Data-Based Ranking of Realised Volatility Estimators*

Andrew J. Patton
London School of Economics
10 April 2007
Preliminary and Incomplete.
Please do not cite without permission

Abstract

I propose a feasible method for ranking realised variance (RV) estimators based on actual returns data. In contrast, most rankings of RV estimators currently in the literature are either graphical in nature, most notably the “volatility signature plot”, or rely on asymptotic approximations of the mean-squared errors of the estimators, or on simulations. The proposed method relies on the existence of a volatility proxy that is unbiased for the variable of interest, and satisfies a certain “zero correlation” condition. The zero correlation condition has some similarities with instrumental variables estimation. The volatility proxy must be unbiased but it does not need to be very precise; a simple and widely-available proxy for conditional variance is the daily squared return. From a small empirical application to IBM volatility estimation, I find that the daily squared return is significantly out-performed by an RV estimator based on intra-daily data, while a simple RV estimator based on 5-minute returns was not significantly out-performed by any of 31 other RV estimators.

Keywords: realised variance, volatility forecasting, microstructure effects.

J.E.L. codes: C52, C22.

*I thank Neil Shephard and Kevin Sheppard for helpful comments and suggestions, without implicating them for any errors. Contact address: Financial Markets Group, London School of Economics, Houghton Street, London WC2A 2AE, U.K. Email: a.patton@lse.ac.uk. A complete draft of paper will be available from http://fmg.lse.ac.uk/~patton.
1 Introduction

Most rankings of realised volatility (RV) estimators in the existing econometrics literature are either graphical in nature (as opposed to formal statistical tests), notably the “volatility signature plot” of Andersen, et al. (2000), or rely on asymptotic approximations of the mean-squared errors of the estimators or on simulations. In this paper I propose a formal ranking method based on actual data, which does not rely on asymptotic approximations or on simulating a “realistic” description of real data. The proposed ranking method relies on the existence of a volatility proxy that is unbiased for the latent target variable, and satisfies an uncorrelatedness condition, described in detail below. This proxy must be unbiased but it may not need to be very precise. A simple and widely-available proxy for conditional variance is the daily squared return 1.

The use of a consistent, data-based ranking of RV estimators has numerous advantages over rankings obtained via asymptotic theory or simulations. Compared with the former, it allows one to examine the finite-sample performance of these estimators, which can differ widely from their asymptotic performance, as noted by Bandi and Russell (2006). Furthermore, much of the asymptotic theory for RV estimators in the presence of market microstructure noise relies on very specific assumptions about the noise process. For example, Hansen and Lunde (2006) consider noise that is iid and additive to the efficient log-price process, or that is mean zero and covariance stationary. Zhang, et al. (2005) also consider the iid noise case, as do Barndorff-Nielsen, et al. (2006) and Bandi and Russell (2006). A data-based ranking has the obvious advantage over a simulation-based approach in that the latter requires a complete specification of the data generating process, and results obtained under one specification/parameterisation need not necessarily hold more generally. The data-based approach presented here allows one to answer the question of immediate interest to users of RV estimators: “which estimator works best on my asset return series, for my sample period?”

The method I propose below relies on the presence of a potentially noisy but conditionally unbiased proxy for the latent target variable. For many assets the squared daily return can reasonably be assumed to be conditionally unbiased: the expected return is generally negligible at the daily frequency, and the impact of market microstructure effects is often also negligible in daily returns.

1 More generally one might consider proxies of the form $Y_t = \sum_{j=-M}^{M} \omega_j r^2_{t-j}$ such as the rolling-window estimator of Foster and Nelson (1996), however we will need to impose $\omega_0 = 0$ for reasons described below.
It should be noted, however, that the presence of jumps in the data generating process will affect the inference obtained using the daily squared return as a proxy: in this case we can compare the RV estimators in terms of their ability to estimate quadratic variation, which is the integrated variance plus the sum of squared jumps in many cases, see Barndorff-Nielsen and Shephard (2006) for example, not in terms of their ability to estimate the integrated variance alone. If an estimator of the integrated variance that is conditionally unbiased even in the presence of jumps is available, however, the methods presented below apply directly.

1.1 Notation

$\theta_t$ is the latent target variable, which I take to be the quadratic variation of some asset price series. I assume that $\theta_t$ is $\mathcal{F}_t$-measurable, though it is not observable to the econometrician.

$X_{it}$, $i = 1, 2, \ldots, n$ are the realised volatility estimators to be ranked. Often these will be the same estimator applied to data sampled at different frequencies, for example: 1-minute returns vs. 30-minute returns. They could also be estimators based on different measures of the price: trades vs. mid-quotes, for example. The estimators could instead be completely different: the multi-scale sub-sampled realised variance estimator of Zhang, 2006, vs. the ‘alternation’ estimator of Large, 2005, or, in the extreme case, a realised variance estimator vs. a GARCH(1,1) forecast.

In order to rank the competing estimators we need some measure of distance from the estimator, $X_{it}$, to the target variable, $\theta_t$. In rankings of estimators based on asymptotic ($m \to \infty$, where $m$ is the number of intra-daily observations) approximations this distance is usually the mean-squared error (MSE). When the two estimators are both consistent this reduces to comparing the asymptotic variances of the two estimators. Barndorff-Nielsen, et al. (2006) provide a detailed study of the asymptotic accuracy of a wide variety of ‘kernel-based’ realised volatility estimators, Hansen and Lunde (2006) study the asymptotic MSE of a variety of estimators under different assumptions on the microstructure noise, while Bandi and Russell (2006) study the finite-sample MSE of some kernel-based realised volatility estimators under the assumption of iid microstructure noise. The extensive simulation study of Gatheral and Oomen (2007) also uses MSE to measure the distance between the estimator and the target variable.

I will consider ranking RV estimators using the average distance between the estimator and the target variable, using the general class of pseudo-distance measures proposed in Patton (2006):
\[ E[L(\theta_t, X_{it}; b)] \] vs. \[ E[L(\theta_t, X_{jt}; b)] \]  

(1)

where \[ L(\theta, X; b) = \tilde{C}(X; b) - \tilde{C}(\theta; b) + C(X; b)(\theta - X) \]  

(2)

with \[ C(z; b) = \begin{cases} 
- (b + 1)^{-1} z^{b+1}, & \text{for } b \notin \{-1, -2\} \\
- \log z, & \text{for } b = -1 \\
z^{-1}, & \text{for } b = -2 
\end{cases} \]

and \[ \tilde{C}(z; b) = \int C(z; b) dz \]

\[ = \begin{cases} 
- (b + 1)^{-1} (b + 2)^{-1} z^{b+2}, & \text{for } b \notin \{-1, -2\} \\
z - z \log z, & \text{for } b = -1 \\
\log z, & \text{for } b = -2 
\end{cases} \]

This class nests MSE as a special case (\( b = 0 \)) and the popular “QLIKE” loss function (\( b = -2 \)), up to location and scale constants in both cases. More generally, the “shape” parameter \( b \) affects the penalty applied to over-estimation compared with under-estimation. This class is well-defined when both \( \theta \) and \( X \) are almost surely strictly positive, which is a reasonable assumption in applications involving realised variance.

Our interest is in measuring the average distance between the estimator and the latent target variable. I will obtain a consistent estimator of this quantity by employing a proxy or instrument for \( \theta_t \), denoted \( Y_t \). The proxy must be observable by the econometrician, for the ranking to be “data-based”, and must satisfy certain unbiasedness and zero correlation conditions. Deriving these conditions and finding a proxy that satisfies them is the main technical challenge in this paper.

