BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2013-49

A Distributional Approach to Realized Volatility

by Selma Chaker and Nour Meddahi

www.bank-banque-canada.ca

Bank of Canada Working Paper 2013-49

December 2013

A Distributional Approach to Realized Volatility

by

Selma Chaker[1] and Nour Meddahi[2]

[1]International Economic Analysis Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
schaker@bankofcanada.ca

[2]Toulouse School of Economics
31000 Toulouse, France
nour.meddahi@tse-fr.eu

ISSN 1701-9397

Acknowledgements

We thank Bruno Feunou for providing us with the data used in this paper. We express our gratitude to Sílvia Gonçalves for valuable feedback. We also thank all the participants of the Humboldt-Copenhagen Conference on Financial Econometrics, Berlin, March 2009. All remaining errors are our own.

Abstract

This paper proposes new measures of the integrated variance, measures which use high-frequency bid-ask spreads and quoted depths. The traditional approach assumes that the mid-quote is a good measure of frictionless price. However, the recent high-frequency econometric literature takes the mid-quote as a noisy measure of the frictionless price and proposes new and robust estimators of the integrated variance. This paper forgoes the common assumption of an additive friction term, and demonstrates how the quoted depth may be used in the construction of refined realized volatility measures under the assumption that the true frictionless price lies between the bid and the ask. More specifically, we make assumptions about the conditional distribution of the frictionless price given the available information, including quotes and depths. This distributional assumption leads to new measures of the integrated variance that explicitly incorporate the depths. We then empirically compare the new measures with the robust ones when dealing with forecasting integrated variance or trading options. We show that, in several cases, the new measures dominate the traditional measures.

JEL classification: C14, C51, C58
Bank classification: Econometric and statistical methods; Financial markets

Résumé

Les auteurs proposent de nouvelles mesures de la variance intégrée qui reposent sur des données à haute fréquence concernant les écarts entre cours acheteur et vendeur et les profondeurs affichées, c.-à-d. les quantités offertes à ces deux cours. Dans l'approche traditionnelle, le cours médian est considéré comme un bon indicateur du prix sans frictions. Or les travaux économétriques récents ayant mis à contribution ce type de données voient dans le cours médian une mesure entachée de bruit et proposent plutôt l'emploi de nouveaux estimateurs de la variance intégrée qui sont robustes. Dans la présente étude, les auteurs n'incluent pas de terme de friction de type additif, contrairement aux modèles classiques, et démontrent que les profondeurs affichées peuvent servir à construire des mesures plus précises de la volatilité réalisée en partant de l'hypothèse que le prix sans frictions se situe entre les cours vendeur et acheteur. Plus précisément, ils formulent des hypothèses au sujet de la distribution conditionnelle du prix sans frictions étant donné l'information disponible, dont les cours et profondeurs affichés. Grâce à ces hypothèses, il est possible d'élaborer des mesures de la variance intégrée qui tiennent compte explicitement des profondeurs affichées. Les auteurs comparent ensuite les nouvelles mesures avec des estimateurs robustes afin d'en évaluer l'apport empirique pour la prévision de la variance intégrée ou les transactions d'options. Dans plusieurs cas, les nouvelles mesures surpassent celles couramment utilisées.

Classification JEL : C14, C51, C58
Classification de la Banque : Méthodes économétriques et statistiques; Marchés financiers

1 Introduction

Measuring volatility using high-frequency data has attracted growing interest since the late 1990s, for many reasons. First, thanks to the increased availability of large data samples, we can observe almost continuous data processes, which in turn justifies the use of the continuous time framework. The Trades and Quotes (TAQ) database[1] usually releases one-second frequency prices and quotes, but recently it has been releasing one-millisecond frequency data. Such an ultra-high-frequency data set opens up research opportunities to explore intraday volatility features and spot volatility estimation. The second major reason for the growing interest is that the model-free approach of the theory of quadratic variation is not vulnerable to model misspecification, as is the case with other approaches from the parametric literature.

In this paper, we are interested in measuring the integrated variance of asset returns using bid and ask prices. We assume conditional distributional assumptions on the frictionless price. The common approach is to assume that the mid-quote price – the bid and ask prices average – is the sum of the frictionless price and a noise term. By making assumptions on the noise, one could derive consistent estimators of the integrated variance; see, for example, Zhang et al. (2005), Zhang (2006), Barndorff-Nielsen et al. (2008), and Jacod et al. (2009). Early assumptions hypothesized an exogenous independent and identically distributed (i.i.d.) dynamic for the noise. Later on, this assumption was relaxed to allow for some forms of endogeneity with the frictionless price and an autocorrelated noise. The problem is that, since noise is not observed it is difficult to be precise about its time-varying characteristics.

