aligned}
$$

We can calculate the expectation of $\widetilde{RR}_i$ as follows,

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{RR}_i\right] &= \mathbb{E}\left[\frac{1}{4\log 2} \sum_{j=1}^{M} \left\{ \log\left(\frac{H_{j\delta}}{L_{j\delta}} \right)^2 + 2\log\left(\frac{H_{j\delta}}{L_{j\delta}} \right) b_2 + b_2^2 \right\} \right] , \\
&= \frac{1}{4\log 2} \sum_{j=1}^{M} \mathbb{E}\left[\log\left(\frac{H_{j\delta}}{L_{j\delta}} \right)^2 \right] + \frac{2b_2}{4\log 2} \sum_{j=1}^{M} \mathbb{E}\left[\log\left(\frac{H_{j\delta}}{L_{j\delta}} \right) \right] + \frac{Mb_2^2}{4\log 2} , \\
&= \frac{1}{4\log 2} \sum_{j=1}^{M} 4\log 2 \frac{\sigma^2}{M} + \frac{b_2}{2\log 2} \sum_{j=1}^{M} 2\sqrt{\frac{2}{\pi}} \left(\frac{\sigma^2}{M} \right)^{1/2} + \frac{Mb_2^2}{4\log 2} , \\
&= \sigma^2 + \sqrt{\frac{2M}{\pi}} \frac{b_2}{\log 2} \sigma + \frac{Mb_2^2}{4\log 2} ,
\end{aligned}
$$

where $b_2 = \log(1+s) - \log(1-s)$ and we have used Eq. (59), linearity of expectation, Lemma A.12 and

$\delta = \frac{1}{M}$. We can calculate the variance of realised range directly from Eq. (59).

$$
\begin{aligned}
\mathrm{Var}\left(\widetilde{RR}_i\right) &= \mathrm{Var}\left(\frac{1}{4\log 2}\sum_{j=1}^{M}\left\{\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2 + 2\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)b_2 + b_2^2\right\}\right) \\
&= \frac{1}{(4\log 2)^2}\sum_{j=1}^{M}\mathrm{Var}\left(\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2 + 2b_2\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right) + b_2^2\right) \\
&= \frac{1}{(4\log 2)^2}\sum_{j=1}^{M}\left\{\mathrm{Var}\left(\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2\right) + 4b_2^2\mathrm{Var}\left(\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)\right)\right. \\
&\qquad\left. + 4b_2\mathrm{Cov}\left(\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2, \log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)\right)\right\}, \qquad (60)
\end{aligned}
$$

where we have used the independence of Brownian Motion increments from Definition A.2. It is easiest to calculate each term in turn. We proceed in this fashion.

$$
\begin{aligned}
\mathrm{Var}\left(\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2\right) &= \mathbb{E}\left[\left(\log(H_{j\delta}) - \log(L_{j\delta})\right)^4\right] - \left(\mathbb{E}\left[\left(\log(H_{j\delta}) - \log(L_{j\delta})\right)^2\right]\right)^2 \qquad (61) \\
&= 9\,\zeta(3)\,\frac{\sigma^4}{M^2} - \left[4\log(2)\right]^2\frac{\sigma^4}{M^2}, \qquad (62)
\end{aligned}
$$

where $\zeta(x)$ is the Riemann zeta function and we have used $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}\left[X\right])^2$ and Lemma A.12.

$$
\begin{aligned}
\mathrm{Var}\left(\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)\right) &= \mathbb{E}\left[\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2\right] - \left(\mathbb{E}\left[\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)\right]\right)^2 \\
&= 4\log 2\frac{\sigma^2}{M} - \left(2\sqrt{\frac{2}{\pi}}\frac{\sigma}{\sqrt{M}}\right)^2 \\
&= \left(4\log 2 - \frac{8}{\pi}\right)\frac{\sigma^2}{M}, \qquad (63)
\end{aligned}
$$

where we have used $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}\left[X\right])^2$, Lemma A.12 and $\delta = \frac{1}{M}$.

$$
\begin{aligned}
\mathrm{Cov}\left(\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2, \log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)\right) &= \mathbb{E}\left[\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^3\right] - \mathbb{E}\left[\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)^2\right]\mathbb{E}\left[\log\left(\frac{H_{j\delta}}{L_{j\delta}}\right)\right] \\
&= \frac{2}{3}\sqrt{2}\pi^{3/2}\frac{\sigma^3}{M^{3/2}} - 4\log 2\frac{\sigma^2}{M}\sqrt{\frac{8}{\pi}}\frac{\sigma}{M} \\
&= \sqrt{8}\left(\frac{\pi^{3/2}}{3} - \frac{4\log 2}{\sqrt{\pi}}\right)\frac{\sigma^3}{M^{3/2}}, \qquad (64)
\end{aligned}
$$

where we have used $\mathrm{Cov}(X, Y) = \mathbb{E}\left[XY\right] - \mathbb{E}\left[X\right]\mathbb{E}\left[Y\right]$, Lemma A.12 and $\delta = \frac{1}{M}$. Substituting Eq. (62),

(63) and (64) into Eq. (60),

$$\begin{aligned}
\mathrm{Var}\Big(\widetilde{RR}_i\Big) &= \frac{1}{(4\log 2)^2}\sum_{j=1}^{M}\left\{9\,\zeta(3)\,\frac{\sigma^4}{M^2} - \Big[4\log(2)\Big]^2\,\frac{\sigma^4}{M^2} + 4b_2^2\left(4\log 2 - \frac{8}{\pi}\right)\frac{\sigma^2}{M}\right.\\
&\qquad\qquad \left. + 4b_2\sqrt{8}\left(\frac{\pi^{3/2}}{3} - \frac{4\log 2}{\sqrt{\pi}}\right)\frac{\sigma^3}{M^{3/2}}\right\},\\
&= \frac{\lambda\sigma^4}{M} + \frac{8\sqrt{2}b_2}{\sqrt{M}}\left(\frac{\pi^{3/2}}{3} - \frac{4\log 2}{\sqrt{\pi}}\right)\sigma^3 + 4b_2^2\left(4\log 2 - \frac{8}{\pi}\right)\sigma^2\,,
\end{aligned}$$

where $\lambda = \frac{9\,\zeta(3) - 4\log(2)}{(4\log(2))^2}$.

Proof of Theorem 2(c). We begin with the form of generalised variance from Eq. (4),

$$\widetilde{GV}_i = \frac{2}{M+1}\sum_{j=1}^{M}\left[\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{0\delta}}\right)\right]^2. \tag{65}$$

Due to its similarity with realised volatility, we can find the expectation by altering Eq. (45), to produce

$$\begin{aligned}
\mathbb{E}\left[\widetilde{GV}_i\right] &= \frac{2}{M+1}\sum_{j=1}^{M}\left\{\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)^2\right] + \frac{4b_1}{9}\right\},\\
&= \frac{2}{M+1}\sum_{j=1}^{M}\mathbb{E}\left[\left(\int_0^{j\delta}\sigma\,dW_s\right)^2\right] + \frac{4b_1}{9}\,,\\
&= \frac{2}{M+1}\sum_{j=1}^{M}\frac{j\sigma^2}{M} + \frac{4b_1}{9}\,,\\
&= \sigma^2 + \frac{8Mb_1}{9\,(M+1)}\,,
\end{aligned} \tag{66}$$

where $b_1 = \log(1+s)^2 + \log(1-s)^2 - \log(1+s)\log(1-s)$ and we have used Eq. (65), linearity of expectation, Lemma A.5 and $\delta = \frac{1}{M}$. Similarly, we can alter Eq. (47) and (48) to suit generalised variance,

$$\begin{aligned}
\widetilde{GV}_i^2 &= \left[\frac{2}{M+1}\sum_{j=1}^{M}\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{0\delta}}\right)^2\right]^2,\\
&= \frac{4}{(M+1)^2}\left[\sum_{j=1}^{M}\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{0\delta}}\right)^4 + 2\sum_{j=1}^{M}\sum_{k=1}^{j-1}\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{0\delta}}\right)^2\log\left(\frac{\widetilde{P}_{(j-k)\delta}}{\widetilde{P}_{0\delta}}\right)^2\right].
\end{aligned}$$

We now turn to the expectation of the square of generalised variance.

$$\mathbb{E}\left[\widetilde{GV}_i^2\right] = \frac{4}{(M+1)^2}\underbrace{\mathbb{E}\left[\sum_{j=1}^{M}\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{0\delta}}\right)^4\right]}_{\text{(G)}}$$
$$+\frac{4}{(M+1)^2}\underbrace{\mathbb{E}\left[2\sum_{j=2}^{M}\sum_{k=1}^{j-1}\log\left(\frac{\widetilde{P}_{j\delta}}{\widetilde{P}_{0\delta}}\right)^2\log\left(\frac{\widetilde{P}_{(j-k)\delta}}{\widetilde{P}_{0\delta}}\right)^2\right]}_{\text{(H)}}, \tag{67}$$

where we have used linearity of expectation. Once again, it is easiest to tackle the derivation term by term.

$$\begin{aligned}
\text{(G)} &= \sum_{j=1}^{M}\left\{\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)^4\right]+\frac{8b_1}{3}\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)^2\right]+\frac{4b_1^2}{9}\right\}\\
&= \sum_{j=1}^{M}\left\{\mathbb{E}\left[\left(\int_0^{j\delta}\sigma\, dW_s\right)^4\right]+\frac{8b_1}{3}\mathbb{E}\left[\left(\int_0^{j\delta}\sigma\, dW_s\right)^2\right]+\frac{4b_1^2}{9}\right\}\\
&= \sum_{j=1}^{M}\left\{\frac{3j^2\sigma^4}{M^2}+\frac{8b_1 j\sigma^2}{3M}+\frac{4b_1^2}{9}\right\}\\
&= \frac{(M+1)(2M+1)}{2M}\sigma^4+\frac{12(M+1)}{9}\sigma^2+\frac{4Mb_1^2}{9},
\end{aligned} \tag{68}$$

where we have used Definition A.3 and Lemmas A.5 and A.6 and $\delta = \frac{1}{M}$.