2 Relation to the ranking of volatility forecasts

Ranking volatility forecasts, as opposed to estimators, has received a lot of attention in the econometrics literature, see Poon and Granger (2003) and Hansen and Lunde (2005) for two recent and comprehensive studies, and this is the natural starting point for considering the ranking realised volatility estimators. Hansen and Lunde (2006) and Patton (2006) show that if:

\[ E[Y_t|\mathcal{F}_{t-1}] = \theta_t \]
(i.e., the proxy is conditionally unbiased for $\theta_t$) then for any pseudo-distance measure in the class in equation (2) rankings based on the proxy are ($T-$asymptotically) equivalent to rankings based on the true unobservable target variable, assuming that the expectations exist. That is,

$$E [L (\theta_t, X_{1t})] \leq E [L (\theta_t, X_{2t})] \Leftrightarrow E [L (Y_t, X_{1t})] \leq E [L (Y_t, X_{2t})]$$

(3)

However, this result does not go through when $(X_{1t}, X_{2t})$ are RV estimators rather than a volatility forecasts. To see this, consider a mean-value expansion of the pseudo-distance measure $L$ given in equation (2):

$$L (Y_t, X_t) = L (\theta_t, X_t) + \frac{\partial L (\theta_t, X_t)}{\partial \theta} (Y_t - \theta_t) + \frac{1}{2} \frac{\partial^2 L (\tilde{\theta}_t, X_t)}{\partial \theta^2} (Y_t - \theta_t)^2$$

where $\tilde{\theta}_t = \lambda_t \theta_t + (1 - \lambda_t) Y_t$, $\lambda_t \in [0, 1]$

then $E_{t-1} [L (Y_t, X_t)] = E_{t-1} [L (\theta_t, X_t)] + E_{t-1} [(C (X_t) - C (\theta_t)) (Y_t - \theta_t)]$

$$- \frac{1}{2} E_{t-1} [C' (\tilde{\theta}_t) (Y_t - \theta_t)^2]$$

(4)

The third term in equation (4) does not depend on $X_t$, and so will not affect the ranking of $(X_{1t}, X_{2t})$. For the ranking obtained using $Y_t$ to be the same as that obtained using $\theta_t$ we need to show that the second term equals zero:

$$E_{t-1} [(C (X_t) - C (\theta_t)) (Y_t - \theta_t)] = 0$$

In the standard case, $X_t$ is a volatility forecast and $\theta_t$ is $\mathcal{F}_{t-1}$-measurable, which allows:

$$E_{t-1} [(C (X_t) - C (\theta_t)) (Y_t - \theta_t)] = (C (X_t) - C (\theta_t)) \cdot E_{t-1} [Y_t - \theta_t] = 0$$

by the conditional unbiasedness of $Y_t$. However, when $X_t$ is a realised volatility estimator we have $(X_t, \theta_t) \in \mathcal{F}_t$ but $(X_t, \theta_t) \notin \mathcal{F}_{t-1}$, and so we cannot take the first term above out of the expectation. In short, the fact that the realised variance estimator of the target variable for day $t$ is only available at the end of day $t$ rules out the direct application of established results for volatility forecast comparison.

If we could assume that

$$\text{Corr}_{t-1} [C (X_t) - C (\theta_t), Y_t - \theta_t] = 0,$$
in addition to $E[Y_t|\mathcal{F}_{t-1}, \theta_t] = \theta_t$, then we would have

$$E_{t-1} [(C(X_t) - C(\theta_t))(Y_t - \theta_t)] = E_{t-1} [(C(X_t) - C(\theta_t)) \cdot E_{t-1}[Y_t|\mathcal{F}_{t-1}, \theta_t] - \theta_t]$$

$$= 0$$

But it is not likely that $Corr_{t-1} [C(X_t) - C(\theta_t), Y_t - \theta_t] = 0$ for all volatility proxies $Y_t$. For example, if $X_{it} = Y_t$ and $L = MSE$, then

$$C(z) = -z$$

so $Corr_{t-1} [C(X_t) - C(\theta_t), Y_t - \theta_t] = Corr_{t-1} [\theta_t - Y_t, Y_t - \theta_t] = -1$

Thus this correlation will in fact equal -1! More generally, we would expect this correlation to be non-zero. It is the correlation between the error in $Y_t$ and something similar to the “generalised forecast error”, see Granger (1999) or Patton and Timmermann (2003), of $X_t$. If the proxy, $Y_t$, and the RV estimator, $X_{it}$, are both using the same or similar information sets then their errors will generally be correlated and this zero correlation restriction will not hold. This reveals the similarity of this problem to instrumental variables estimation: correlation between the error in the RV estimator and the error in the proxy leads to invalid inference.

3 Data-based ranking of RV estimators

I present results under two broad sets of assumptions: the first allows for general behaviour in the target variable, $\theta_t$, but restricts the behaviour of the RV estimators, $X_{it}$. The second set of assumptions allows for general behavior of the RV estimators, at the cost of imposing some restrictions on the behaviour of the target variable. We present both sets of results as in different applications one set of assumptions may be more palatable than the other.

3.1 Rankings based on assumptions about the RV estimators

This section presents results for data-based ranking of RV estimators that hold when we can assume that the time series behaviour of the bias in the RV estimators satisfies restriction given below.

Assumption T1: $\theta_t$ is a mean stationary process and $\theta_t > 0$ a.s.

Assumption P1: $Y_t$ is a mean stationary process with $E[Y_t] = E[\theta_t]$.

Assumption P2: $Y_t \in \mathcal{F}_{t-1}$ and $Y_t > 0$ a.s.
**Assumption R1:** \[ E[X_{it}|\mathcal{F}_{t-1}] = \theta_t + c_i \theta_i^k \quad \forall \ i, \] where \( k \) is known, and \( \max_i |c_i| < \infty \).

The first two of these assumptions are standard, with only unconditional unbiasedness of \( Y_t \) required (rather than conditional unbiasedness). Assumption P2 requires that the proxy is almost surely positive, a standard assumption for a volatility proxy, and is measurable at time \( t-1 \), which is non-standard. We would usually consider a proxy for \( \theta_t \) as being something measured on day \( t \), such as the squared returns from day \( t \). Assumption P2 suggests instead to use the first lag of the daily squared return, or longer lags, or perhaps combinations of lags. (We will consider optimal choices of proxies below.) The result below shows that using lagged squared returns can be useful in obtaining a data-based ranking of RV estimators.

Assumption R1 is the key assumption for this result. It requires that the bias in the RV estimators be proportional to some power of the target variable, with a common power but potentially different proportionality constants. This nests the interesting special cases where all RV estimators are unbiased \( (c_i = 0 \ \forall \ i) \), or where all RV estimators have some biases that are constant through time \( (k = 0) \) but which can differ across estimators \( (c_i \neq c_j) \). It also allows the biases in the RV estimators to be proportional to some power of \( \theta_t \), for example \( \theta_t^2 \). This might be of interest as in many cases asymptotic variance of many RV estimators is related to integrated quarticity, which is in turn related to the square of integrated variance, see Barndorff-Nielsen and Shephard (2004) for example.

**Proposition 1** Let assumptions T1, P1, P2 and R1 hold. Then

\[
E[L(\theta_t, X_{1t}; b)] - E[L(\theta_t, X_{2t}; b)] = E[L(Y_t, X_{1t}; b)] - E[L(Y_t, X_{2t}; b)], \quad \text{if } b = k = 0
\]

\[
\approx E[L(Y_t, X_{1t}; b)] - E[L(Y_t, X_{2t}; b)], \quad \text{if } b = k^{-1} \text{ and } k \neq 0
\]

for any two RV estimators, \( X_{1t} \) and \( X_{2t} \), where \( k \) is from assumption R1.

All proofs are presented in the Appendix.

This proposition shows that if we can make some assumption about the time series behaviour of the bias in the competing RV estimators\(^2\), then there exists a unique pseudo-distance measure from equation (2) that yields a feasible data-based ranking of RV estimators. For example, if... 