This paper follows a novel approach. As a first attempt, we derive bounds on the integrated variance when assuming that the frictionless price lies between the bid and the ask prices. Such a non-point identification (also known as partial identification) approach was initiated by Manski (2003) and later surveyed by Tamer (2010). Unfortunately, this approach leads to wide bounds, implying that one needs to

[1]The TAQ database is a collection of intraday trades and quotes for all securities listed on the New York Stock Exchange, American Stock Exchange, Nasdaq National Market System and SmallCap issues. TAQ provides historical tick-by-tick data of all stocks listed on NYSE back to 1993.

make additional assumptions. Our main approach consists of making distributional assumptions on the frictionless price conditioned on quoted data (the bid, ask and depths). We then derive new realized volatility measures. One important feature of the new measures is the explicit presence of the bid-ask spread variable. So far, in market microstructure theory, the spread has been only implicitly shown to impound information about volatility; see Hasbrouck (1999) for a bid and ask model that features ARCH volatility effects.

We consider a range of distributions and evaluate them through three criteria. First, we analyze the ability to capture noise at high frequency using the signature plot that was first introduced in Andersen et al. (1999). Indeed, the signature plot – drawing the realized variance against sampling frequencies – gives an idea of the magnitude of the noise in the data for each sampling frequency. Second, we examine the forecast performance of the integrated volatility measures that we propose, both in-sample and out-of-sample. For instance, Andersen et al. (2003) evaluate the forecasting abilities of the standard realized variance, and Aït-Sahalia and Mancini (2008) study the forecasting of integrated volatility using the robust-to-noise estimator: the two time-scales estimator. Third, we use the proposed integrated volatility measures to quantify the pecuniary gain or loss for option pricing in a hypothetical market, as in Bandi et al. (2008). We show that some new measures outperform the existing ones.

We carry out our analysis by adding the quoted depths (the ask volume and the bid volume) to the conditioning information set. The ask (bid) volume is the maximum number of shares to buy (sell) at the ask (bid) price. The quoted depths reveal information about the stock liquidity and inventory control. For instance, Kavajecz (1999) shows that changes in quoted depths are consistent with market makers managing their inventory as well as having expectations about the stock's future value. Consequently, using the depths may lessen the microstructure frictions effect. The bounded distributions for the frictionless price that we use are the uniform and the triangular over the bid-ask interval. We also accommodate the normal distribution to a bounded support. We explicitly model the correlation between successive prices.

In the empirical section, we use data from the Alcoa stock traded on the New York Stock Exchange during the January 2009 – March 2011 period. We find that the best measures stemming from the forecasting exercise are different from those based on the option-trading exercise. Moreover, the new realized measures can outperform the traditional robust-to-noise volatility estimators.

The rest of this paper is structured as follows. In section 2, we present the common realized measures, the Mincer-Zarnowitz regression for forecasting evaluation, and the option-trading exercise. In section 3, we use the bid-ask prices as bounds for the unobserved frictionless price, in order to derive bounds for the traditional realized volatility measure. Section 4 states the distributional assumptions and the new volatility measures that they imply. We also assess the forecasting performance of each new realized measure. We explore the value of the volume information in section 5. Finally, section 6 concludes.

2 The Forecasting Performance of the Realized Measures

In this paper, we focus on daily integrated volatility. This frequency is of interest to analyze daily patterns of volatility and option pricing, as in Hansen and Lunde (2006) and Bandi et al. (2008). Weekly or monthly frequencies could also be examined using the same framework.

In what follows, t stands for the day. We observe a sample of size N of intraday bids and asks denoted b_{t-1+ih}, a_{t-1+ih} in log terms, where $i = 1..N$ and h is the sampling frequency. The logarithm of the frictionless price is latent and denoted p_{t-1+ih}. Throughout the paper, we simplify the notation by letting b_i, a_i and p_i refer to $b_{t-1+i/N}$, $a_{t-1+i/N}$ and $p_{t-1+i/N}$, respectively. The intraday return is given by

$$r_i = p_i - p_{i-1}. \tag{1}$$

We suppose that the frictionless price follows a semimartingale given by

$$dp_s = \mu_s ds + \sigma_s dW_s, \tag{2}$$

where W_s is a Wiener process and σ_t is a *càdlàg* volatility function. The object of interest is the integrated variance for a given day, defined as

$$IV_t = \int_{t-1}^{t} \sigma_s^2 ds. \tag{3}$$

The realized variance is defined as

$$RV_t(h) = \sum_{i=1}^{1/h} r_{t-1+ih}^2, \tag{4}$$

where $r_{t-1+ih} = p_{t-1+ih} - p_{t-1+(i-1)h}$. The realized variance computed with the highest-frequency returns would be a consistent estimator of the integrated variance if the observed price were equal to the frictionless price; see Jacod (1994), and Barndorff-Nielsen and Shephard (2002) for a proof of the consistency and the limit theory of the realized volatility.