$$
\begin{aligned}
(\mathrm{H}) &= 2\sum_{j=2}^{M}\sum_{k=1}^{j-1}\left\{\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)^2\log\left(\frac{P_{(j-k)\delta}}{P_{0\delta}}\right)^2\right]+\frac{36b_1}{81}\left(\mathbb{E}\left[\log\left(\frac{P_{j\delta}}{P_{0\delta}}\right)^2\right]+\mathbb{E}\left[\log\left(\frac{P_{(j-k)\delta}}{P_{0\delta}}\right)^2\right]\right)+\frac{16b_1^2}{81}\right\}\\
&= 2\sum_{j=2}^{M}\sum_{k=1}^{j-1}\left\{\mathbb{E}\left[\left(\int_0^{j\delta}\sigma\,dW_s\right)^2\left(\int_0^{(j-k)\delta}\sigma\,dW_s\right)^2\right]+\frac{36b_1}{81}\mathbb{E}\left[\left(\int_0^{j\delta}\sigma\,dW_s\right)^2+\left(\int_0^{(j-k)\delta}\sigma\,dW_s\right)^2\right]+\frac{16b_1^2}{81}\right\}\\
&= 2\sum_{j=2}^{M}\sum_{k=1}^{j-1}\left\{\frac{(j-k)\,(3j-2k)}{M^2}\sigma^4+\frac{36b_1}{81}\left(\frac{j\sigma^2}{M}+\frac{(j-k)\,\sigma^2}{M}\right)+\frac{16b_1^2}{81}\right\}\\
&= \frac{(M-1)\,(M+1)\,(7M+2)}{12M}\sigma^4+\frac{4\,(M-1)\,(M+1)\,b_1}{9}\sigma^2+\frac{16\,(M-1)\,Mb_1^2}{81}\,, \qquad (69)
\end{aligned}
$$

where we have used Definition A.3 and Lemmas A.5 and A.9 and $\delta=\frac{1}{M}$. We substitute Eq. (68) and (69) into Eq. (67).

$$
\begin{aligned}
\mathbb{E}\left[\widetilde{GV}_i^2\right] &= \frac{4}{(M+1)^2}\left[\frac{(M+1)\left(7M^2+7M+4\right)}{12M}\sigma^4+\frac{4\,(M+1)\,(M+2)\,b_1}{9}\sigma^2+\frac{4\,(4M+5)\,Mb_1^2}{81}\right]\\
&= \frac{7M^2+7M+4}{3M\,(M+1)}\sigma^4+\frac{16\,(M+2)\,b_1}{9\,(M+1)}+\frac{16M\,(4M+5)\,b_1^2}{81\,(M+1)^2} \qquad (70)\\
\mathrm{Var}\left(\widetilde{GV}_i\right) &= \mathbb{E}\left[\widetilde{GV}_i^2\right]-\left(\mathbb{E}\left[\widetilde{GV}_i\right]\right)^2\\
&= \frac{7M^2+7M+4}{3M\,(M+1)}\sigma^4+\frac{16\,(M+2)\,b_1}{9\,(M+1)}+\frac{16M\,(4M+5)\,b_1^2}{81\,(M+1)^2}-\left(\sigma^2+\frac{8Mb_1}{9\,(M+1)}\right)^2\\
&= \frac{4\left(M^2+M+1\right)}{3M\,(M+1)}\sigma^4+\frac{32b_1}{9\,(M+1)}\sigma^2+\frac{80Mb_1^2}{81\,(M+1)^2}\,, \qquad (71)
\end{aligned}
$$

where we have used Eq. (66) and (70).

Finally, we can evaluate the limit as M tends to infinity.

$$\begin{aligned}
\lim_{M\to\infty} \operatorname{Var}\left(\widetilde{GV}_i\right) &= \lim_{M\to\infty}\left[\frac{4\left(M^2+M+1\right)}{3M\left(M+1\right)}\sigma^4 + \frac{32b_1}{9\left(M+1\right)}\sigma^2 + \frac{80Mb_1^2}{81\left(M+1\right)^2}\right] \\
&= \lim_{M\to\infty} \operatorname{Var}\left(\widehat{GV}_i\right) \\
&= \frac{4\sigma^4}{3}
\end{aligned} \tag{72}$$

where we have used Eq. (71) and Eq. (33). Note that the first term in Eq. (72) is equal to $\operatorname{Var}\left(\widehat{GV}_i\right)$, the second term has M in the denominator so it tends to zero and in the third term, $\frac{M}{(M+1)^2} \to 0$, as $M \to \infty$.

B Simulation Appendix

The estimators considered are detailed in Section 2.2. Extensive analysis of the performance of each of the estimators under Geometric Brownian Motion is provided in Section 3.4. The inferences made on the three other price or volatility processes are extremely similar. This leads to the conclusion that market microstructures are a more important determinant of the behaviour of each of the estimators than the introduction of different price or volatility processes.

The Merton (1976) Jump Diffusion process was simulated using,

$$dP_t = (\mu - \lambda k)\, P_t\, dt + \sigma_{mer} P_t\, dW_t + (y_t - 1)\, P_t\, dJ_t\,,$$

where $\mu = 0$ and $\sigma_{mer} = \sigma_{gbm}$, $\lambda = 4$ is the mean number of jumps per day, the size of a jump is y_t, $\log(y_t) \sim$ i.i.d. $\mathcal{N}\big(0.0018,\, 10^{-5}\big)$, $k = \mathbb{E}\,[y_t - 1]$, W_t is a Wiener process and dJ_t is a Poisson process such that the probability of one jump during time interval is $\lambda\, dt$.

This paper employed the following GARCH(1,1) specification for prices,

$$\begin{aligned} dP_t &= \mu P_t\, dt + \sigma_{GARCH,t} P_t\, dW_t\,, \\ \sigma^2_{GARCH,t} &= \alpha_0 + \alpha_1 a^2_{t-1} + \beta_1 \sigma^2_{GARCH,t-1}\,, \\ a_t &= \sigma_{GARCH,t}\, \varepsilon_t\,, \qquad\qquad \varepsilon_t \sim \mathcal{N}(0,\, 1)\;. \end{aligned}$$

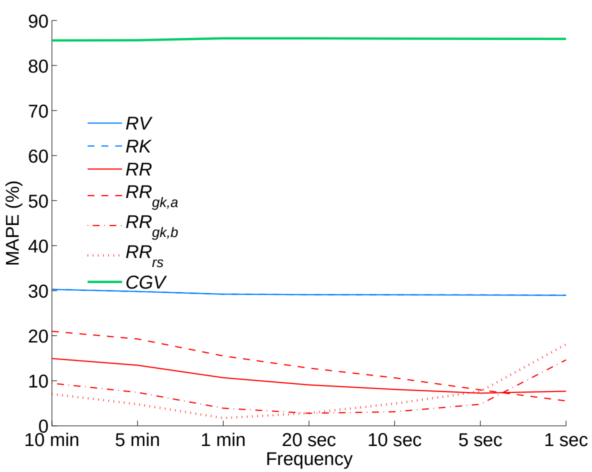
The drift parameter was again set to zero, however, the volatility was determined by the GARCH equation from a GARCH(1,1) model. The persistence parameter, β_1 was set to 0.999995, to allow for interday volatility effects; that is, the daily volatility would change between days and α_1 was set to satisfy the constraint that $\alpha_1 + \beta_1 < 1$. α_0 was chosen so that the long run and the initial volatilities were set equal to σ^2_{gbm}.

The final model simulated was the Heston (1993) Stochastic Volatility process. The model is stated by Wilmott (2007) as

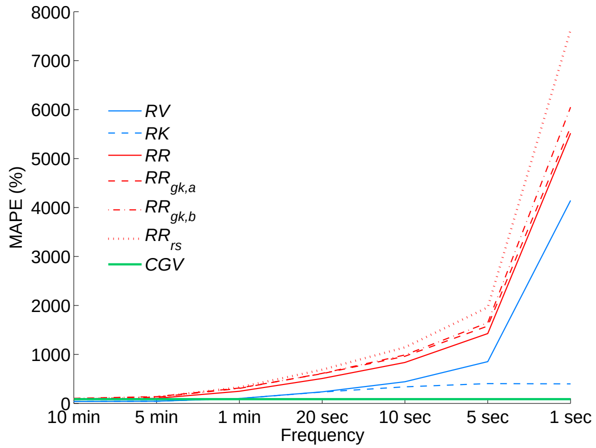
$$\begin{aligned} dP_t &= \mu P_t\, dt + \sigma_{Heston,t} P_t\, dW_t\,, \\ d\sigma^2_{Heston,t} &= \kappa \left(\theta - \sigma^2_{Heston,t}\right)\, dt + \xi \sigma_{Heston,t}\, dW^v_t\,, \\ dW_t\, dW^v_t &= \rho\, dt\,, \end{aligned}$$

The long-run volatility, θ, and the initial volatility is again set to σ^2_{gbm}. The speed of reversion, κ, set to 0.00006, to allow for interday variance effects. The Feller condition states that if $2\theta\kappa > \xi^2$ then the series σ^2_t will always be positive (Cox et al., 1985). To satisfy this condition, ξ was set to 5×10^{-8}. The third equation induces a correlation between the Wiener process in the price process and the Wiener process in the CIR volatility process; ρ was set to -0.7.

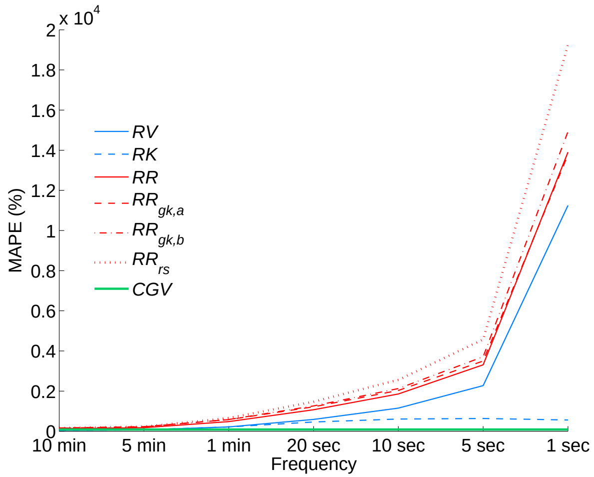
The process which is most different is the Merton Jump Diffusion Process. The impact of the introduction of this price process is most clear in Figures 4(a) and 4(d); the MAPE of realised volatility does not improve with frequency. In addition, the behaviour of the realised range estimators is different to that previously seen. The introduction of jumps also increases the MAPE of each estimator by a fair margin; the error of generalised variance increases by approximately 15% in the ideal case, compared to the Geometric Brownian Motion results. The generalised variance estimator again, however, shows its robustness to the increasing effects of market frictions when the frequency of observations is increased.