\(^2\)Note, importantly, that we do not need to assume anything about the behaviour of the biases as \( m \) (the number of intra-daily observations) varies. Different choices of \( m \) correspond to different RV estimators in this framework and no relation between the competing estimators is imposed.
the biases in the RV estimators are constant through time, then we can rank the RV estimators using MSE \((b = 0)\). If the biases in the RV estimators are proportional to \(\theta^2_t\), then the QLIKE pseudo-distance measure \((b = -2)\) should be used to rank the RV estimators.

3.2 Rankings based on assumptions about the target variable

The result from the previous section relied on a rather specific assumption about the time series properties of the biases in the competing RV estimators. In this section I do away with such assumptions by imposing some structure on the time series dynamics of the target variable, \(\theta_t\). Numerous papers on the conditional variance (see Bollerslev, et al., 1994, Engle and Patton, 2001, and Andersen, et al., 2005 for example), or integrated variance (see Andersen, et al., 2003 and 2005) have reported that these quantities are very persistent, close to being random walks. Wright (1999) provides thorough evidence against the presence of a unit root in daily conditional variance for many stocks, however, despite this, it has proven to be a good approximation in many cases. The popular RiskMetrics model, for example, is based on a unit root assumption for the conditional variance. Given this, consider the following assumption:

**Assumption T2:** \(\theta_t = \theta_{t-1} + \eta_t\), with \(E[\eta_t|\mathcal{F}_{t-1}] = 0\).

In the proof of the following proposition I need to strengthen the unconditional unbiasedness assumption in P1 to the standard conditional unbiasedness assumption. We will denote the conditionally unbiased proxy as \(\tilde{\theta}_t\), rather than \(Y_t\), as below I will consider taking linear combinations of unbiased proxies to improve the power of tests in finite samples.

**Assumption P1':** \(\tilde{\theta}_t = \theta_t + \nu_t\), with \(E[\nu_t|\mathcal{F}_{t-1}, \theta_t] = 0\), and \(\tilde{\theta}_t > 0\) a.s.

For the proposition below I again consider using a proxy for \(\theta_t\) that is not measured on day \(t\), but instead of considering lags of \(\tilde{\theta}_t\) it turns out to be best to consider leads of \(\tilde{\theta}_t\). The reason for this becomes clear in the proof.

**Assumption P2':** \(Y_t = \sum_{i=1}^{J} \omega_i \tilde{\theta}_{t+i}\), where \(1 \leq J < \infty\), \(\omega_i \geq 0 \forall i\) and \(\sum_{i=1}^{J} \omega_i = 1\).

**Proposition 2** Let assumptions T2, P1' and P2' hold. Then

\[
E[L(\theta_t, X_{1t}; b)] - E[L(\theta_t, X_{2t}; b)] = E[L(Y_t, X_{1t}; b)] - E[L(Y_t, X_{2t}; b)]
\]

for any two RV estimators, \(X_{1t}\) and \(X_{2t}\), and any \(b\).
In some ways the above result is substantially more general than that in Proposition 1. The assumption that \( \theta_t \) follows a random walk allows us to leave the bias, if any, in the RV estimators completely unspecified: it can be constant, time-varying, a function of different powers of \( \theta_t \), or a function of other variables altogether. Furthermore, Proposition 2 shows that any pseudo-distance measure from the class in equation (2) may be used, according to the preferences of the user of the RV estimator. Note that for a formal RV comparison test to be implemented we will need certain moment conditions to be satisfied and this may restrict the choice of pseudo-distance measure. These moment conditions are discussed in the next section.

An alternative motivation for the empirical approach suggested by Proposition 2 is based on the asymptotics of rolling window volatility estimators given in Foster and Nelson (1996), and used in Fleming, et al. (2001), amongst many other applications. Foster and Nelson show that, under some conditions, estimators such as those covered in assumption P2’ converge to true (spot) variance as the length of the period \( H = 1/m \) (one day, in our case) goes to zero and as the number of intra-period observations, \( m \), goes to infinity.

Before moving on, it is worth considering how the above proposition changes when the target variable is only “close to” a random walk. To that end, consider the following modification the random walk assumption:

**Assumption T2’**: \( \theta_t = \mu + \phi (\theta_{t-1} - \mu) + \eta_t \), with \( E[\eta_t|\mathcal{F}_{t-1}] = 0 \), and \( \mu \equiv \bar{\theta} \delta \), \( \phi \equiv 1 - \delta \), where \( \delta \) is a small positive constant.

Under this weaker assumption on the time series dynamics of the target variable I obtain the following result. For simplicity I restrict the proxy to be a simple lead of \( \bar{\theta}_t \).

**Proposition 3** Let assumptions T2’ and P1’ hold, and set \( Y_t = \bar{\theta}_{t+1} \). (i) Then

\[
E[L(\theta_t, X_{1t}; b)] - E[L(\theta_t, X_{2t}; b)] = E[L(Y_t, X_{1t}; b)] - E[L(Y_t, X_{2t}; b)] \\
- \delta E[\theta_t (C(X_{1t}; b) - C(X_{2t}; b))] \\
+ \bar{\theta} \delta^2 E[C(X_{1t}; b) - C(X_{2t}; b)]
\]

for any two RV estimators, \( X_{1t} \) and \( X_{2t} \), and any \( b \).
(ii) If we further assume R1, with \( c_2 - c_1 \equiv \epsilon \), then

\[
\delta E \left[ \theta_t (C (X_{1t}; b) - C (X_{2t}; b)) \right] = \delta \epsilon E \left[ \theta_t^{b+1} \right], \text{ for } b = 0 \\
\approx \delta \epsilon E \left[ \theta_t^{b+1+b} \right], \text{ for } b \neq 0
\]

and \( \tilde{\delta} \delta^2 E [C (X_{1t}; b) - C (\theta_t; b)] = \tilde{\delta} \delta^2 \epsilon E \left[ \theta_t^{b} \right], \text{ for } b = 0 \\
\approx \tilde{\delta} \delta^2 \epsilon E \left[ \theta_t^{b+b} \right], \text{ for } b \neq 0
\]

The first part of Proposition 3 shows explicitly the extra terms that appear when the target variable follows an AR(1) rather than a random walk. The second part of the proposition provides some idea of the magnitudes of these terms as a function of \( \delta \), which measures how close to a random walk the target variable is, and \( \epsilon \), which measures the difference in the proportionality constants in the biases of the two RV estimators. Empirically it is widely found that \( \delta \) is positive but small. Thus the third term in part (i) is \( O (\delta^2) \) and may be negligible. The second term is \( O (\delta) \), and becomes \( O (\delta \epsilon) \) when we impose some structure on the biases in the RV estimators. If \( \delta \) is small, and we think \( \epsilon \) is small, then this term will also be negligible.

Proposition 3 provides some reassurance of the empirical usefulness of the ranking method suggested by Proposition 2: if the target variable is close to a random walk, and/or the RV estimators being compared have similar biases, then ranking RV estimators by using a lead of a conditionally unbiased proxy for \( \theta_t \) in conjunction with a pseudo-distance measure from equation (2) will yield the same ranking as if \( \theta_t \) was directly observable.

Proposition 2 above suggests the use of a convex combination of leads of \( \tilde{\theta}_t \), but gives no guidance on how many leads, \( J \), to consider or on the appropriate weights to apply to each lead individually. While the weighting function could theoretically have \( J - 1 \) free parameters (it must sum to one, pinning down the \( J^{th} \) weight) let us simplify the problem and consider only equally-weighted proxies. In that case, the problem reduces to choosing \( J \), the number of leads to combine.