Let m_{t-1+ih} and s_{t-1+ih} denote the mid-quotes and the spread, respectively. Thus we have

$$m_{t-1+ih} = \frac{a_{t-1+ih} + b_{t-1+ih}}{2}, \tag{5}$$

and

$$s_{t-1+ih} = a_{t-1+ih} - b_{t-1+ih}. \tag{6}$$

In this paper, we make assumptions about the distribution of the frictionless price conditioning on the quotes data. These data include bid, ask prices and quoted depths. The ask (bid) depth specifies the maximum quantity for which the ask (bid) price applies. Such an assumption does not conflict with the previous semimartingale assumption for the price. We work in a discrete time setting, since we make a distributional hypothesis about successive intraday prices.

2.1 The common realized measures

The realized variance defined in equation (4) is an inconsistent estimator of the integrated variance because of the market microstructure noise that contaminates frictionless prices. Empirical evidence of the noise is the signature plot introduced by Andersen et al. (1999). The signature plot draws a sample average of the daily realized measure of volatility as a function of the underlying returns sampling frequency. A graph that explodes at high frequencies is evidence of market microstructure noise severity. At low frequencies, the plot converges to the integrated variance measure and the noise effect disappears.

If the highest-frequency returns are used to compute the realized variance, we obtain the RV_t^{all} estimator given by

$$RV_t^{all} = \sum_{i=1}^{N} r_i^2. \tag{7}$$

Figure 1 shows the signature plot of the realized variance. We use Alcoa data covering the January 2009 – March 2011 period. We indeed notice that the noise causes a considerable bias at high frequencies.

If a low-frequency $\bar{h}$ is used to compute the returns, we obtain the following estimator:

$$RV_t^{low} = RV_t(\bar{h}) = \sum_{i=1}^{1/\bar{h}} r_{t-1+i\bar{h}}^2. \tag{8}$$

The two time-scales estimator of Zhang et al. (2005), which is consistent under i.i.d. market microstructure noise, is defined as

$$RV_t^{TS} = RV_t^{average} - \frac{\overline{N}}{N} RV_t^{all}, \tag{9}$$

where the average estimator $RV^{average}$ is the mean of several sparse estimators. Formally, it is given by

$$RV_t^{average} = \frac{1}{K} \sum_{k=1}^{K} RV_t^{(k)}, \tag{10}$$

where $RV_t^{(k)} = \sum_{i=1}^{N-k+1} r_{t-1+(i+k-1)h}^2$, and h is the sampling frequency.

The kernel estimator of Barndorff-Nielsen et al. (2008) achieves a faster rate of convergence than the RV^{TS}, and is defined as

$$RV_t^{kernel} = \gamma_0 + \sum_{l=1}^{L} f\left(\frac{l-1}{L}\right) \{\gamma_l + \gamma_{-l}\}, \tag{11}$$

where $\gamma_l = \sum_{j=1}^{N} r_{t-1+jh} r_{t-1+(j-l)h}$, $f(x) = (1 - cos\pi(1-x)^2)/2$, and L is the bandwidth.

The pre-averaging estimator introduced by Jacod et al. (2009) is robust to heteroscedastic noise, and achieves an optimal rate of convergence. We denote RV_t^{pre} the pre-averaging estimator, given by

$$RV_t^{pre} = \sum_{i=0}^{N-k} \left\{ \sum_{j=1}^{k} \phi\left(\frac{j}{k}\right) r_{i+j} \right\}^2 - \frac{6}{\theta^2} RV_t^{all},$$

where $\frac{k}{\sqrt{N}} = \theta + \mathcal{O}(N^{-1/4})$ for some $\theta > 0$, and $\phi(x) = min(x, 1-x)$.

2.2 Mincer-Zarnowitz regression

In order to assess the forecasting performance of a realized measure RM_t, we use a Mincer and Zarnowitz (1969) regression, given by

$$IV_{t+1} = \alpha + \beta RM_{t+1|t} + \eta_{t+1}, \tag{12}$$

where $RM_{t+1|t}$ is the forecast at time t of IV_{t+1} using the measure RM, and η_{t+1} is an error term. A forecast $RM_{t+1|t}$ is good if $\alpha = 0$, $\beta = 1$ and there is a high R^2. In the empirical results, we report only the R^2 because we find that none of our realized measures results in both an α significantly not different from 0 and a β significantly not different from 1. Since the dependent variable is latent, we use RV_{t+1}^{low} as a proxy for IV_{t+1}. For the longer forecasting horizon H, the Mincer-Zarnowitz regression is given by

$$IV_{t:t+H} = \alpha + \beta RM_{t+H|t} + \eta_{t+H}, \tag{13}$$

where $IV_{t:t+H} = \int_t^{t+H} \sigma_s^2 ds$. The term $RM_{t+H|t}$ refers to the forecast at time t of $IV_{t:t+H}$ using the measure RM. We denote the error term η_{t+H}.