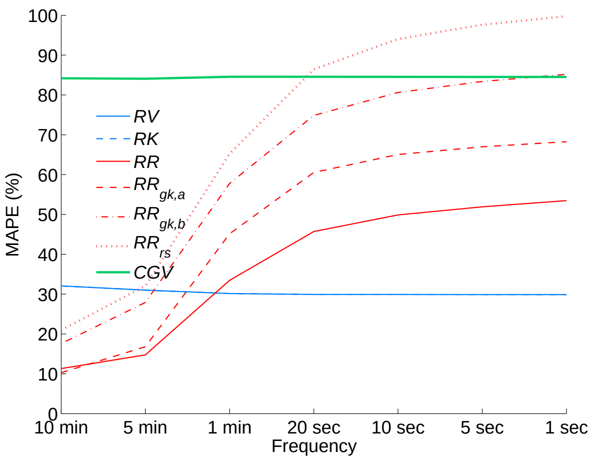
(a) No Market Microstructures



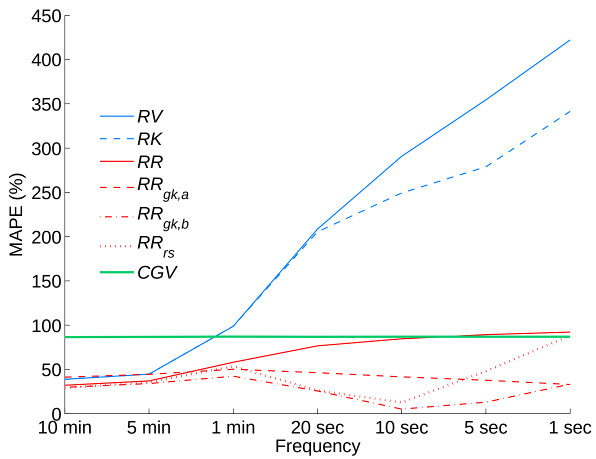
(b) Bid-Ask Bounce



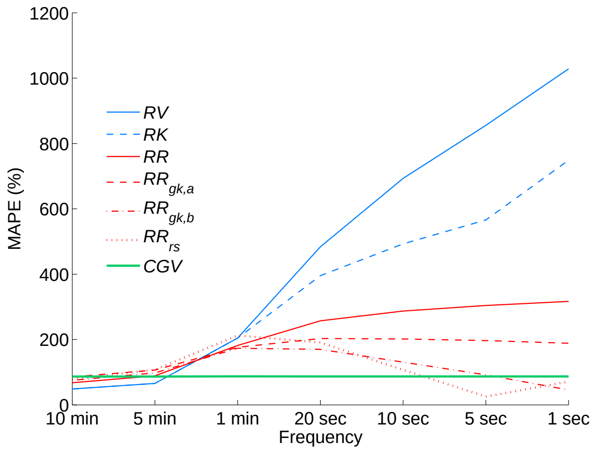
(c) A&F Bid-Ask Bounce



(d) Non-Synchronous Trading



(e) Bid-Ask Bounce and Non-Synchronous Trading



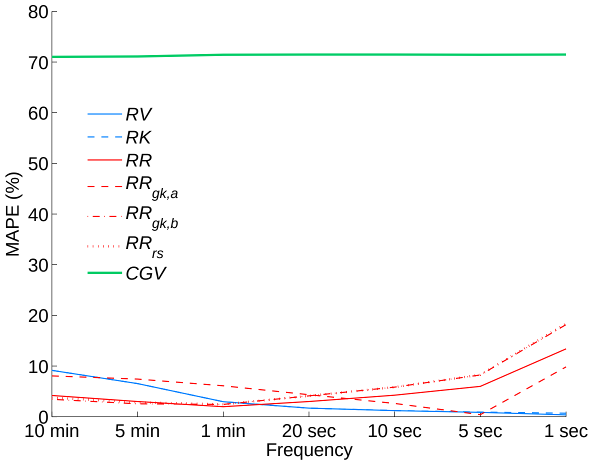
(f) A&F Bid-Ask Bounce and Non-Synchronous Trading

Figure 4: Accuracy in Simulation under Merton Jump Diffusion Process. The estimators are labelled as in Section 2.2.

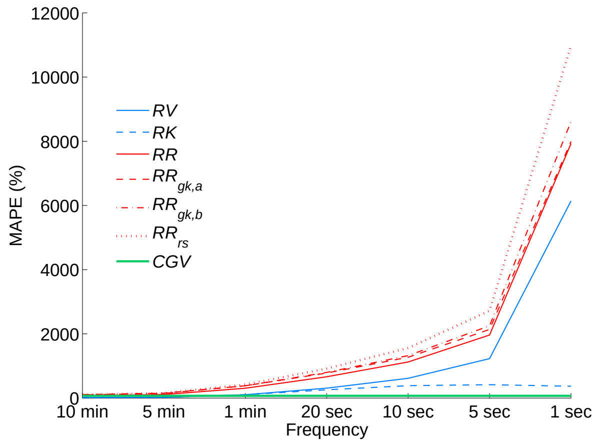
	(a) No Market Microstructures							(b) Bid-Ask Bounce							(c) A&F Bid-Ask Bounce						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	29.03	29.08	7.68	**5.53**	14.63	18.09	86.43	4143.10	398.04	5517.71	5642.52	6049.00	7626.86	**85.60**	11340.77	560.48	14012.67	13885.15	15044.97	19396.44	**85.38**
5 sec	29.11	29.13	7.28	8.00	**4.77**	7.62	86.45	852.85	403.54	1425.92	1580.33	1647.40	1961.34	**85.50**	2291.66	638.43	3334.47	3522.58	3737.11	4606.21	**85.54**
10 sec	29.15	29.14	8.11	10.68	**3.13**	4.93	86.48	441.68	337.29	836.18	963.28	988.74	1144.06	**85.68**	1160.40	607.54	1868.95	2044.44	2142.79	2576.13	**85.45**
20 sec	29.19	29.19	9.10	12.83	**2.79**	2.82	86.54	235.71	231.76	507.37	607.13	612.41	688.52	**85.51**	595.11	463.17	1079.44	1225.77	1266.58	1481.48	**85.51**
1 min	29.28	29.28	10.71	15.54	3.92	**1.72**	86.57	98.56	98.56	247.48	312.62	304.98	328.09	**85.61**	218.41	218.41	484.87	584.40	587.79	657.06	**85.68**
5 min	29.87	29.87	13.48	19.32	7.45	**4.80**	86.14	**44.01**	**44.01**	103.07	138.07	125.91	127.94	85.53	**68.20**	**68.20**	178.40	231.56	221.09	232.30	85.50
10 min	30.37	30.37	14.97	21.00	9.47	**7.06**	86.12	**38.42**	**38.42**	75.92	103.30	90.76	90.32	85.64	**49.41**	**49.41**	124.23	164.92	153.12	157.32	85.68

	(d) Non-Synchronous Trading							(e) Bid-Ask Bounce and Non-Synchronous Trading							(f) A&F Bid-Ask Bounce and Non-Synchronous Trading						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	29.88	**29.87**	53.49	68.27	85.18	99.77	84.68	422.20	341.41	92.19	**32.92**	33.34	88.23	86.77	1036.80	753.62	319.63	190.80	**46.82**	71.16	86.92
5 sec	**29.88**	29.88	51.91	67.00	83.37	97.62	84.68	354.48	278.78	89.22	37.75	**12.95**	47.91	86.77	862.39	569.32	307.15	199.12	93.29	**26.42**	86.98
10 sec	**29.90**	29.90	49.88	65.05	80.61	94.02	84.70	290.64	249.19	84.53	41.57	**5.14**	12.66	86.80	698.72	495.07	290.00	203.97	132.30	108.70	**86.91**
20 sec	**29.92**	**29.92**	45.73	60.54	74.86	86.44	84.72	208.52	205.37	76.57	46.23	**25.64**	26.39	86.73	487.77	398.33	259.93	205.11	171.98	193.02	**86.99**
1 min	**30.13**	**30.13**	33.46	45.14	57.79	65.24	84.72	98.78	98.78	58.02	50.15	**42.28**	53.44	86.92	206.06	206.06	183.62	177.96	174.96	214.69	**87.19**
5 min	30.98	30.98	**14.75**	16.81	27.93	32.05	84.23	44.70	44.70	36.96	44.46	**34.00**	35.65	86.64	**66.17**	**66.17**	89.55	107.28	98.59	108.29	86.79
10 min	32.06	32.06	11.31	**10.28**	17.62	20.99	84.34	38.84	38.84	32.10	41.20	29.73	**29.15**	86.38	**48.96**	**48.96**	68.08	86.15	75.51	79.12	86.52

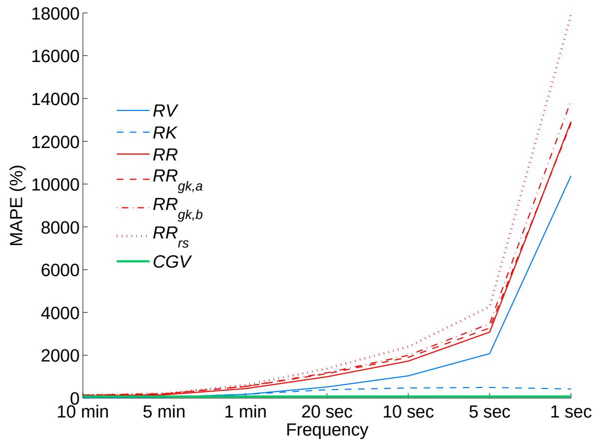
Table 5: Accuracy in Simulation under Merton Jump Diffusion Process. The estimators are labelled as in Section 2.2. This table gives the MAPE (%) values observed in Figure 4. The most accurate estimator for each frequency is in bold.