**Proposition 4** Let \( P1', P2' \) and \( T2' \) hold. Then, imposing \( \omega_i = J^{-1} \forall i \), the variance of the error in the proxy for a given value of \( J \) is

\[
V [Y_t - \theta_t] = \frac{1}{J} \sigma^2 + \frac{(J + 1)(2J + 1)}{6J} \rho^2 + \left( 1 + \frac{1}{J} \right) \sigma^2
\]
The number of leads that minimises the variance of the measurement error in $Y_t$ is given by

$$J^* = \sqrt{\frac{1 + 6k + 6\rho \sqrt{k}}{2}}$$

where $k = \sigma^2_p / \sigma^2_\eta$ and $\rho \equiv Corr [\nu_t, \eta_t]$

When we constrain $J^*$ to be an integer between 1 and 10000, the optimal values are:

<table>
<thead>
<tr>
<th>$J^*$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>-0.9</td>
</tr>
<tr>
<td>0.0001</td>
<td>1</td>
</tr>
<tr>
<td>0.1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>100</td>
<td>17</td>
</tr>
<tr>
<td>10,000</td>
<td>172</td>
</tr>
</tbody>
</table>

These results reveal that the optimal integer values of $J$ do not vary greatly with $\rho$, though they do change with $k$. When $k < 1$, it is intuitively clear that we should use only one lead of $\tilde{\theta}_t$, as in that case $\tilde{\theta}_t$ is a relatively accurate estimator of $\theta_t$ and the gains from smoothing are low. When $k \geq 1$, there is potentially some benefit to smoothing the proxy across a range of leads of $\tilde{\theta}_t$. Only for very large values of $k$ do we average across more than a few leads of $\tilde{\theta}_t$.

It should be noted that the above result for the optimal value for $J$ is very sensitive to the random walk assumption for $\theta_t$: if $\theta_t$ is actually slowly mean-reverting then using leads of 100 or more periods will yield misleading results. In practice, it may be best to limit the value of $J$ to be no more than 5 or 10 for daily data, depending on the estimated persistence in the latent target variable.

Finally, I present the most general result, which allows the target variable to follow almost any stationary AR(1) process:

**Assumption T2**: $\theta_t = \mu + \phi (\theta_{t-1} - \mu) + \eta_t$, with $E [\eta_t | F_{t-1}] = 0$, and $|\phi| < 1$, $\phi \neq 0$.

The following result uses an instrumental variables estimator to address the correlation between the error in the proxy and the error in the RV estimator.
Proposition 5 Let assumptions T2" and P1' hold, and set $Y_t = \tilde{\theta}_{t+1}$. (i) Then

$$E [L (\theta_t, X_{1t}; b)] - E [L (\theta_t, X_{2t}; b)] = E [L (Y_t, X_{1t}; b)] - E [L (Y_t, X_{2t}; b)]$$

$$+ \frac{1 - \phi}{\phi} E [(C (X_{1t}) - C (X_{2t})) Y_t]$$

for any two RV estimators, $X_{1t}$ and $X_{2t}$, and any $b$.

(ii) Under standard regularity conditions, we have

$$\frac{1}{T} \sum_{t=1}^{T} \{L (Y_t, X_{1t}; b) - L (Y_t, X_{2t}; b)\}$$

$$+ \frac{1 - \hat{\phi}_T}{\hat{\phi}_T} \frac{1}{T} \sum_{t=1}^{T} \{(C (X_{1t}) - C (X_{2t})) Y_t\}$$

$$\to E [L (\theta_t, X_{1t}; b)] - E [L (\theta_t, X_{2t}; b)], \text{ as } T \to \infty$$

where $\hat{\phi}_T \equiv \frac{\frac{1}{T-2} \sum_{t=1}^{T-2} Y_{t+2} Y_t - \left(\frac{1}{T} \sum_{t=1}^{T} Y_t\right)^2}{\frac{1}{T-1} \sum_{t=1}^{T-1} Y_{t+1} Y_t - \left(\frac{1}{T} \sum_{t=1}^{T} Y_t\right)^2}$

Proposition 5 relaxes the assumption of a random walk, which introduces a bias term to the expected loss computed using the proxy, even when the error in the proxy is uncorrelated with the error in the RV estimators (e.g., when a lead of $\tilde{\theta}_t$ is used). This bias term, however, can be consistently estimated under the assumption that the target variable follows a stationary, non-trivial AR(1) process using an instrumental variables estimator. The cost of the added flexibility in allowing for a general AR(1) process for the target variable is the added estimation error induced by having to estimate the AR(1) parameter: this estimation error will lead to reduced power to distinguish between competing RV estimators than would otherwise be the case.

4 Simulation study

To be added.

5 Empirical application

In this section I consider the problem of estimating the quadratic variation of the daily return on IBM, using data from the TAQ database over the period from January 1993 to May 1998, yielding
1364 daily returns. This sample period was used in the Andersen, et al. (2001b) study of realised variance of equity returns. We use “last price” interpolation of the trade price series to create a series of 30-second prices, denoted \( \{ r_{t,j}, \ j = 1, 2, \ldots, 780 \}^T_{t=1} \). From this series I compute two simple types of realised volatility estimators:

\[
RV_t^{(m)} = \sum_{j=1}^{m} r_{t,j,m}^2
\]

\[
RVAC1_t^{(m)} = \sum_{j=1}^{m} r_{t,j,m}^2 + 2 \sum_{j=2}^{m} r_{t,j,m} r_{t,j-1,m}
\]

where \( r_{t,j,m} \) is the \( h \)-minute return, computed from the original 30-second return series. The first set of RV estimators includes standard realised variances computed for various values of \( m \). The second set are the “RV-AC1” estimators used in French, et al. (1987), and studied in Hansen and Lunde (2006) and Bandi and Russell (2006).

I consider all values for \( h \) that are multiples of one-half (so that I can evenly aggregate these from our original 30-second return series) and that divide evenly into 390, the number of minutes in a trade day on the New York stock exchange, which is open from 9.30am to 4pm. This yields 17 sampling frequencies: \( h = 0.5, 1, 2, 3, 5, 6, 10, 13, 15, 26, 30, 39, 65, 78, 130, 195, \) and 390 minutes, the final value for \( h \) corresponding to simply using the open-to-close return. The total number of RV estimators considered is thus 32: 17 standard RV estimators and 15 RVAC1 estimators (for \( h = 390 \) and \( h = 195 \) the RVAC1 estimator corresponds exactly to the squared open-to-close return). I impose an “insanity filter” on the RVAC1 estimator, as it is not guaranteed to yield positive estimates for finite \( m \): on days when the RVAC1 estimate is negative the estimate is replaced with the standard RV estimate with the same sampling frequency\(^3\). Figure 1 presents the volatility signature plot and a plot of the standard deviation of these estimators.

[ INSERT FIGURE 1 ABOUT HERE ]

Tables 1 to 3 present the first empirical contribution of this paper. These tables present the

\(^3\)Depending on the sampling frequency, between zero and 6% of days (zero to 85 observations) yielded negative RVAC1 estimates. Most sampling frequencies yielded only 5 to 20 negative estimates out of 1364 days in the sample.
average loss, under MSE and QLIKE, of the 32 RV estimators relative to the average loss incurred using the squared open-to-close (“daily”) return. A negative value indicates that the daily squared return was out-performed, while a positive value indicates the opposite. In all cases the proxy is the squared daily return. The three tables show the estimated average losses under three assumptions on the DGP: Table 1 is based on an AR(1) assumption for the latent target variable (Assumption T2”) and computes the loss differentials using the consistent estimator presented in Proposition 5. Table 2 is based on a random walk assumption for the latent target variable (Assumption T2) and computes the loss differentials using the estimator presented in Proposition 2. Table 3 is based on the (incorrect) assumption that the measurement error in the proxy is uncorrelated with the errors in the RV estimators, and uses the contemporaneous value of the proxy rather than a one-period lead as in the former two cases. The results from these three tables are summarised in Table 4, and depicted in Figure 2.

Under the AR(1) assumption for the target variable, the best two estimators according to both MSE and QLIKE are the RV estimators based on 30-second and 1-minute returns. The worst two estimators under MSE are daily squared returns and the RVAC1 based on 195-minute returns (i.e., half-day returns), while under QLIKE the two worst estimators are daily squared returns and the RVAC1 estimator using 30-minute returns.