The forecasting model that we use is an AR(3) with RV^{low} as the dependent variable and a 100-day rolling window. We conduct both an in-sample and out-of-sample forecasting exercise. In Table 1, we report the R^2 for the in-sample and out-of-sample forecasts for a one-day and 5-day horizon. The results show that the pre-averaging estimator achieves the highest R^2 whether in-sample or out-of-sample for the short forecasting horizon. In addition, the realized variance RV^{all} has the least R^2. However, the R^2 is close for the estimators RV^{all}, RV^{TS}, RV^{kernel} and RV^{pre}. For the longer horizon of five days, we find that the overall forecasting performance of the four estimators has improved upon the short horizon. Finally, the RV^{TS} becomes the best forecast whether in-sample or out-of-sample when the forecasting horizon equals five days.

2.3 Option trading

In this section, we evaluate the proposed integrated volatility estimates in the context of the profits from option pricing and trading. Using alternative forecasts obtained in the previous section, agents price short-term options on Alcoa stock before trading with each other at average prices. The average profits are used as the criteria to evaluate alternative volatility estimates and the corresponding forecasts. We construct a hypothetical option market as in Bandi, Russell and Yang (2008) in order to quantify the economic gain or loss for using alternative integrated volatility forecasts. The methodology was first proposed by Engle et al. (1990). Consider a simple market of only two traders such that each one uses a different volatility forecast. The trades are conducted at the midpoint of the two traders' prices. The trader with the highest-volatility forecast will buy a call and a put option from a counterpart. The average dollar profits obtained from the trading represent the metric used to evaluate alternative variance forecasts. If the high-volatility forecast is accurate, the straddle is underpriced. Then, the trader who buys the straddle is expected to make a profit.

In this paper, our artificial market has as many traders as alternative forecasts. Each trader uses a different measure from the set of realized measures.

First, each trader constructs an out-of sample, one-day-ahead variance forecast using daily variances series and computes a predicted Black-Scholes option price. We focus on an at-the-money price of a one-day or 5-day option on a one-dollar share of Alcoa. The risk-free rate is taken to be zero.

Second, the pairwise trades take place. For two given traders, if the forecast of the first is higher than the midpoint of the forecasts of the two traders, then the option is perceived as underpriced. Thus the first trader will buy a straddle (one call and one put) from a counterpart. Then the positions are hedged using the deltas of the options.

Finally, we compute the profits or losses. Each trader averages the profits or losses from the pairwise trading. We report the average profits across all days in the sample.

The option trading and profit results are computed as in the following three steps:

1- Compute the option price. Let σ_t denote the volatility forecast for a given measure. The Black-Scholes option price P_t is given by

$$P_t = 2\Phi(\frac{1}{2}\sigma_t) - 1,$$

where Φ is the cumulative normal distribution.

2- Compute the profit for each trader. The daily profit for a trader who buys the straddle is

$$| R_t | -2P_t + R_t(1 - 2\Phi(\frac{1}{2}\sigma_t)),$$

where the last term corresponds to the hedging, and R_t is the daily return for day t. The daily profit for a trader who sells the straddle is

$$2P_t - | R_t | -R_t(1 - 2\Phi(\frac{1}{2}\sigma_t)).$$

3- Average the profits and obtain the metric.
We use the out-of-sample forecasts of section 2.2. The profits in cents are reported in Table 2 when all the realized measures are used in the trading game. We find that the agents using the four traditional measures RV^{all}, RV^{TS}, RV^{kernel} and RV^{pre} endure losses. For the one-day horizon, the RV^{pre} is the worst estimator, whereas it becomes the best at the 5-day horizon. The inverse is observed for RV^{all}, where it is the best at the short horizon but the worst at the long horizon.

3 Simple Bounds

We make the assumption that the frictionless price is bounded by the bid and the ask. This assumption is restrictive, since the frictionless price could be less than the bid or higher than the ask. Formally, we have
Assumption A $b_i \leq p_i \leq a_i$, $i = 1...N$.

Under Assumption A, we derive bounds for the realized variance in the following proposition.

Proposition 1 *Under Assumption A,*

$$RV_t^{inf} \leq RV_t \leq RV_t^{sup},$$

where

$$\begin{aligned} RV_t^{sup} &= \sum_{i=1}^{1/h} \overline{r}_i^2, \\ RV_t^{inf} &= \sum_{i=1}^{1/h} \underline{r}_i^2, \end{aligned} \tag{14}$$

and

$$\overline{r}_i^2 = Max\left\{(b_i - a_{i-1})^2; (a_i - b_{i-1})^2\right\},$$
$$\underline{r}_i^2 = \left\{ \begin{array}{ll} 0 & if\ (b_i - a_{i-1})(a_i - b_{i-1}) \leq 0 \\ Min\left\{(b_i - a_{i-1})^2; (a_i - b_{i-1})^2\right\} & else. \end{array} \right\} \quad (15)$$