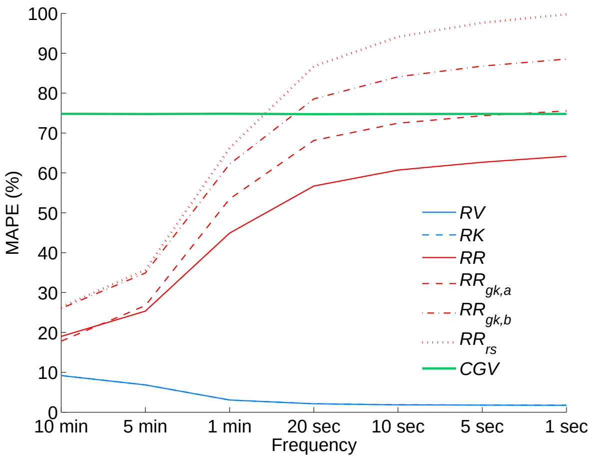
(a) No Market Microstructures



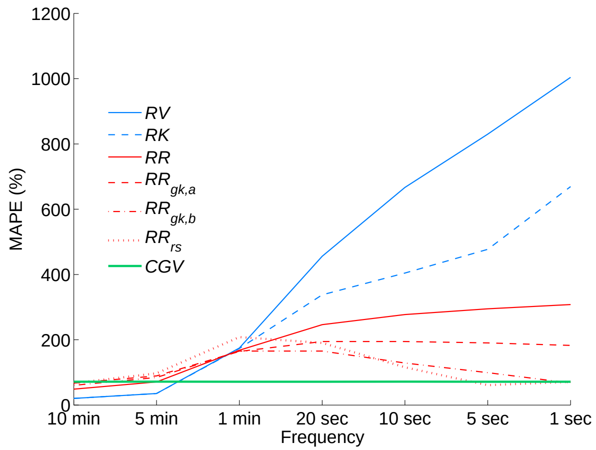
(b) Bid-Ask Bounce



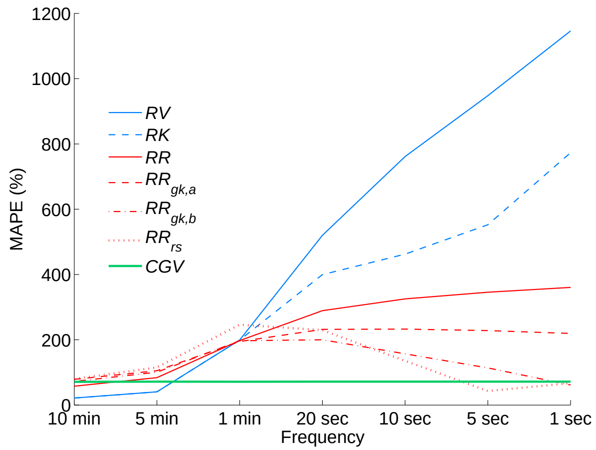
(c) A&F Bid-Ask Bounce



(d) Non-Synchronous Trading



(e) Bid-Ask Bounce and Non-Synchronous Trading



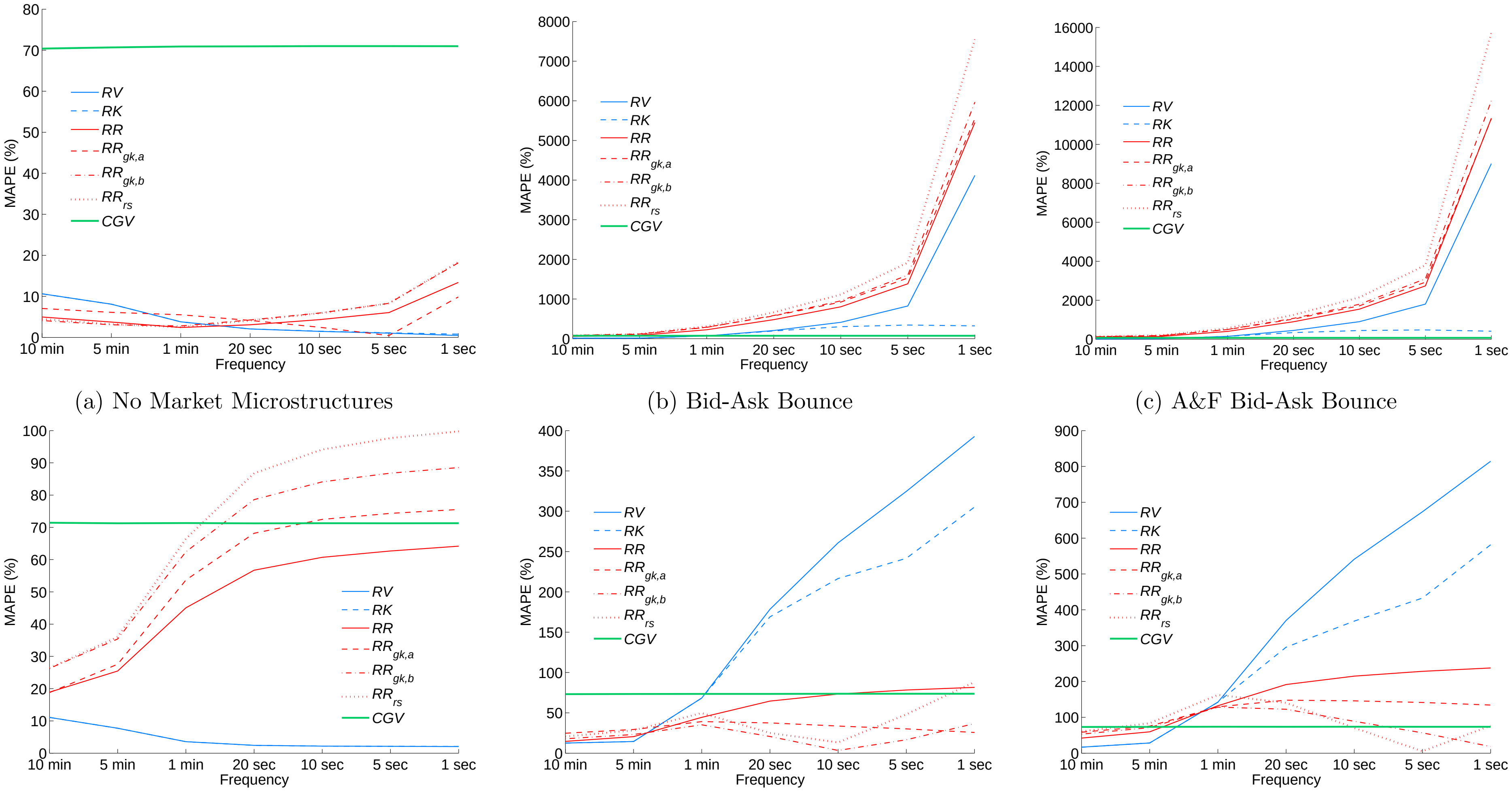
(f) A&F Bid-Ask Bounce and Non-Synchronous Trading

Figure 5: Accuracy in Simulation under GARCH(1,1) Volatility Process. The estimators are labelled as in Section 2.2.

	(a) No Market Microstructures							(b) Bid-Ask Bounce							(c) A&F Bid-Ask Bounce						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	**0.36**	0.68	13.39	9.83	18.18	18.36	71.36	6277.95	369.89	8090.79	8154.50	8791.32	11206.37	**68.88**	10391.22	421.94	12930.81	12844.96	13911.78	17911.71	**68.73**
5 sec	0.86	0.92	5.99	**0.41**	8.23	8.27	71.33	1255.11	418.66	2001.60	2174.66	2290.02	2774.39	**68.87**	2077.99	495.84	3084.63	3273.82	3473.11	4274.42	**68.75**
10 sec	**1.20**	1.22	4.24	2.61	5.84	5.84	71.36	627.72	387.10	1143.96	1292.08	1343.36	1585.52	**68.95**	1038.56	468.72	1728.40	1901.97	1994.86	2396.05	**68.63**
20 sec	**1.69**	**1.69**	3.00	4.35	4.13	4.13	71.37	313.76	260.56	673.14	792.03	812.01	933.25	**68.96**	519.57	388.67	995.69	1139.54	1179.63	1380.17	**68.85**
1 min	2.95	2.95	**1.99**	6.10	2.43	2.45	71.33	104.78	104.78	308.60	387.78	387.41	427.87	**68.96**	173.57	173.57	441.38	539.00	544.85	611.49	**68.55**
5 min	6.51	6.51	2.99	7.42	**2.55**	2.85	70.95	**20.96**	**20.96**	110.95	153.45	145.83	153.95	68.53	**34.72**	**34.72**	152.27	204.32	197.73	211.08	68.30
10 min	9.13	9.13	4.18	8.05	**3.44**	3.77	70.86	**12.65**	**12.65**	73.73	106.65	98.49	102.68	68.57	**18.20**	**18.20**	99.90	139.80	131.94	138.80	68.48

	(d) Non-Synchronous Trading							(e) Bid-Ask Bounce and Non-Synchronous Trading							(f) A&F Bid-Ask Bounce and Non-Synchronous Trading						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	**1.75**	1.77	64.17	75.56	88.55	99.77	74.79	1025.65	680.95	315.84	188.26	**69.29**	70.62	71.38	1145.93	772.99	360.40	219.22	**61.55**	67.26	71.41
5 sec	**1.78**	1.80	62.68	74.35	86.79	97.66	74.79	847.75	483.95	302.54	196.12	102.03	**62.28**	71.40	947.52	551.97	345.47	227.95	113.52	**42.97**	71.46
10 sec	**1.85**	1.87	60.72	72.48	84.11	94.11	74.77	681.16	407.75	284.49	200.29	132.76	119.09	**71.44**	761.16	462.03	325.04	232.42	156.76	135.55	**71.54**
20 sec	**2.11**	**2.11**	56.71	68.15	78.55	86.73	74.73	465.77	340.34	252.69	200.00	170.43	196.01	**71.40**	519.83	399.43	288.97	231.65	199.84	229.45	**71.54**
1 min	**3.06**	**3.06**	44.91	53.48	62.21	66.26	74.83	178.64	176.81	172.16	168.61	169.66	213.77	**71.38**	199.23	199.11	197.42	194.63	196.73	246.30	**71.38**
5 min	**6.84**	**6.84**	25.35	26.76	34.90	35.68	74.75	**35.98**	**35.98**	72.34	91.26	86.43	99.46	71.50	**40.31**	**40.31**	83.87	105.20	100.71	115.30	71.44
10 min	**9.18**	**9.18**	19.02	17.84	26.04	26.38	74.81	**20.54**	**20.54**	49.60	68.30	61.60	68.37	71.18	**21.37**	**21.37**	57.78	78.78	72.25	79.91	70.97

Table 6: Accuracy in Simulation under GARCH(1,1) Volatility Process. The estimators are labelled as in Section 2.2. This table gives the MAPE (%) values observed in Figure 5. The most accurate estimator for each frequency is in bold.