Under the random walk assumption for the target variable, the best RV estimators according to MSE and QLIKE are the RV estimators based on 30-minute and 2-minute returns respectively. The second-best estimators are the RVAC1 based on 2-minute and 30-second returns. The worst estimators are similar to those under the AR(1) assumption: daily squared returns, and RVAC1 based either on 130-minute or 30-minute returns.

To illustrate the distortions caused by neglecting correlation between the error in the proxy and the RV estimators, I also present the ranking obtained under the naïve assumption that this correlation is zero. The resulting ranking suggests that daily squared returns are the best estimator of daily quadratic variation amongst all RV and RVAC1 estimators, which is driven purely by the fact that the correlation between the measurement errors goes to unity for the standard RV estimator when \( m = 1 \); far from the assumption that it is zero.
In Table 5 I present the results of formal comparisons of the 32 RV estimators considered in this empirical application. To do this I implement the “reality check” of White (2000), and a refinement of this test proposed by Hansen (2005). The reality check is a means of testing the null:

\[ H_0 : E[L(\theta_t, X_{0t})] \leq E[L(\theta_t, X_{it})], \text{ for all } i = 1, 2, \ldots, K \]

vs.

\[ H_a : E[L(\theta_t, X_{0t})] > E[L(\theta_t, X_{it})] \text{ for some } i \]

where \( X_{0t} \) is some benchmark RV estimator. That is, we test whether the benchmark RV estimator generates losses that are weakly smaller in expectation than any competing RV estimator. The null hypothesis contains \( K \) weak inequalities, and the critical values for this test can be easily obtained using a bootstrap procedure. Using the bootstrap also simplifies accounting for the impact of the instrumental variables estimator of the AR(1) coefficient on the asymptotic distribution of the test statistic, see Corollary 2.7 of White (2000). Hansen’s (2005) refinement of the White’s reality check involves a form of “trimming” to limit the impact of very poor estimators and studentising the test statistic; both of these refinements should lead to improved power to reject the null.

I consider three benchmark estimators of daily quadratic variation: the daily squared return, a standard RV estimator based on 5-minute returns, and the estimated volatility obtained from a Normal GARCH(1,1) model applied to the open-to-close return series. I present results under both the AR(1) assumption and the random walk assumption, which allows for some insight into the impact of estimation error in the AR(1) parameter estimate on the power of the test. Finally, I consider two proxies: the squared daily return, and a standard RV estimator based on 3 intra-daily returns \((h = 130)\). This latter estimator is approximately unbiased and is about 40% less volatile than daily squared returns, according to the plots in Figure 1, and so may lead to more powerful inference.

Table 5 reveals that the daily squared return can be rejected as being significantly beaten by some alternative RV estimator in many cases: for all cases under the random walk assumption it is rejected, and for both proxies using Hansen’s test under MSE. When using White’s test and the AR(1) assumption the daily squared return cannot be rejected either under MSE or QLIKE. This is perhaps indicative of low power for this application.
The standard RV estimator based on 5-minute returns is rarely rejected using either White’s or Hansen’s test, suggesting that for this sample period none of the competing RV or RVAC1 estimators were significantly better than this simple estimator. This finding provides some support for the rule-of-thumb that a simple 5-minute RV estimator works well in practice.

For comparison, I also considered the estimated volatility from a simple GARCH(1,1) model as a measure of daily quadratic variation. This estimator is almost certainly biased relative to RV estimators based on the current day’s information, as the GARCH estimate for day $t$ uses only data up until day $t - 1$, however the GARCH estimates will be smoother than the RV estimates, perhaps allowing for some bias-variance trade-off. This is indeed what is found: in only one case can the GARCH estimator be rejected in favour of one of the RV or RVAC1 estimators.

Overall, this small empirical application suggests that it is difficult to beat simple estimates of daily quadratic variation. A standard RV estimator based on 5-minute returns, and even an estimate obtained from a GARCH(1,1) model are not significantly out-performed by more sophisticated estimators based on higher frequency data. It remains to be seen whether this conclusion holds for other assets in other sample periods.

6 Conclusion

This paper considers the problem of comparing realised variance (RV) estimators. I propose a method for formally ranking RV estimators that does not rely on simulations, detailed assumptions about the market microstructure noise process, or on “large $m$” (or “continuous record”) asymptotics, though my method does rely on “large $T$” asymptotics. By either imposing some assumptions on the time series dynamics of the biases in the RV estimators, or by imposing a rather weak assumption on the time series dynamics of the latent target variable (quadratic variation or integrated variance), I present results that allow for a consistent estimate of the ranking of competing RV estimators. These results can be used in formal Diebold-Mariano (1995) pair-wise comparisons of RV estimators, or comparisons involving multiple estimators, such as the “reality check” of White (2000) or its refinement by Hansen (2005). In a small empirical application to IBM equity return volatility, I find reasonable evidence that the daily squared return is out-performed as a measure of quadratic variation by RV estimators based on higher frequency data. However, I find little evidence that a simple RV estimator constructed using 5-minute returns is out-performed
either by estimators using even higher frequency data, or by an estimator designed to be robust to market microstructure noise.
7 Appendix: Proofs

Proof of Proposition 1. Consider a first-order Taylor series expansion of $C(X_t; b)$ and assume $b \notin \{0, -1, -2\}$.

$$C(X_t; b) \approx C(\theta_t; b) + C'(\theta_t; b)(X_t - \theta_t)$$

so

$$E[(C(X_t; b) - C(\theta_t; b))(Y_t - \theta_t)] \approx E[C'(\theta_t; b)(X_t - \theta_t)(Y_t - \theta_t)]$$

$$= -E[\theta_t^{-k}(X_t - \theta_t)(Y_t - \theta_t)]$$

Under assumption P2 we have:

$$E[\theta_t^{-k}(X_t - \theta_t)(Y_t - \theta_t)] \approx E[\theta_t^{-k}(E[X_t|\theta_t, F_{t-1}] - \theta_t)(Y_t - \theta_t)]$$

$$= E[\theta_t^{-k}(c_t\theta_t^k)(Y_t - \theta_t)]$$

$$= c_t E[Y_t - \theta_t]$$

$$= 0$$

Thus we have

$$E[L(Y_t, X_{1t}; b)] = E[L(\theta_t, X_t; b)] + E[(C(X_t; b) - C(\theta_t; b))(Y_t - \theta_t)]$$

$$- \frac{1}{2}E[C'(\bar{\theta}_t; b)(Y_t - \theta_t)^2]$$

$$\approx E[L(\theta_t, X_t; b)] - \frac{1}{2}E[C'(\bar{\theta}_t; b)(Y_t - \theta_t)^2]$$

and so

$$E[L(Y_t, X_{1t}; b)] - E[L(Y_t, X_{2t}; b)] \approx E[L(\theta_t, X_{1t}; b)] - E[L(\theta_t, X_{2t}; b)]$$

up to the error term from the first-order Taylor series expansion of $C(X_t; b)$ around $C(\theta_t; b)$. The cases for $b = -1, -2$ go through similarly.