The bounds derived in Proposition 1 are not tight to be informative for the volatility level. Indeed, they are based on very weak assumptions. We draw the signature plot of RV^{inf} and RV^{sup} in Figure 2. At high frequencies, the interval $[RV^{inf}, RV^{sup}]$ is very wide; at low frequencies, this interval becomes narrow. For the forecasting results, we find in Table 1 that the bound RV^{sup} beats all the traditional measures RV^{all}, RV^{TS}, RV^{kernel} and RV^{pre}, whether in-sample, out-of-sample, short- or long-horizon forecasting. The lower-bound RV^{inf} has the worst results compared to the traditional measures at the short horizon, but beats them at the long forecasting horizon. The profits from the option-trading exercise are reported in Table 2. At the long horizon, the upper-bound RV^{sup} achieves a significant profit, whereas the lower-bound RV^{inf} has a considerable loss.

We also report in Table 3 the profits from option trading when the agents use the traditional measures RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and the new measure RV^{inf}. The objective of this alternative game is that by restricting the participating agents to only those using the traditional measures and RV^{inf}, we actually evaluate the performance of RV^{inf} compared to the traditional measures only, instead of all the measures of this paper. We find that the agent using RV^{inf} endures losses both at the short and long horizon and is dominated by the traditional measures. In Table 4, we report the results for the option trading between the agents using the aforementioned four measures in addition to RV^{sup}. This new measure beats the traditional measures at the 5-day horizon. However, the agent using RV^{sup} endures more losses than the agents using the traditional measures.

In the following section, our goal is to examine some distributional restrictions for the price. In fact, imposing many more restrictions may provide better volatility estimators.

4 A Distributional Approach

In this section, we impose more restrictions on the distribution of the price. We evaluate the new realized measures using signature plots, the Mincer-Zarnowitz regression for forecasting performance and the option-trading outcome.

4.1 A Dirac measure

If we assume that the frictionless price p_i follows a Dirac measure in the mid-quotes, we obtain the usual expression for the realized volatility,

$$RV_t^{all} = \sum_{i=1}^{N} (m_i - m_{i-1})^2. \tag{16}$$

If we assume that the frictionless price p_i follows a Dirac measure in the bid, we obtain the measure

$$RV_t^{bid} = \sum_{i=1}^{N} (b_i - b_{i-1})^2. \tag{17}$$

The same applies for the ask, and the corresponding measure is given by

$$RV_t^{ask} = \sum_{i=1}^{N} (a_i - a_{i-1})^2. \tag{18}$$

The signature plots in Figure 3 of RV_t^{bid} and RV_t^{ask} are very close and more noisy than RV_t^{all} at high frequencies. Table 1 reports the forecasting results, showing that RV_t^{bid} and RV_t^{ask} have a similar forecasting performance to the traditional realized measures. However, the option-trading profits/losses in Table 2 are positive for the 5-day horizon, contrary to the negative outcome of the traditional realized measures. At the short horizon, the traders using RV_t^{bid} and RV_t^{ask} endure losses comparable to those of the traditional measures.

In the option-trading game where only the agents using the traditional measures and RV_t^{bid} play, we find that the bid-based measure beats the traditional measures at the long horizon. The results are reported in Table 5. RV_t^{ask} achieves similar results to RV_t^{ask}, as shown in Table 6.

4.2 Univariate distributions

We take the set of intraday quotes as the conditioning set,

$$I = \{b_j, a_j, j = 1, ..., N\}. \tag{19}$$

The components of the squared return conditional expectation are derived as follows:

$$\begin{aligned} E[r_i^2 \mid I] &= (E[r_i \mid I])^2 + Var[r_i \mid I] \\ &= (E[p_i \mid I] - E[p_{i-1} \mid I])^2 \\ &+ Var[p_i \mid I] + Var[p_{i-1} \mid I] - 2Cov[p_i, p_{i-1} \mid I]. \end{aligned} \tag{20}$$

We make the following assumption.

Assumption B

Conditional on I, r_{t-1+ih} and $p_{t-1+(i-1)h}$ are independent.

Assumption B specifies that any intraday return is conditionally independent from the previous price. In Proposition 2, we use Assumption B and expression (20) to derive the conditional expectation of the realized variance.

Proposition 2 *Under Assumption A and B,*

$$\begin{aligned} &E[RV_t(h) \mid I] \\ &= \sum_{i=1}^{1/h} (E[p_{t-1+ih} \mid I] - E[p_{t-1+(i-1)h} \mid I])^2 + Var[p_{t-1+ih} \mid I] + Var[p_{t-1+(i-1)h} \mid I] \\ &- 2Min\left(\sqrt{\frac{Var[p_{t-1+(i-1)h} \mid I]}{Var[p_{t-1+ih} \mid I]}}, 1\right) \sqrt{Var[p_{t-1+ih} \mid I]}\sqrt{Var[p_{t-1+(i-1)h} \mid I]}. \end{aligned} \tag{21}$$

Equation (21) is only a function of the expectation and the variance. Therefore, by varying the distributional hypothesis regarding the intraday price, we obtain different estimators. We define the resulting realized measure as

$$RM_t = E[RV_t \mid I]. \tag{22}$$

We specifically examine the realized measures based on uniform and triangular distributional assumptions.