(a) No Market Microstructures

(b) Bid-Ask Bounce

(c) A&F Bid-Ask Bounce

(d) Non-Synchronous Trading

(e) Bid-Ask Bounce and Non-Synchronous Trading

(f) A&F Bid-Ask Bounce and Non-Synchronous Trading

Figure 6: Accuracy in Simulation under Heston Stochastic Volatility process. The estimators are labelled as in Section 2.2.

	(a) No Market Microstructures							(b) Bid-Ask Bounce							(c) A&F Bid-Ask Bounce						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	**0.46**	0.78	13.39	9.84	18.18	18.36	71.48	4119.31	322.36	5451.18	5554.48	5966.49	7549.95	**72.72**	8905.18	409.73	11216.80	11197.41	12109.66	15539.93	**72.73**
5 sec	1.03	1.11	6.04	**0.44**	8.30	8.30	71.49	824.01	343.23	1384.86	1530.16	1601.73	1919.51	**72.82**	1781.04	474.80	2712.03	2900.94	3071.51	3762.37	**72.78**
10 sec	**1.46**	1.51	4.31	2.48	5.93	5.87	71.48	412.04	301.32	802.50	923.45	953.32	1112.21	**72.83**	890.50	446.16	1531.54	1699.61	1779.38	2127.72	**72.60**
20 sec	**2.06**	**2.06**	3.08	4.13	4.25	4.12	71.44	206.12	193.13	479.13	575.00	584.60	664.21	**72.85**	445.32	340.42	890.25	1027.26	1062.16	1238.95	**72.46**
1 min	3.80	3.80	**2.41**	5.52	2.79	2.49	71.41	**68.78**	**68.78**	224.87	288.41	285.09	311.92	72.72	148.86	148.86	401.45	492.81	498.99	561.75	**72.86**
5 min	8.08	8.08	3.74	6.11	**3.07**	3.23	71.17	**14.15**	**14.15**	83.68	117.77	110.85	117.48	73.17	**29.31**	**29.31**	143.84	191.42	188.26	206.04	72.94
10 min	10.62	10.62	4.97	7.06	**4.09**	4.46	70.89	**12.09**	**12.09**	56.77	83.49	76.36	80.31	72.98	**16.61**	**16.61**	96.98	133.66	129.07	140.01	72.59

	(d) Non-Synchronous Trading							(e) Bid-Ask Bounce and Non-Synchronous Trading							(f) A&F Bid-Ask Bounce and Non-Synchronous Trading						
	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$	$\widehat{RV}$	$\widehat{RK}$	$\widehat{RR}$	$\widehat{RR}_{gk,a}$	$\widehat{RR}_{gk,b}$	$\widehat{RR}_{rs}$	$\widehat{GV}$
1 sec	**2.06**	2.08	64.16	75.55	88.55	99.77	71.54	392.80	305.29	81.66	**25.71**	36.71	88.25	73.88	805.22	576.07	234.35	131.79	**17.16**	76.74	73.82
5 sec	2.12	**2.12**	62.67	74.35	86.80	97.67	71.52	324.68	241.65	78.30	30.17	**16.62**	48.07	73.90	665.89	429.00	224.92	139.09	55.05	**5.12**	73.81
10 sec	**2.18**	2.18	60.71	72.48	84.12	94.14	71.53	260.90	216.58	73.28	33.63	**3.37**	13.27	73.86	534.92	365.92	211.74	143.53	87.06	**68.71**	73.91
20 sec	**2.42**	**2.42**	56.72	68.16	78.59	86.80	71.48	178.41	169.02	64.64	37.53	**20.72**	25.08	73.76	366.03	292.57	188.70	145.56	120.24	137.90	**73.81**
1 min	**3.56**	**3.56**	45.05	53.68	62.42	66.48	71.58	68.41	68.41	44.40	39.14	**35.13**	49.78	73.77	140.60	140.60	130.50	127.35	126.61	160.23	**73.91**
5 min	**7.71**	**7.71**	25.46	27.58	35.41	36.04	71.49	**14.42**	**14.42**	20.45	29.39	23.18	28.23	73.60	**28.07**	**28.07**	58.63	74.62	70.46	82.55	73.59
10 min	**11.11**	**11.11**	18.91	18.80	26.30	26.38	71.66	**12.58**	**12.58**	14.66	24.74	17.80	20.64	73.40	**16.53**	**16.53**	41.61	57.72	52.45	59.96	73.44

Table 7: Accuracy in Simulation under Heston Stochastic Volatility process. The estimators are labelled as in Section 2.2. This table gives the MAPE (%) values observed in Figure 6. The most accurate estimator for each frequency is in bold.

C Empirical Appendix

Table 8: List of Companies Included in the Empirical Study. This consists of companies listed on the ASX 20 as of 13/05/14. The GICS Industry Group gives an indication of the sector for each stock while the First Date is the date of the first data available. The date range of the empirical data set was from 01/01/10 until 31/12/13, a period of 999 working days. Highlighted in grey are the companies for which data within this date range in unavailable. N/A means that the SIRCA Australian Equities database could not find the data.

Code	Company Name	GICS Industry Group	First Date
AMP	AMP (Ltd.)	Insurance	6/05/1998
ANZ	Australia and New Zealand Banking Group (Ltd.)	Banks	1/06/1981
BHP	BHP Billiton (Ltd.)	Materials	2/01/1980
BXB	Brambles (Ltd.)	Commercial & Professional Services	30/10/2006
CBA	Commonwealth Bank of Australia	Banks	12/09/1991
CSL	CSL (Ltd.)	Pharmaceuticals, Biotechnology & Life Sciences	1/11/2007
MQG	Macquarie Group (Ltd.)	Diversified Financials	11/10/2007
NAB	National Australia Bank (Ltd.)	Banks	17/09/1981
NCM	Newcrest Mining (Ltd.)	Materials	15/10/1992
ORG	Origin Energy (Ltd.)	Energy	9/03/2000
QBE	QBE Insurance Group (Ltd.)	Insurance	2/01/1980
RIO	Rio Tinto (Ltd.)	Materials	4/06/1997
STO	Santos (Ltd.)	Energy	2/01/1980
SUN	Suncorp Group (Ltd.)	Insurance	12/01/2011
TLS	Telstra Corporation (Ltd.)	Telecommunication Services	20/05/2008
WBC	Westpac Banking Corporation	Banks	17/09/1981
WDC	Westfield Group	Real Estate	N/A
WES	Wesfarmers (Ltd.)	Food & Staples Retailing	08/11/2013
WOW	Woolworths (Ltd.)	Food & Staples Retailing	23/07/1993
WPL	Woodside Petroleum (Ltd.)	Energy	2/01/1980

Table 9: Unconditional Distributions of Annualised Percentage Volatility Estimates for AMP - CSL

	AMP			ANZ			BHP			BXB			CBA			CSL		
	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$
Mean																		
1 sec	44.44	4.89	2.93	6.85	0.82	2.24	4.24	0.57	1.56	21.47	2.69	4.05	3.53	0.50	1.84	5.07	0.76	2.55
10 sec	31.17	14.84	2.94	4.82	2.44	2.23	2.85	1.55	1.55	15.90	7.54	4.05	2.79	1.44	1.84	4.35	2.15	2.55
1 min	13.67	10.51	2.91	3.17	2.42	2.20	2.02	1.60	1.54	9.10	6.29	4.00	2.27	1.68	1.81	3.63	2.53	2.51
5 min	5.63	6.07	2.73	2.54	2.41	2.08	1.71	1.65	1.49	5.45	5.03	3.88	1.96	1.82	1.73	2.91	2.76	2.37
Variance																		
1 sec	416.96	17.18	15.43	18.58	0.39	15.35	6.78	0.22	7.24	110.30	4.20	49.77	7.30	0.18	13.95	9.35	0.29	10.38
10 sec	158.70	47.48	15.58	7.70	2.45	15.25	2.30	0.92	7.26	52.27	13.36	49.92	3.38	1.16	13.79	5.96	1.63	10.41
1 min	26.10	18.28	15.77	4.69	2.80	15.16	1.70	1.06	7.13	17.93	8.96	53.60	3.03	1.66	13.73	4.30	2.24	10.15
5 min	11.50	9.26	13.79	6.09	3.78	13.49	1.99	1.42	6.83	14.99	8.98	58.62	3.62	2.31	12.73	3.75	2.67	9.72
Skewness																		
1 sec	1.70	2.11	4.74	4.24	5.62	10.21	5.35	8.00	7.90	1.25	1.87	9.97	7.38	8.47	11.92	3.66	3.09	3.00
10 sec	1.39	1.72	4.83	3.80	5.05	10.27	3.66	5.41	7.87	1.13	1.29	9.96	4.52	6.72	11.91	3.52	3.56	3.02
1 min	2.07	2.34	5.07	4.28	4.99	10.39	3.54	4.69	7.64	1.97	2.11	11.91	4.83	5.54	12.04	2.76	3.23	3.05
5 min	4.51	4.20	5.08	7.11	5.70	10.65	4.58	4.94	7.82	4.91	3.79	13.40	7.15	5.88	12.60	2.48	2.68	3.42
Kurtosis																		
1 sec	8.93	9.50	40.37	40.74	64.43	166.78	61.12	120.34	107.27	5.79	8.15	171.34	110.17	136.44	220.27	30.40	20.44	15.73
10 sec	7.33	9.44	42.06	36.65	57.37	168.65	30.93	65.01	106.03	5.80	6.33	170.59	49.42	95.77	222.71	31.56	31.48	15.81
1 min	14.57	17.64	46.04	41.34	54.22	171.22	26.86	51.13	99.25	13.84	14.38	232.92	52.97	67.97	224.06	19.17	26.36	16.49
5 min	47.35	39.64	46.31	87.51	63.65	185.00	43.95	55.92	101.66	52.82	34.58	277.93	99.38	73.32	248.26	12.78	17.58	21.11
Minimum																		
1 sec	10.82	0.42	0.06	1.49	0.14	0.02	1.11	0.12	0.01	4.92	0.27	0.03	0.67	0.07	0.02	1.11	0.09	0.03
10 sec	8.62	3.66	0.06	1.16	0.61	0.02	0.73	0.40	0.01	3.97	1.77	0.04	0.51	0.28	0.04	0.90	0.46	0.02
1 min	4.43	3.12	0.07	0.71	0.63	0.02	0.44	0.41	0.02	1.93	1.45	0.04	0.39	0.31	0.03	0.54	0.45	0.02
5 min	1.10	2.02	0.08	0.31	0.52	0.02	0.24	0.39	0.01	0.52	0.87	0.02	0.24	0.32	0.02	0.25	0.37	0.02
Maximum																		
1 sec	199.61	30.36	48.73	63.70	10.24	78.26	43.31	9.01	48.37	82.92	16.49	145.11	51.77	8.59	80.67	42.04	5.75	28.24
10 sec	122.87	71.17	48.75	41.03	25.62	78.28	21.43	16.37	48.21	64.38	32.55	145.12	29.29	19.99	80.86	35.09	18.18	27.35
1 min	64.03	55.91	48.79	31.61	26.47	78.27	16.79	16.53	46.80	50.19	35.14	163.28	27.69	21.89	80.47	25.35	20.21	28.75
5 min	53.18	46.16	49.01	40.25	30.61	76.39	20.77	19.30	45.79	60.23	42.94	178.20	34.65	26.00	80.31	16.29	19.00	31.59