When $k = b = 0$, we have $C(z; b) = -z$, and so

$$E[(C(X_t; b) - C(\theta_t; b))(Y_t - \theta_t)] = E[(\theta_t - X_t)(Y_t - \theta_t)]$$

$$= E[(E[X_t|\theta_t, F_{t-1}] - \theta_t)(Y_t - \theta_t)]$$

$$= E[c_t(Y_t - \theta_t)]$$

$$= 0$$

without any Taylor series approximation. Thus we have

$$E[L(Y_t, X_{1t}; b)] - E[L(Y_t, X_{2t}; b)] = E[L(\theta_t, X_{1t}; b)] - E[L(\theta_t, X_{2t}; b)].$$

$\blacksquare$
Proof of Proposition 2. Consider again the expectation second term in the mean-value expansion of \( L(Y_t, X_t; b) \) around \( L(\theta_t, X_t; b) \):

\[
E \left[ (C(X_t; b) - C(\theta_t; b)) (Y_t - \theta_t) \right] \\
= E \left[ (C(X_t; b) - C(\theta_t; b)) \left( \sum_{i=1}^{J} \omega_i \tilde{\theta}_{t+i} - \theta_t \right) \right] \\
= E \left[ (C(X_t; b) - C(\theta_t; b)) \left( \sum_{i=1}^{J} \omega_i \left( \left( \theta_t + \sum_{j=1}^{i} \eta_{t+j} \right) + \nu_{t+i} \right) - \theta_t \right) \right] \\
= E \left[ (C(X_t; b) - C(\theta_t; b)) \left( \sum_{i=1}^{J} \omega_i \sum_{j=1}^{i} \eta_{t+j} + \sum_{i=1}^{J} \omega_i \nu_{t+i} \right) \right] \\
= E \left[ (C(X_t; b) - C(\theta_t; b)) \left( \sum_{i=1}^{J} \omega_i \sum_{j=1}^{i} E[\eta_{t+j}|\mathcal{F}_t] + \sum_{i=1}^{J} \omega_i E[\nu_{t+i}|\mathcal{F}_t] \right) \right] \\
= 0
\]

This then yields

\[
E[L(Y_t, X_{1t}; b)] - E[L(Y_t, X_{2t}; b)] = E[L(\theta_t, X_{1t}; b)] - E[L(\theta_t, X_{2t}; b)]
\]

using the same calculations as in the proof of Proposition 1. □

Proof of Proposition 3. (i) Consider once more the expectation second term in the mean-value expansion of \( L(Y_t, X_t; b) \) around \( L(\theta_t, X_t; b) \):

\[
E \left[ (C(X_t; b) - C(\theta_t; b)) (Y_t - \theta_t) \right] \\
= E \left[ (C(X_t; b) - C(\theta_t; b)) \left( \tilde{\theta}_{t+1} - \theta_t \right) \right] \\
= E \left[ (C(X_t; b) - C(\theta_t; b)) \left( \bar{\theta} \delta + (1 - \delta) (\theta_t - \bar{\theta} \delta) + \eta_{t+1} + \nu_{t+1} - \theta_t \right) \right] \\
= E \left[ (C(X_t; b) - C(\theta_t; b)) \left( \bar{\theta} \delta^2 - \delta \theta_t + \eta_{t+1} + \nu_{t+1} \right) \right] \\
= \bar{\theta} \delta^2 E[\theta_t (C(X_t; b) - C(\theta_t; b))] - \delta E[\theta_t (C(X_t; b) - C(\theta_t; b))] \\
= -\delta E[\theta_t (C(X_t; b) - C(\theta_t; b))] + o(\delta)
\]

(ii) First consider the case that \( b = 0 \):

\[
E \left[ \theta_t (C(X_{1t}; b) - C(X_{2t}; b)) \right] = -E \left[ \theta_t (X_{1t} - X_{2t}) \right] \\
= -E \left[ \theta_t \left( \theta_t + c_1 \theta_t^k + \xi_{1t} - \theta_t - c_2 \theta_t^k - \xi_{2t} \right) \right] \\
\]

where \( E[\xi_{it} | \theta_t, \mathcal{F}_{t-1}] = 0, i = 1, 2 \)

so \( E[\theta_t (C(X_{1t}; b) - C(X_{2t}; b))] = -\epsilon E[\theta_t^{k+1}] \)
For cases with \( b \neq 0 \) I employ a first-order Taylor series approximation of \( C (X_{it}; b) \) around \( C (\theta_t; b) \)

\[
C (X_{it}; b) \approx C (\theta_t; b) + C' (\theta_t; b) (X_{it} - \theta_t)
\]

so

\[
E \left[ \theta_t (C (X_{1t}; b) - C (X_{2t}; b)) \right] \approx E \left[ \theta_t \left( -\theta_t^b (X_{1t} - \theta_t) + \theta_t^b (X_{2t} - \theta_t) \right) \right]
\]

\[
= -E \left[ \theta_t^{b+1} (X_{1t} - X_{2t}) \right]
\]

\[
= -\epsilon E \left[ \theta_t^{k+1+b} \right]
\]

Proof of Proposition 4.

\[
Y_t = \frac{1}{J} \sum_{i=1}^{J} (\theta_{t+i} + \nu_{t+i})
\]

\[
= \frac{1}{J} \sum_{i=1}^{J} \left( \nu_{t+i} + \theta_t + \sum_{k=1}^{i} \eta_{t+k} \right)
\]

so

\[
Y_t - \theta_t = \frac{1}{J} \sum_{i=1}^{J} \nu_{t+i} + \frac{1}{J} \sum_{i=1}^{J} \sum_{k=1}^{i} \eta_{t+k}
\]

\[
= \frac{1}{J} \sum_{i=1}^{J} \nu_{t+i} + \frac{1}{J} \sum_{i=1}^{J} (J + 1 - i) \eta_{t+i}
\]

From this expression I compute the MSE of \( Y_t \), as a function of \( J \) and \( \sigma_{\eta}^2 \equiv E \left[ \eta_t^2 \right], \sigma_{\nu}^2 \equiv E \left[ \nu_t^2 \right] \) and \( \sigma_{\eta\nu} \equiv E \left[ \eta_t \nu_t \right] \)

\[
E \left[ (Y_t - \theta_t)^2 \right] = \frac{1}{J^2} E \left[ \left( \sum_{i=1}^{J} \nu_{t+i} \right)^2 \right] + \frac{1}{J^2} E \left[ \left( \sum_{i=1}^{J} (J + 1 - i) \eta_{t+i} \right)^2 \right]
\]

\[
+ \frac{2}{J^2} E \left[ \left( \sum_{i=1}^{J} \nu_{t+i} \right) \left( \sum_{i=1}^{J} (J + 1 - i) \eta_{t+i} \right) \right]
\]

\[
= \frac{1}{J} \sigma_{\nu}^2 + \frac{1}{J^2} \sigma_{\eta}^2 \sum_{i=1}^{J} (J + 1 - i)^2 + \frac{2}{J^2} \sigma_{\eta\nu} \sum_{i=1}^{J} (J + 1 - i)
\]

\[
= \frac{1}{J} \sigma_{\nu}^2 + \frac{(J + 1) (2J + 1)}{6J} \sigma_{\eta}^2 + \left( 1 + \frac{1}{J} \right) \sigma_{\eta\nu}
\]

This expression reveals the competing influences of the three terms: the first term is decreasing in \( J \), the second term is increasing in \( J \), and the third term is approximately flat in \( J \) for large values of \( J \); for small \( J \) it is decreasing in \( J \).
w.l.o.g., let $\sigma^2 = k\sigma^2_\eta$ and so $\sigma_{\eta
u} \equiv \rho \sqrt{k} \sigma^2_\eta$. In that case the first-order condition for $J^*$ becomes:

$$0 = -\frac{\sigma^2_\eta}{6J^2} \left( -1 + 2J^* - 6k - 6\rho \sqrt{k} \right)$$

so $J^* = \sqrt{\frac{1 + 6k + 6\rho \sqrt{k}}{2}}$

The optimal value of $J$ for various values of $k$ and $\rho$ is given below:

<table>
<thead>
<tr>
<th>$J^*$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>-0.9</td>
</tr>
<tr>
<td>0.0001</td>
<td>0.69</td>
</tr>
<tr>
<td>0.1</td>
<td>n/a</td>
</tr>
<tr>
<td>1</td>
<td>0.89</td>
</tr>
<tr>
<td>10</td>
<td>4.69</td>
</tr>
<tr>
<td>100</td>
<td>16.54</td>
</tr>
<tr>
<td>10,000</td>
<td>172.43</td>
</tr>
</tbody>
</table>

When we constrain $J^*$ to be an integer between 1 and 10000, the optimal values are those presented in the statement of the proposition. □