4.2.1 The uniform distribution

We make the assumption that the intraday price follows a uniform distribution. Formally,

$$p_i \mid I \sim \ Uniform[b_i, a_i]. \tag{23}$$

The uniform distribution is such that all intervals of the same length on the distribution's support $[b_i, a_i]$ are equally probable. The first two moments are given by

$$\begin{aligned} E[p_i \mid I] &= m_i, \\ Var[p_i \mid I] &= \frac{s_i^2}{12}. \end{aligned} \tag{24}$$

We define $RV^{uniform}$ using equations (21) and (22); thus,

$$RV_t^{uniform} = \sum_{i=1}^{1/h} \left\{ (m_i - m_{i-1})^2 + v_i + v_{i-1} - 2Min\{\sqrt{\frac{v_{i-1}}{v_i}} 1\} \sqrt{v_i}\sqrt{v_{i-1}} \right\}, \tag{25}$$

where $v_i = \frac{s_i^2}{12}$. The bid-ask spread appears in the new realized variance $RV^{uniform}$. The spread is a friction measure that is not yet explored, to our knowledge, in the high-frequency literature on measuring volatility.

The empirical results of $RV^{uniform}$ show a similar forecasting performance to the traditional realized measures, as reported in Table 1. However, in the short horizon, the trader using $RV^{uniform}$ endures the smallest loss among the realized measures introduced so far, as shown in Table 2. At the 5-day horizon, the trader using $RV^{uniform}$ achieves a profit. The signature plots in Figures 1 and 3 show that, at high frequencies, $RV^{uniform}$ is more noisy than the realized variance RV^{all}.

If only the agents using the traditional measures and $RV^{uniform}$ play, this new measure generates net profits for the agent using it, whereas the other agents endure losses at the long horizon, as shown in Table 7.

4.2.2 The triangular distribution

Let the frictionless price follow a centred triangular distribution,

$$p_i \mid I \sim \ Centred\ Triangular[b_i, a_i]. \tag{26}$$

This distribution has an affine probability density function. The mid-quotes is the most probable of the distribution support. The expectation and variance expressions are, respectively,

$$\begin{aligned} E[p_i \mid I] &= m_i, \\ Var[p_i \mid I] &= \frac{b_i^2 + a_i^2 + m_i^2 - b_i a_i - b_i m_i - m_i a_i}{18}. \end{aligned} \tag{27}$$

For the uniform distribution, we define $RV^{triangular}$ using equations (21) and (22) as

$$RV_t^{triangular} = \sum_{i=1}^{1/h} \left\{ (m_i - m_{i-1})^2 + v_i + v_{i-1} - 2Min\{\sqrt{\frac{v_{i-1}}{v_i}}; 1\}\sqrt{v_i}\sqrt{v_{i-1}} \right\}, \quad (28)$$

where

$$v_i = \frac{b_i^2 + a_i^2 + m_i^2 - b_i a_i - b_i m_i - m_i a_i}{18}. \quad (29)$$

The new realized measure $RV^{triangular}$ uses the bid, ask and mid-quotes. Assuming that the frictionless price has a centred triangular distribution means that the mid-quotes is the most probable value for the frictionless price. However, the other values in the bid-ask interval are realized with non-zero probability. In addition, the least probabilities are obtained near the edge of $[b_i, a_i]$.

Empirically, we find that $RV^{triangular}$ is less noisy than the univariate-based measures RV^{bid}, RV^{ask} and $RV^{uniform}$, as shown in the signature plots of Figure 3. The forecasting performance of $RV^{triangular}$, measured by the R^2 of the Mincer-Zarnowitz regression, is similar to the univariate-based measures reported in Table 1. However, the trader using $RV^{triangular}$ endures losses at the 5-day horizon in the option-trading exercise, contrary to the other univariate-based measures (Table 2). At the one-day horizon, $RV^{triangular}$ gives fewer losses than the traditional realized measures.

For the option-trading game where only the agents using the traditional measures and $RV^{triangular}$ play, we also find that this new measure beats the traditional ones at the long horizon; see Table 8 for the results.