Table 10: Unconditional Distributions of Annualised Percentage Volatility Estimates for MQG - RIO

	MQG			NAB			NCM			ORG			QBE			RIO		
	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$
Mean																		
1 sec	6.88	0.94	4.59	6.55	0.81	2.41	8.08	1.18	5.42	10.17	1.15	3.50	11.36	1.41	3.89	3.79	0.60	2.39
10 sec	5.68	2.79	4.59	4.74	2.40	2.40	6.35	3.25	5.38	8.16	3.79	3.48	8.33	4.12	3.87	2.93	1.57	2.37
1 min	4.63	3.18	4.54	3.30	2.48	2.35	5.13	3.66	5.24	5.89	3.88	3.42	5.41	4.00	3.83	2.37	1.80	2.34
5 min	4.09	3.56	4.30	2.69	2.53	2.21	4.80	4.13	4.91	4.53	3.95	3.24	4.21	3.92	3.75	2.17	1.98	2.25
Variance																		
1 sec	29.34	0.78	59.48	19.63	0.90	20.49	46.83	2.56	141.89	29.92	0.74	32.47	59.60	2.18	34.48	6.35	0.20	24.79
10 sec	15.56	4.53	59.31	7.44	2.58	20.35	22.57	8.22	139.51	18.65	4.76	32.28	23.84	8.13	33.72	2.89	1.03	24.66
1 min	11.06	5.59	57.36	5.06	3.13	19.48	17.34	9.94	134.04	13.95	6.47	30.19	12.01	8.78	32.35	2.38	1.41	24.38
5 min	11.54	7.47	53.73	5.87	3.93	17.76	28.70	15.18	125.16	18.49	9.77	28.93	17.02	11.52	49.08	3.04	1.98	22.46
Skewness																		
1 sec	4.17	4.79	7.57	6.22	16.09	11.49	3.91	10.04	8.94	4.09	3.02	9.72	5.16	7.58	5.47	4.03	4.32	11.93
10 sec	3.37	3.77	7.62	4.12	6.05	11.50	3.12	4.50	8.99	4.47	4.36	9.80	5.52	5.26	5.37	2.55	3.65	11.94
1 min	3.71	3.78	7.21	3.94	5.15	11.64	4.29	4.61	9.23	5.33	5.51	9.88	4.27	5.46	5.22	2.98	3.63	11.79
5 min	3.94	3.93	7.64	5.59	4.94	11.92	6.77	5.43	9.90	10.48	7.41	11.24	7.95	6.08	10.80	3.83	3.59	12.01
Kurtosis																		
1 sec	32.15	42.02	97.10	75.73	358.86	202.78	27.81	154.55	109.27	39.54	18.73	162.21	45.54	98.68	48.40	35.34	40.48	219.54
10 sec	24.65	28.71	98.65	43.58	76.85	203.72	18.77	35.68	111.21	47.69	44.62	164.25	62.48	49.41	46.86	17.25	31.98	220.15
1 min	29.65	31.58	87.86	34.72	56.46	206.87	34.96	38.13	117.52	64.50	67.08	171.28	34.81	51.03	45.65	22.92	33.25	214.19
5 min	31.54	33.32	97.19	59.79	50.64	216.22	77.37	52.78	132.21	199.56	113.42	218.42	108.76	58.60	186.68	35.54	32.42	224.97
Minimum																		
1 sec	1.10	0.06	0.02	1.57	0.14	0.03	1.89	0.14	0.04	2.39	0.09	0.06	3.33	0.22	0.05	0.80	0.07	0.02
10 sec	1.00	0.39	0.02	1.33	0.61	0.03	1.39	0.71	0.04	2.51	0.84	0.06	2.89	1.21	0.03	0.74	0.34	0.02
1 min	0.77	0.45	0.03	0.74	0.57	0.02	0.89	0.69	0.05	1.62	1.02	0.05	1.30	0.96	0.05	0.43	0.34	0.02
5 min	0.36	0.52	0.02	0.39	0.52	0.02	0.45	0.63	0.04	0.83	0.82	0.04	0.55	0.80	0.03	0.25	0.32	0.02
Maximum																		
1 sec	67.76	11.06	127.92	75.46	23.59	93.94	78.20	30.94	191.58	83.06	8.88	114.64	97.43	25.57	72.13	34.09	6.36	107.72
10 sec	49.90	26.93	128.46	42.53	28.07	93.98	44.90	33.35	192.27	68.97	33.93	114.64	79.01	38.91	72.23	20.17	13.94	107.57
1 min	43.25	32.09	119.60	30.97	28.22	92.04	48.53	39.07	192.83	63.23	43.23	112.78	41.95	40.85	73.46	19.70	16.75	106.05
5 min	44.28	37.33	118.70	36.78	30.06	89.47	85.65	55.83	192.19	94.18	60.04	117.94	74.07	43.08	144.54	24.79	19.73	103.47

Table 11: Unconditional Distributions of Annualised Percentage Volatility Estimates for STO - WPL

	STO			TLS			WBC			WOW			WPL		
	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$	$\widehat{RV}_i$	$\widehat{RR}_i$	$\widehat{GV}_i$
Mean															
1 sec	11.63	1.40	3.93	159.18	19.54	1.82	7.11	0.87	2.73	4.73	0.62	1.54	5.44	0.73	2.92
10 sec	9.16	4.32	3.91	91.35	49.64	1.83	5.01	2.55	2.73	3.68	1.81	1.55	4.39	2.18	2.89
1 min	6.40	4.37	3.82	24.75	23.92	1.77	3.41	2.58	2.68	2.55	1.86	1.52	3.48	2.49	2.83
5 min	5.03	4.45	3.65	6.23	7.51	1.65	2.78	2.62	2.50	1.91	1.85	1.48	3.12	2.75	2.67
Variance															
1 sec	42.66	1.21	41.64	10071.33	392.72	7.45	21.86	0.53	21.51	6.80	0.22	5.30	23.00	0.51	48.64
10 sec	24.36	7.12	41.30	3150.01	940.18	7.60	7.83	2.79	21.37	3.53	1.06	5.14	12.50	3.53	48.03
1 min	15.00	8.77	39.74	227.41	134.37	7.39	5.20	3.11	20.88	2.19	1.27	5.13	10.28	4.53	48.59
5 min	18.91	12.04	38.23	15.43	10.00	7.42	6.08	4.13	16.78	2.81	1.58	4.99	14.83	6.97	46.93
Skewness															
1 sec	5.31	4.57	8.84	0.58	2.13	4.59	4.47	6.01	9.17	4.62	3.97	5.05	9.25	9.07	15.20
10 sec	5.55	6.41	8.78	0.38	0.43	4.53	3.18	4.80	9.07	4.38	4.73	4.86	8.65	8.41	15.65
1 min	6.55	7.72	9.15	0.39	0.02	4.81	3.54	4.13	8.85	4.04	4.88	4.99	10.37	7.94	15.31
5 min	8.82	8.75	9.63	1.12	1.03	5.32	4.69	4.28	7.10	7.69	4.70	5.63	11.84	9.91	15.96
Kurtosis															
1 sec	64.36	39.63	133.76	2.81	10.73	33.27	44.06	71.93	141.37	49.74	28.25	39.95	136.14	142.24	329.81
10 sec	70.41	86.48	133.77	2.07	2.34	31.85	26.90	51.73	138.23	48.37	52.26	37.58	125.41	117.92	346.83
1 min	95.06	121.55	143.51	1.88	1.80	35.47	29.71	38.88	129.95	38.86	56.34	39.27	170.11	108.69	330.92
5 min	151.25	150.86	157.32	4.84	6.16	41.29	41.73	37.95	81.59	108.99	46.12	51.51	196.51	161.98	353.32
Minimum															
1 sec	3.19	0.17	0.05	15.91	0.65	0.05	1.65	0.14	0.04	1.45	0.11	0.03	1.41	0.09	0.04
10 sec	2.83	1.31	0.05	8.28	5.15	0.07	1.25	0.63	0.04	1.11	0.48	0.04	1.05	0.50	0.04
1 min	1.88	1.37	0.03	2.81	3.34	0.05	0.78	0.63	0.04	0.63	0.51	0.03	0.63	0.43	0.04
5 min	0.90	1.06	0.05	0.59	1.81	0.03	0.37	0.50	0.05	0.30	0.38	0.02	0.28	0.35	0.03
Maximum															
1 sec	111.82	14.69	122.39	544.96	170.63	29.52	69.42	12.24	88.39	41.84	5.07	26.58	87.89	13.93	167.68
10 sec	86.87	48.58	122.58	290.40	158.78	29.57	38.24	26.36	87.31	30.46	16.75	25.83	63.59	33.62	168.82
1 min	72.82	58.54	122.37	65.11	52.39	29.63	31.06	25.77	84.27	21.87	18.62	25.96	64.57	37.84	167.49
5 min	89.36	71.76	122.83	26.59	26.59	30.22	31.30	28.04	60.81	30.87	18.39	27.71	73.81	53.09	167.27