**Proof of Proposition 5.** (i) Recall the second-order mean-value expansion of the loss function

$$E [L (Y_t, X_{1t}; b)] = E [L (\theta_t, X_{1t}; b)] + E [(C (X_{1t}; b) - C (\theta_t; b)) (Y_t - \theta_t)]$$

$$\quad - \frac{1}{2} E \left[ C' (\tilde{\theta}_t; b) (Y_t - \theta_t)^2 \right]$$

so $E [L (Y_t, X_{1t}; b)] - E [L (Y_t, X_{2t}; b)] = E [L (\theta_t, X_{1t}; b)] - E [L (\theta_t, X_{2t}; b)]$

$$\quad + E [(C (X_{1t}; b) - C (X_{2t}; b)) (Y_t - \theta_t)]$$

If $Y_t = \tilde{\theta}_{t+1}$ and assumption T2” is satisfied, then:

$$E [(C (X_{1t}; b) - C (X_{2t}; b)) Y_t] = E [(C (X_{1t}; b) - C (X_{2t}; b)) (\theta_{t+1} + \nu_{t+1})]$$

$$= E [(C (X_{1t}; b) - C (X_{2t}; b)) \theta_{t+1}]$$

$$= E [(C (X_{1t}; b) - C (X_{2t}; b)) E_t [\theta_{t+1}]]$$

$$= \phi E [(C (X_{1t}; b) - C (X_{2t}; b)) \theta_t]$$

20
\[ E [L(\theta_t, X_{1t}; b)] - E [L(\theta_t, X_{2t}; b)] = E [L(Y_t, X_{1t}; b)] - E [L(Y_t, X_{2t}; b)] + \frac{1 - \phi}{\phi} E [(C(X_{1t}; b) - C(X_{2t}; b)) Y_t] \]

(ii) To implement this method, we need a consistent estimator of \( \phi \). I obtain this via an instrumental variables estimator:

\[
Cov[Y_t, Y_{t+1}] = E \left[ \theta_t \hat{\theta}_{t+1} \right] - E \left[ \hat{\theta}_t \right]^2 = E [(\theta_t + \nu_t) (\theta_{t+1} + \nu_{t+1})] - E [(\theta_t + \nu_t)^2] = E [\theta_t \theta_{t+1}] + E [\nu_t \nu_{t+1}] + E [\nu_{t+1} \theta_t] + E [\nu_t \theta_{t+1}] - (E [\theta_t] + E [\nu_t])^2 = E [\theta_t \theta_{t+1}] - E [\theta_t]^2 + E [\nu_t (\mu + \phi (\theta_t - \mu) + \eta_{t+1})] = E [\theta_t \theta_{t+1}] - E [\theta_t]^2 + E [\nu_t] \mu + \phi E [\nu_t (\theta_t - \mu)] + E [\nu_t \eta_{t+1}] = E [\theta_t \theta_{t+1}] - E [\theta_t]^2 \equiv Cov[\theta_t, \theta_{t+1}] = \phi V[\theta_t]
\]

Similarly I find

\[
Cov[Y_t, Y_{t+2}] = E \left[ \theta_t \hat{\theta}_{t+2} \right] - E \left[ \hat{\theta}_t \right]^2 = Cov[\theta_t, \theta_{t+2}] = \phi^2 V[\theta_t]
\]

and so

\[
\frac{Cov[Y_t, Y_{t+2}]}{Cov[Y_t, Y_{t+1}]} = \frac{\phi^2 V[\theta_t]}{\phi V[\theta_t]} = \phi
\]

A consistent estimator of \( \phi \) is thus

\[
\hat{\phi}_T \equiv \frac{\frac{1}{T-2} \sum_{t=1}^{T-2} Y_{t+2} Y_t - \left( \frac{1}{T} \sum_{t=1}^{T} Y_t \right)^2}{\frac{1}{T-1} \sum_{t=1}^{T-1} Y_{t+1} Y_t - \left( \frac{1}{T} \sum_{t=1}^{T} Y_t \right)^2}
\]

\[
\rightarrow_{p} \frac{Cov[Y_t, Y_{t+2}]}{Cov[Y_t, Y_{t+1}]}, \text{ as } T \rightarrow \infty
\]

\[
= \phi
\]
<table>
<thead>
<tr>
<th>$h$</th>
<th>MSE</th>
<th>QLIKE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RV</td>
<td>RVAC1</td>
</tr>
<tr>
<td>0.5</td>
<td>-13.71</td>
<td>-9.04</td>
</tr>
<tr>
<td>1</td>
<td>-12.49</td>
<td>-5.66</td>
</tr>
<tr>
<td>2</td>
<td>-10.24</td>
<td>-8.21</td>
</tr>
<tr>
<td>3</td>
<td>-8.86</td>
<td>-6.79</td>
</tr>
<tr>
<td>5</td>
<td>-9.72</td>
<td>-8.86</td>
</tr>
<tr>
<td>6</td>
<td>-8.00</td>
<td>-7.21</td>
</tr>
<tr>
<td>10</td>
<td>-8.77</td>
<td>-8.26</td>
</tr>
<tr>
<td>13</td>
<td>-7.39</td>
<td>-8.50</td>
</tr>
<tr>
<td>15</td>
<td>-8.60</td>
<td>-9.22</td>
</tr>
<tr>
<td>26</td>
<td>-8.67</td>
<td>-8.01</td>
</tr>
<tr>
<td>30</td>
<td>-10.1</td>
<td>-10.20</td>
</tr>
<tr>
<td>65</td>
<td>-6.48</td>
<td>-9.58</td>
</tr>
<tr>
<td>78</td>
<td>-6.79</td>
<td>-8.87</td>
</tr>
<tr>
<td>130</td>
<td>-7.66</td>
<td>-7.41</td>
</tr>
<tr>
<td>195</td>
<td>-8.18</td>
<td>n/a</td>
</tr>
<tr>
<td>390</td>
<td>0.00</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Notes: This table presents the estimated mean of the difference between the loss incurred using a given RV or RVAC1 estimator and the loss incurred using the squared open-to-close return, using the bias-adjustment from Proposition 5. The best forecast for a given loss function is in bold; the second-best is in italics.
Table 2: Average loss relative to daily squared returns, under the RW assumption

<table>
<thead>
<tr>
<th>RW</th>
<th>( MSE )</th>
<th>( QLIKE )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h )</td>
<td>RV</td>
<td>RVAC1</td>
</tr>
<tr>
<td>0.5</td>
<td>-9.85</td>
<td>-16.65</td>
</tr>
<tr>
<td>1</td>
<td>-14.1</td>
<td>-16.85</td>
</tr>
<tr>
<td>2</td>
<td>-14.66</td>
<td><strong>-17.21</strong></td>
</tr>
<tr>
<td>3</td>
<td>-16.61</td>
<td>-16.46</td>
</tr>
<tr>
<td>5</td>
<td>-17.13</td>
<td>-16.74</td>
</tr>
<tr>
<td>6</td>
<td>-16.53</td>
<td>-16.44</td>
</tr>
<tr>
<td>10</td>
<td>-16.75</td>
<td>-16.07</td>
</tr>
<tr>
<td>13</td>
<td>-16.73</td>
<td>-15.83</td>
</tr>
<tr>
<td>15</td>
<td>-17.01</td>
<td>-16.08</td>
</tr>
<tr>
<td>26</td>
<td>-16.81</td>
<td>-14.34</td>
</tr>
<tr>
<td>30</td>
<td><strong>-17.36</strong></td>
<td>-15.08</td>
</tr>
<tr>
<td>39</td>
<td>-14.89</td>
<td>-13.57</td>
</tr>
<tr>
<td>65</td>
<td>-13.57</td>
<td>-12.97</td>
</tr>
<tr>
<td>78</td>
<td>-12.78</td>
<td>-11.97</td>
</tr>
<tr>
<td>195</td>
<td>-10.83</td>
<td>n/a</td>
</tr>
<tr>
<td>390</td>
<td>0.00</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Notes: This table presents the estimated mean of the difference between the loss incurred using a given RV or RVAC1 estimator and the loss incurred using the squared open-to-close return, assuming that the target variable follows a random walk. The best forecast for a given loss function is in bold; the second-best is in italics.
Table 3: Average loss relative to daily squared returns, ignoring correlation in measurement errors