4.3 Bivariate distributions

In this section, we do not assume the independent form of successive intraday prices specified by Assumption B. Rather, we specify the joint distribution of each successive intraday price and use general equation (20) to find the realized measure expression. We denote $\rho(h)$ the correlation between two intraday prices p_{t-1+ih} and $p_{t-1+(i-1)h}$. We assume that

$$(i)\rho(.) \in [0,1], decreasing,$$
$$(ii)\rho(0) = 1; lim_{h\to\infty}\rho(h) = 0.$$

Assumption (i) implies that the correlation parameter decreases as the time interval between successive observations becomes larger. At the limit, (ii) assumes a zero correlation if the intraday prices are sampled at a very low frequency.

4.3.1 The bivariate normal distribution

In this section, we assume a joint normal distribution for successive prices. Although the normal distribution does not have a bounded support, we parameterize it to have very slim tails, as if the distribution had almost bounded support (consistent with Assumption A). Formally, we assume that

$$\left[\begin{array}{c} p_i \\ p_{i-1} \end{array}\right] \mid I \sim \mathcal{N}(\left[\begin{array}{c} m_i \\ m_{i-1} \end{array}\right], \left(\begin{array}{cc} v_i & c_{i;i-1} \\ c_{i;i-1} & v_{i-1} \end{array}\right)), \tag{30}$$

where

$$c_{i,i-1} = \rho(h)\sqrt{v_i v_{i-1}}, \tag{31}$$

and

$$\begin{aligned} v_i &= \lambda^2 s_i^2, \\ v_{i-1} &= \lambda^2 s_{i-1}^2. \end{aligned} \tag{32}$$

λ is a constant such that $P[b_i \leq p_i \leq a_i]$=0.99 and $P[b_{i-1} \leq p_{i-1} \leq a_{i-1}]$=0.99, i.e. $\lambda = 0.19$.

We define the measure $RV^{corr.Normal}$, using equations (20) and (22) by,

$$RV^{corr.Normal} = \sum_{i=1}^{1/h} (m_i - m_{i-1})^2 + v_i + v_{i-1} - 2c_{i,i-1}, \tag{33}$$

where the variance and covariance are given in (31) and (32).

The new measure $RV^{corr.Normal}$ is a function of the friction measure (the spread) and the correlation parameter that we specify ad hoc.

The empirical performance of the realized measure $RV^{corr.Normal}$ is similar to the traditional RV^{all} when one looks at the signature plot; the Mincer-Zarnowitz regression results; and the short horizon outcome of the option-trading game. See, respectively, Figures 1 and 4; Table 1; and Table 2. However, the trader using $RV^{corr.Normal}$ endures much less loss than the trader using RV^{all} for the long horizon. When only the agents using the traditional measures and $RV^{corr.Normal}$ trade options, we notice in Table 9 that this new measure beats the traditional ones both at the short and long horizons. The agent using $RV^{corr.Normal}$ always achieves a profit.

4.3.2 The bivariate uniform distribution

We assume the bivariate distribution:

$$\begin{aligned}
&p_i \mid I \sim \ Uniform[b_i, a_i], \\
&p_{i-1} \mid I \sim \ Uniform[b_{i-1}, a_{i-1}], \\
&Cov[p_i, p_{i-1} \mid I] = \frac{\rho(h)}{12} s_i s_{i-1}.
\end{aligned}$$

Using equations (20) and (22), we define the measure $RV^{corr.Uniform}$ by

$$RV^{corr.Uniform} = \sum_{i=1}^{1/h} (m_i - m_{i-1})^2 + \frac{s_i^2}{12} + \frac{s_{i-1}^2}{12} - 2\frac{\rho(h)}{12} s_i s_{i-1}. \tag{34}$$

The signature plot and the forecasting results of $RV^{corr.Uniform}$ are similar to the $RV^{corr.Normal}$, as shown in Figure 4 and Table 1. For the short horizon, $RV^{corr.Uniform}$ achieves the smallest loss compared to the measures introduced so far, including the traditional measures as reported in Table 2. We also notice that the long-horizon loss of $RV^{corr.Uniform}$ is smaller than that for $RV^{corr.Normal}$.
In Table 10 we report the results for the restricted game where only agents using the traditional measures and $RV^{corr.Uniform}$ play. At the long horizon, this measure beats all the traditional ones, but it generates small losses at the short horizon.

4.3.3 The bivariate triangular distribution

We assume that successive intraday prices follow the joint distribution, given by

$$\begin{aligned}
&p_i \mid I \sim \ Triangular\{[b_i, a_i]; m_i\}, \\
&p_{i-1} \mid I \sim \ Triangular\{[b_{i-1}, a_{i-1}]; m_{i-1}\}, \\
&Cov[p_i, p_{i-1} \mid I] = \frac{\rho(h)}{\sqrt{a_i - b_i}\sqrt{a_{i-1} - b_{i-1}}}[\\
&\frac{1}{18}((m_i - b_i)^{3/2}(m_{i-1} - b_{i-1})^{3/2} + (a_i - m_i)^{3/2}(a_{i-1} - m_{i-1})^{3/2}) \\
&+ (\frac{\pi}{8} + \frac{4}{9})((m_i - b_i)^{3/2}(a_{i-1} - m_{i-1})^{3/2} + (a_i - m_i)^{3/2}(m_{i-1} - b_{i-1})^{3/2})].
\end{aligned} \tag{35}$$