Table 12: Distributions of Standardised Daily Returns for AMP - CSL

	AMP				ANZ				BHP				BXB				CBA				CSL			
	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$
Mean																								
1 sec	-0.04	-0.05	-0.18	0.57	0.03	0.43	1.36	1.11	-0.01	0.18	0.71	1.05	0.03	0.15	0.59	0.71	0.03	0.66	1.83	1.44	0.07	0.72	1.63	0.87
10 sec	-0.04	-0.05	-0.08	0.59	0.03	0.53	0.74	1.08	-0.01	0.23	0.35	1.11	0.03	0.17	0.25	0.68	0.03	0.74	1.05	1.43	0.07	0.77	1.06	0.88
1 min	-0.04	-0.06	-0.09	0.55	0.03	0.80	0.83	0.97	-0.01	0.41	0.40	1.18	0.03	0.26	0.29	0.55	0.03	0.98	1.06	1.50	0.07	0.86	0.99	0.72
5 min	-0.04	0.03	-0.04	0.48	0.03	1.06	0.95	1.28	-0.01	0.68	0.57	1.39	0.03	0.41	0.36	0.49	0.03	1.20	1.12	1.68	0.07	0.98	0.95	0.74
Variance																								
1 sec	2.27	12.58	156.17	311.33	1.86	63.37	543.05	385.16	2.06	115.82	905.65	947.76	2.42	29.01	297.31	281.93	1.32	91.38	668.61	364.55	1.84	95.07	666.71	400.92
10 sec	2.27	17.52	37.29	314.36	1.86	93.08	178.79	383.65	2.06	178.55	323.99	963.45	2.42	38.27	81.95	281.31	1.32	119.72	227.52	367.40	1.84	112.27	222.76	402.47
1 min	2.27	38.60	49.66	308.06	1.86	150.97	186.53	419.32	2.06	280.34	330.16	970.48	2.42	64.44	93.07	285.56	1.32	158.02	203.24	371.31	1.84	138.12	190.23	409.63
5 min	2.27	93.43	82.23	343.26	1.86	201.55	197.11	423.87	2.06	361.03	341.14	1045.41	2.42	114.73	115.73	326.03	1.32	192.44	193.92	404.10	1.84	184.98	175.67	431.31
Skewness																								
1 sec	-1.13	-0.51	-0.27	-0.15	-0.12	-0.14	-0.08	-1.15	-0.12	-0.03	0.00	-0.37	0.03	0.07	0.10	0.12	-0.15	-0.13	-0.08	-0.12	0.04	0.21	0.17	-0.52
10 sec	-1.13	-0.48	-0.43	-0.10	-0.12	-0.21	-0.16	-0.93	-0.12	-0.04	-0.02	-0.28	0.03	0.05	0.07	0.12	-0.15	-0.15	-0.13	-0.13	0.04	0.15	0.18	-0.77
1 min	-1.13	-0.44	-0.37	-0.14	-0.12	-0.30	-0.20	-1.33	-0.12	-0.04	-0.03	-0.06	0.03	0.05	0.06	0.13	-0.15	-0.16	-0.15	-0.01	0.04	0.15	0.15	-0.61
5 min	-1.13	-0.21	-0.22	-0.42	-0.12	-0.29	-0.26	-0.80	-0.12	-0.02	0.00	-0.41	0.03	0.04	0.05	0.22	-0.15	-0.17	-0.17	0.10	0.04	-0.28	0.05	-0.24
Kurtosis																								
1 sec	11.87	6.33	5.61	5.47	4.98	4.00	3.61	14.38	3.55	2.70	2.66	9.67	4.08	3.55	4.29	5.12	4.18	3.38	3.24	6.76	4.71	3.86	3.54	12.29
10 sec	11.87	6.19	5.77	5.51	4.98	4.69	4.11	11.57	3.55	2.96	2.76	9.83	4.08	3.46	3.49	5.14	4.18	3.62	3.38	6.67	4.71	4.27	3.89	14.95
1 min	11.87	5.24	4.84	5.15	4.98	5.54	4.48	16.93	3.55	3.18	2.90	8.20	4.08	3.24	3.21	4.94	4.18	3.92	3.57	6.57	4.71	4.51	3.94	13.97
5 min	11.87	3.41	3.50	6.53	4.98	5.18	4.79	10.44	3.55	3.22	3.00	11.49	4.08	3.40	3.06	8.35	4.18	4.17	3.79	6.91	4.71	7.38	4.18	10.07
Jarque-Bera Test																								
1 sec	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.002	**0.135**	**0.085**	0.001	0.001	0.004	0.001	0.001	0.001	0.015	**0.161**	0.001	0.001	0.001	0.001	0.001
10 sec	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	**0.500**	**0.266**	0.001	0.001	0.014	0.008	0.001	0.001	0.001	0.016	0.001	0.001	0.001	0.001	0.001
1 min	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	**0.416**	**0.500**	0.001	0.001	**0.220**	**0.280**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
5 min	0.001	0.003	0.001	0.001	0.001	0.001	0.001	0.001	0.002	**0.344**	**0.500**	0.001	0.001	0.033	**0.500**	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Minimum																								
1 sec	-14.62	-26.52	-74.22	-92.38	-6.41	-38.71	-101.37	-196.60	-5.24	-32.04	-81.51	-252.03	-6.23	-19.69	-72.14	-73.22	-4.76	-34.78	-87.75	-94.74	-7.28	-42.70	-86.90	-171.03
10 sec	-14.62	-31.05	-43.23	-92.49	-6.41	-52.29	-65.75	-178.64	-5.24	-46.15	-55.85	-252.02	-6.23	-21.87	-31.25	-73.16	-4.76	-43.71	-52.97	-94.83	-7.28	-51.07	-65.06	-192.59
1 min	-14.62	-42.31	-45.50	-87.06	-6.41	-78.24	-72.34	-217.16	-5.24	-61.50	-59.58	-213.57	-6.23	-26.00	-30.35	-71.57	-4.76	-57.20	-55.64	-91.24	-7.28	-58.68	-63.11	-189.00
5 min	-14.62	-45.09	-44.50	-119.56	-6.41	-87.84	-79.86	-180.12	-5.24	-71.38	-60.19	-286.05	-6.23	-39.05	-33.83	-96.74	-4.76	-72.12	-63.04	-93.59	-7.28	-109.54	-72.47	-166.02
Maximum																								
1 sec	4.89	13.29	52.36	80.88	5.43	32.45	77.04	76.62	4.98	32.67	84.36	130.76	6.93	18.75	70.60	80.17	4.70	32.73	81.03	113.80	6.79	46.65	115.67	135.62
10 sec	4.89	15.69	23.92	80.86	5.43	42.43	53.07	84.92	4.98	41.38	53.32	138.09	6.93	21.60	31.06	80.43	4.70	36.20	50.06	113.98	6.79	53.28	73.46	135.85
1 min	4.89	21.79	25.33	82.83	5.43	50.85	52.35	84.52	4.98	56.74	54.13	155.97	6.93	28.43	33.60	71.87	4.70	39.89	46.02	104.79	6.79	60.66	67.33	134.18
5 min	4.89	29.89	29.42	65.83	5.43	47.06	47.40	87.77	4.98	59.75	60.01	149.84	6.93	35.59	37.04	133.66	4.70	45.12	43.44	120.43	6.79	58.10	59.29	128.91