<table>
<thead>
<tr>
<th>Naïve</th>
<th>MSE</th>
<th>QLIKE</th>
</tr>
</thead>
<tbody>
<tr>
<td>h</td>
<td>RV</td>
<td>RVACT</td>
</tr>
<tr>
<td>0.5</td>
<td>19.42</td>
<td>19.24</td>
</tr>
<tr>
<td>1</td>
<td>17.42</td>
<td>18.60</td>
</tr>
<tr>
<td>2</td>
<td>16.53</td>
<td>16.19</td>
</tr>
<tr>
<td>3</td>
<td>16.06</td>
<td>15.56</td>
</tr>
<tr>
<td>5</td>
<td>16.15</td>
<td>14.43</td>
</tr>
<tr>
<td>6</td>
<td>15.15</td>
<td>15.10</td>
</tr>
<tr>
<td>10</td>
<td>14.94</td>
<td>13.68</td>
</tr>
<tr>
<td>13</td>
<td>15.78</td>
<td>13.17</td>
</tr>
<tr>
<td>15</td>
<td>15.46</td>
<td>13.22</td>
</tr>
<tr>
<td>26</td>
<td>14.43</td>
<td>11.70</td>
</tr>
<tr>
<td>30</td>
<td>15.36</td>
<td>12.02</td>
</tr>
<tr>
<td>39</td>
<td>12.95</td>
<td>10.11</td>
</tr>
<tr>
<td>65</td>
<td>11.84</td>
<td>9.35</td>
</tr>
<tr>
<td>78</td>
<td>10.49</td>
<td>7.84</td>
</tr>
<tr>
<td>130</td>
<td>9.76</td>
<td>4.81</td>
</tr>
<tr>
<td>195</td>
<td>7.52</td>
<td>n/a</td>
</tr>
<tr>
<td>390</td>
<td>0.00</td>
<td>n/a</td>
</tr>
</tbody>
</table>

Notes: This table presents the estimated mean of the difference between the loss incurred using a given RV or RVAC1 estimator and the loss incurred using the squared open-to-close return, ignoring the correlation between the measurement error in the proxy and the error in the RV estimators. The best forecast for a given loss function is in bold; the second-best is in italics.
Table 4: Optimal RV estimators

<table>
<thead>
<tr>
<th>DGP assumption</th>
<th>AR(1)</th>
<th>RW</th>
<th>Naive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>RV</td>
<td>h = 0.5</td>
<td>RV</td>
</tr>
<tr>
<td>2nd best</td>
<td>RV</td>
<td>h = 1</td>
<td>RVAC1</td>
</tr>
<tr>
<td>2nd worst</td>
<td>RVAC1</td>
<td>h = 1</td>
<td>RVAC1</td>
</tr>
<tr>
<td>Worst</td>
<td>RV</td>
<td>daily</td>
<td>RV</td>
</tr>
</tbody>
</table>

Notes: This table presents the best, second-best, worst and second-worst RV estimators across the 17 standard RV and 16 RVAC1 estimators considered, according to the MSE and QLIKE loss functions. Three assumptions on the DGP are considered: the AR(1) assumption corresponds to Assumption T2”, while the RW assumption corresponds to T2. The “naïve” case incorrectly assumes zero correlation between the proxy error and the RV estimators. In all cases a one-period lead of the squared open-to-close return was used as the instrument.
Table 5: P-values from “reality check” tests

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>DGP</th>
<th>Proxy</th>
<th><strong>MSE</strong></th>
<th></th>
<th><strong>QLIKE</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>White Hansen</td>
<td></td>
<td>White Hansen</td>
<td></td>
</tr>
<tr>
<td>Daily</td>
<td>AR(1)</td>
<td>Daily</td>
<td>0.166</td>
<td><strong>0.096</strong></td>
<td>0.282</td>
<td>0.224</td>
</tr>
<tr>
<td>Daily</td>
<td>AR(1)</td>
<td>130-min RV</td>
<td>0.107</td>
<td><strong>0.039</strong></td>
<td>0.388</td>
<td>0.126</td>
</tr>
<tr>
<td>Daily</td>
<td>RW</td>
<td>Daily</td>
<td><strong>0.025</strong></td>
<td>0.014</td>
<td><strong>0.000</strong></td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td>Daily</td>
<td>RW</td>
<td>130-min RV</td>
<td><strong>0.047</strong></td>
<td><strong>0.023</strong></td>
<td><strong>0.006</strong></td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td>5-min RV</td>
<td>AR(1)</td>
<td>Daily</td>
<td>0.684</td>
<td>0.656</td>
<td>0.933</td>
<td>0.248</td>
</tr>
<tr>
<td>5-min RV</td>
<td>AR(1)</td>
<td>130-min RV</td>
<td>0.893</td>
<td>0.742</td>
<td>0.949</td>
<td>0.215</td>
</tr>
<tr>
<td>5-min RV</td>
<td>RW</td>
<td>Daily</td>
<td>0.315</td>
<td>0.658</td>
<td>0.939</td>
<td><strong>0.025</strong></td>
</tr>
<tr>
<td>5-min RV</td>
<td>RW</td>
<td>130-min RV</td>
<td>0.820</td>
<td><strong>0.073</strong></td>
<td>0.969</td>
<td>0.191</td>
</tr>
<tr>
<td>GARCH</td>
<td>AR(1)</td>
<td>Daily</td>
<td>0.454</td>
<td>0.398</td>
<td>0.933</td>
<td>0.331</td>
</tr>
<tr>
<td>GARCH</td>
<td>AR(1)</td>
<td>130-min RV</td>
<td>0.686</td>
<td>0.459</td>
<td>0.975</td>
<td>0.417</td>
</tr>
<tr>
<td>GARCH</td>
<td>RW</td>
<td>Daily</td>
<td><strong>0.058</strong></td>
<td>0.293</td>
<td>0.981</td>
<td>0.643</td>
</tr>
<tr>
<td>GARCH</td>
<td>RW</td>
<td>130-min RV</td>
<td>0.609</td>
<td>0.148</td>
<td>1.000</td>
<td>0.643</td>
</tr>
</tbody>
</table>

Notes: This table presents the $p$-values from the reality check of White (2000), and those from Hansen’s (2005) refinement of the reality check, under MSE and QLIKE loss. Two assumptions on the DGP for the target variable: the AR(1) assumption corresponds to Assumption T2”, while the RW assumption corresponds to T2. Three benchmark forecasts are considered: RV based on $h = 390$ (denoted “daily”), RV based on $h = 5$ minutes, and the estimated volatility produced by a GARCH(1,1) model estimated on the full sample of open-to-close returns. Two variables were considered as the proxy: the squared open-to-close return (denoted “daily”) and standard realised variance based on 3 returns per day (denoted “130-min RV”). In both cases the instrument was the one-period lead of the proxy variable. A $p$-value of less than 0.10 indicates that the benchmark RV estimator is significantly beaten by one of the competing RV estimators at the 10% level.
References


Figure 1: Volatility signature plot (top panel) for RV and RV-AC1 estimators, and plot of RV and RV-AC1 standard deviation (lower panel). The 95% confidence interval (CI) in the upper panel is a CI for the daily squared return at the far right, and a CI for the mean difference between the daily squared return and the RV(m) or RVAC1(m) estimator for the remaining values of m.
Figure 2: Differences in average loss across sampling frequencies for the standard RV estimator and the RV-AC1 estimator, for MSE and QLIKE loss, according to different assumptions about the DGP. A negative loss differential means that the estimator out-performs the 5-min RV estimator.