Table 13: Distributions of Standardised Returns for Stocks MQG - RIO

	MQG				NAB				NCM				ORG				QBE				RIO			
	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$
Mean																								
1 sec	0.01	0.39	0.93	0.40	0.02	0.42	1.34	1.25	-0.15	-0.60	-1.62	-1.09	-0.02	-0.04	-0.26	0.26	-0.08	-0.13	-0.39	-0.41	-0.01	0.40	1.13	1.49
10 sec	0.01	0.41	0.59	0.23	0.02	0.51	0.73	1.16	-0.15	-0.68	-0.96	-0.99	-0.02	-0.03	-0.07	0.17	-0.08	-0.20	-0.23	-0.43	-0.01	0.47	0.65	1.47
1 min	0.01	0.54	0.58	0.49	0.02	0.76	0.81	1.11	-0.15	-0.77	-0.95	-1.23	-0.02	-0.03	-0.06	0.20	-0.08	-0.15	-0.19	-0.48	-0.01	0.60	0.67	1.42
5 min	0.01	0.63	0.60	0.30	0.02	0.96	0.92	1.55	-0.15	-0.89	-0.95	-1.23	-0.02	0.00	-0.03	0.25	-0.08	-0.12	-0.14	-0.67	-0.01	0.90	0.75	1.33
Variance																								
1 sec	3.62	128.05	1032.97	447.52	2.04	69.92	596.13	418.58	5.32	153.98	1139.15	486.18	2.23	50.71	498.67	242.69	4.03	66.68	591.05	387.53	3.10	202.18	1332.33	943.97
10 sec	3.62	157.19	320.59	458.99	2.04	100.25	193.11	402.09	5.32	198.07	385.56	495.14	2.23	63.62	136.81	246.14	4.03	94.49	185.77	397.92	3.10	274.73	501.54	946.86
1 min	3.62	198.26	283.17	460.08	2.04	153.57	194.66	416.07	5.32	256.02	344.72	518.47	2.23	90.96	134.97	247.95	4.03	152.37	194.61	387.67	3.10	367.57	457.82	930.38
5 min	3.62	233.00	256.11	499.21	2.04	204.51	201.05	469.07	5.32	293.80	312.76	595.91	2.23	125.86	137.21	286.12	4.03	215.55	207.64	423.20	3.10	426.63	436.09	1017.09
Skewness																								
1 sec	0.22	0.15	0.12	0.11	-0.34	-0.24	-0.18	-1.33	0.00	0.13	0.18	0.02	-0.20	-0.05	-0.06	-0.15	-2.50	-0.07	0.20	-0.66	-0.09	0.08	0.06	0.84
10 sec	0.22	0.15	0.14	-0.93	-0.34	-0.33	-0.26	-1.08	0.00	0.12	0.13	0.03	-0.20	-0.08	-0.06	-0.33	-2.50	-0.40	-0.06	-0.78	-0.09	0.06	0.06	0.76
1 min	0.22	0.19	0.16	0.25	-0.34	-0.37	-0.30	-1.30	0.00	0.14	0.12	-0.23	-0.20	-0.11	-0.08	-0.22	-2.50	-0.07	0.00	-0.84	-0.09	0.03	0.05	0.36
5 min	0.22	0.23	0.18	0.10	-0.34	-0.48	-0.37	-1.18	0.00	0.08	0.10	-0.05	-0.20	-0.06	-0.08	0.27	-2.50	-0.34	0.03	-0.86	-0.09	0.07	0.06	0.25
Kurtosis																								
1 sec	4.91	2.97	3.01	9.77	5.28	3.71	3.51	17.54	5.27	2.99	2.94	5.02	4.92	3.22	3.07	5.47	31.36	6.66	5.29	8.47	3.85	2.89	2.83	10.89
10 sec	4.91	3.15	3.01	17.07	5.28	4.20	3.81	13.80	5.27	3.18	2.98	5.09	4.92	3.34	3.13	5.92	31.36	9.18	6.72	10.00	3.85	3.15	2.94	10.56
1 min	4.91	3.21	3.05	7.58	5.28	4.65	4.14	16.77	5.27	3.34	3.03	6.28	4.92	3.39	3.12	5.54	31.36	8.40	6.77	8.35	3.85	3.35	3.10	6.89
5 min	4.91	3.19	3.09	8.31	5.28	5.47	4.63	17.63	5.27	3.35	3.11	5.74	4.92	3.20	3.14	8.86	31.36	9.64	6.70	10.37	3.85	3.15	3.16	8.54
Jarque-Bera Test																								
1 sec	0.001	**0.130**	**0.316**	0.001	0.001	0.001	0.002	0.001	0.001	**0.228**	**0.064**	0.001	0.001	**0.303**	**0.500**	0.001	0.001	0.001	0.001	0.001	0.001	**0.471**	**0.409**	0.001
10 sec	0.001	**0.084**	**0.167**	0.001	0.001	0.001	0.001	0.001	0.001	**0.164**	**0.222**	0.001	0.001	**0.058**	**0.500**	0.001	0.001	0.001	0.001	0.001	0.001	**0.473**	**0.500**	0.001
1 min	0.001	0.025	**0.107**	0.001	0.001	0.001	0.001	0.001	0.001	0.022	**0.312**	0.001	0.001	0.021	**0.448**	0.001	0.001	0.001	0.001	0.001	0.001	**0.075**	**0.500**	0.001
5 min	0.001	0.009	0.049	0.001	0.001	0.001	0.001	0.001	0.001	0.046	**0.343**	0.001	0.001	**0.314**	**0.388**	0.001	0.001	0.001	0.001	0.001	0.001	**0.419**	**0.433**	0.001
Minimum																								
1 sec	-7.04	-35.91	-108.53	-154.68	-6.49	-32.17	-94.95	-201.17	-9.33	-39.75	-101.26	-103.71	-8.43	-26.72	-73.54	-98.50	-24.57	-56.11	-122.45	-117.41	-6.88	-40.96	-116.79	-147.82
10 sec	-7.04	-38.72	-56.52	-224.60	-6.49	-41.47	-55.02	-179.95	-9.33	-48.12	-60.05	-112.47	-8.43	-28.49	-39.38	-104.60	-24.57	-77.16	-90.18	-140.26	-6.88	-55.25	-69.19	-136.31
1 min	-7.04	-45.84	-49.92	-118.02	-6.49	-58.68	-61.55	-200.44	-9.33	-52.92	-58.03	-154.34	-8.43	-35.54	-38.31	-97.46	-24.57	-82.25	-87.13	-112.15	-6.88	-67.87	-69.54	-148.73
5 min	-7.04	-49.38	-48.03	-134.89	-6.49	-76.34	-67.55	-216.36	-9.33	-66.01	-58.97	-130.34	-8.43	-37.66	-38.16	-104.88	-24.57	-121.77	-87.63	-140.10	-6.88	-66.57	-70.65	-158.38
Maximum																								
1 sec	10.60	40.34	97.71	131.03	6.42	25.94	76.26	86.50	10.77	40.73	105.24	106.65	5.28	21.91	70.69	63.64	9.26	48.34	153.12	107.24	5.48	44.68	105.85	266.94
10 sec	10.60	45.65	57.96	88.01	6.42	29.68	41.34	79.17	10.77	47.90	62.31	103.76	5.28	25.04	34.69	67.23	9.26	57.34	83.28	107.08	5.48	55.72	68.95	265.54
1 min	10.60	50.23	56.30	109.25	6.42	36.25	41.44	87.45	10.77	60.79	59.62	110.17	5.28	28.84	32.80	67.03	9.26	83.59	89.34	86.40	5.48	64.48	68.60	178.08
5 min	10.60	62.61	56.63	126.31	6.42	38.99	41.05	100.76	10.77	70.52	63.04	114.58	5.28	33.94	33.51	133.28	9.26	87.03	91.43	123.69	5.48	59.32	64.70	240.32

Table 14: Distributions of Standardised Returns for Stocks STO - WPL

	STO				TLC				WBC				WOW				WPL			
	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$	r_t	$\frac{r_t}{\sqrt{\widehat{RV}_i}}$	$\frac{r_t}{\sqrt{\widehat{RR}_i}}$	$\frac{r_t}{\sqrt{\widehat{GV}_i}}$
										Mean										
1 sec	0.00	0.10	0.29	0.68	0.04	0.11	0.45	1.44	0.02	0.42	1.24	1.20	0.02	0.26	0.76	0.12	-0.02	-0.02	-0.11	0.81
10 sec	0.00	0.13	0.19	0.65	0.04	0.14	0.19	1.48	0.02	0.50	0.71	1.22	0.02	0.32	0.44	0.15	-0.02	0.00	-0.01	0.81
1 min	0.00	0.23	0.24	0.78	0.04	0.26	0.24	1.40	0.02	0.74	0.81	1.33	0.02	0.42	0.47	0.05	-0.02	0.06	0.03	0.91
5 min	0.00	0.29	0.30	0.62	0.04	0.58	0.43	1.22	0.02	1.00	0.93	1.38	0.02	0.61	0.53	-0.05	-0.02	0.04	0.07	0.67
										Variance										
1 sec	3.10	59.05	535.05	320.59	1.36	2.46	31.79	250.47	2.08	67.16	575.54	362.27	0.95	48.72	401.74	231.42	2.30	93.56	727.84	353.97
10 sec	3.10	76.03	160.05	325.87	1.36	4.47	8.11	259.46	2.08	95.94	186.10	363.31	0.95	63.90	128.03	225.41	2.30	120.75	235.35	358.98
1 min	3.10	112.94	160.07	344.14	1.36	16.22	14.79	262.61	2.08	149.00	188.96	341.48	0.95	95.90	127.30	246.98	2.30	156.77	209.71	368.87
5 min	3.10	150.29	160.55	364.28	1.36	55.61	40.06	312.99	2.08	192.98	193.10	431.55	0.95	136.69	132.38	273.43	2.30	182.22	192.53	401.25
										Skewness										
1 sec	-0.15	0.11	0.07	0.16	-1.21	-0.23	0.21	-0.75	-0.24	-0.12	-0.10	-1.43	-0.16	0.10	0.18	-0.50	0.05	0.15	0.10	0.12
10 sec	-0.15	0.08	0.12	0.15	-1.21	-0.48	-0.28	-1.17	-0.24	-0.17	-0.14	-1.43	-0.16	0.09	0.12	-0.44	0.05	0.09	0.13	0.10
1 min	-0.15	0.09	0.12	0.04	-1.21	-0.61	-0.50	-1.48	-0.24	-0.18	-0.15	-0.29	-0.16	0.08	0.10	-0.58	0.05	0.15	0.13	0.19
5 min	-0.15	0.10	0.13	-0.05	-1.21	-0.38	-0.54	-2.25	-0.24	-0.13	-0.14	-1.22	-0.16	0.06	0.07	-0.55	0.05	0.07	0.10	-0.08
										Kurtosis										
1 sec	7.01	4.50	3.96	6.54	13.11	6.24	7.10	8.95	4.83	3.41	3.29	19.09	6.11	3.76	3.46	5.01	6.99	4.03	3.47	7.63
10 sec	7.01	4.88	4.49	6.42	13.11	7.69	6.70	13.61	4.83	3.57	3.37	19.12	6.11	4.05	3.73	4.75	6.99	4.97	3.99	7.47
1 min	7.01	5.41	4.56	7.24	13.11	7.97	6.98	15.79	4.83	3.63	3.43	6.03	6.11	4.11	3.80	5.83	6.99	4.60	4.14	6.96
5 min	7.01	4.29	4.36	7.49	13.11	5.40	5.77	22.18	4.83	3.54	3.49	18.49	6.11	4.07	3.92	5.94	6.99	4.01	3.97	7.20
										Jarque-Bera Test										
1 sec	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.013	**0.065**	0.001	0.001	0.001	0.003	0.001	0.001	0.001	0.008	0.001
10 sec	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.014	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
1 min	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.007	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
5 min	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.003	0.004	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
										Minimum										
1 sec	-11.49	-32.19	-77.91	-100.53	-10.64	-9.05	-26.38	-119.05	-6.72	-26.73	-84.00	-210.66	-5.71	-28.64	-75.43	-89.33	-8.57	-42.30	-101.21	-86.44
10 sec	-11.49	-43.89	-54.31	-100.50	-10.64	-14.86	-17.33	-147.83	-6.72	-33.19	-46.40	-210.84	-5.71	-33.59	-44.96	-89.30	-8.57	-60.82	-64.85	-86.27
1 min	-11.49	-58.79	-58.50	-121.78	-10.64	-29.17	-25.61	-146.64	-6.72	-47.60	-47.37	-106.18	-5.71	-41.23	-43.47	-92.65	-8.57	-61.55	-65.58	-87.23
5 min	-11.49	-53.97	-55.83	-124.99	-10.64	-38.92	-35.72	-188.45	-6.72	-61.28	-53.86	-217.74	-5.71	-61.59	-49.23	-103.21	-8.57	-60.56	-61.69	-111.56
										Maximum										
1 sec	10.22	42.38	121.67	100.17	5.30	8.19	38.28	79.50	5.87	28.03	76.12	71.41	5.16	24.01	67.24	49.56	9.17	43.28	99.59	128.24
10 sec	10.22	48.07	69.53	100.37	5.30	10.14	15.02	71.23	5.87	35.32	46.68	75.55	5.16	28.18	38.54	49.56	9.17	54.89	68.74	128.19
1 min	10.22	61.14	69.10	89.88	5.30	16.89	17.75	79.40	5.87	44.65	50.69	78.12	5.16	36.08	39.27	62.26	9.17	57.07	65.47	126.51
5 min	10.22	57.22	65.84	104.98	5.30	35.23	25.26	55.66	5.87	48.39	53.65	105.04	5.16	39.73	40.47	71.63	9.17	55.73	59.53	116.30