Using equations (20) and (22), we define the measure $RV^{corr.Triangular}$ by

$$RV^{corr.Triangular} = \sum_{i=1}^{1/h} (m_i - m_{i-1})^2 + v_i + v_{i-1} - 2Cov[p_i, p_{i-1} \mid I], \tag{36}$$

where

$$v_i = \frac{b_i^2 + a_i^2 + m_i^2 - b_i a_i - b_i m_i - m_i a_i}{18}, \tag{37}$$

and the covariance expression is given in (35).
The signature plot of $RV^{corr.Triangular}$ in Figure 4 shows more bias at high frequencies than the other bivariate uniform and normal distribution-based measures, and even the traditional realized variance RV^{all}. The forecasting performance of $RV^{corr.Triangular}$ as reported in Table 1 is better than $RV^{corr.Normal}$ and $RV^{corr.Uniform}$, whether at short or long horizons, and in-sample or out-of-sample. The trader using $RV^{corr.Triangular}$ achieves the best profit compared to the overall realized measures of this paper for the long horizon, as shown in Table 2.
If only the agents using the traditional measures and $RV^{corr.Triangular}$ play, we still find that the new measure beats the traditional ones at the long horizon, as shown in Table 11.

5 The Volume Information

In order to explore the volume information, we include the intraday quoted depths in conditioning set I.

5.1 The Dirac distribution

We define a weighted volume measure $RV_t^{depths.Weighted}$, as in Gatheral and Oomen (2010), by

$$RV_t^{depths.Weighted} = \sum_{i=1}^{N} \left(\frac{V_i^B a_i + V_i^A b_i}{V_i^A + V_i^B} - \frac{V_{i-1}^B a_{i-1} + V_{i-1}^A b_{i-1}}{V_{i-1}^A + V_{i-1}^B} \right)^2, \tag{38}$$

where V^B (V^A) denotes the bid depth (ask depth).
The forecasting results using Alcoa data in Table 1 show that $RV_t^{depths.Weighted}$ has the highest R^2 whether in-sample or out-of-sample, and for short or long horizons,

among the other Dirac-based measures RV^{all}, RV^{bid} and RV^{ask}. The signature plots depicted in Figures 1, 3 and 5 show evidence that $RV_t^{depths.Weighted}$ is less noisy than RV^{bid} and RV^{ask}, but more noisy than RV^{all} at high frequencies. For the option-trading exercise, $RV_t^{depths.Weighted}$ is the unique realized measure that achieves profits for its user at both short and long horizons (Table 2).

In Table 12 we report the profits for option trading of the agents using the four traditional measures and $RV_t^{depths.Weighted}$ only. Unlike the traditional measures, $RV_t^{depths.Weighted}$ achieves profits both at the short and long horizon.

5.2 The univariate triangular distribution

We assume the price distribution,

$$P_i \mid I \sim \ Non - Centred\ Triangular[b_i, a_i]. \tag{39}$$

We denote c_i the mode, or the most probable value of the distribution support $[b_i, a_i]$. The first moments are then given by

$$\begin{aligned} E[p_i \mid I] &= \frac{a_i + b_i + c_i}{3}, \\ Var[p_i \mid I] &= \frac{b_i^2 + a_i^2 + c_i^2 - b_i a_i - b_i c_i - c_i a_i}{18}. \end{aligned} \tag{40}$$

Using the quoted depths, we incorporate the volume information in the mode expression. Let's denote the volume increments by $\Delta V_i^A = V_i^A - V_{i-1}^A$ and $\Delta V_i^B = V_i^B - V_{i-1}^B$. We set the mode in the following way:

$$c_i = \left\{ \begin{array}{ll} \frac{|\Delta V_i^B|}{|\Delta V_i^A| + |\Delta V_i^B|} b_i + \frac{|\Delta V_i^A|}{|\Delta V_i^A| + |\Delta V_i^B|} a_i & if\ \Delta V_i^A \Delta V_i^B \neq 0 \\ 0.95 b_i + 0.05 a_i & if\ \Delta V_i^A = 0; \Delta V_i^B \neq 0 \\ 0.05 b_i + 0.95 a_i & if\ \Delta V_i^A \neq 0; \Delta V_i^B = 0 \\ 0.5 b_i + 0.5 a_i & if\ \Delta V_i^A = 0; \Delta V_i^B = 0. \end{array} \right\} \tag{41}$$

We then define $RV^{depths.Triangular}$ using equation (20) and by applying Proposition 2:

$$\begin{aligned} &RV^{depths.Triangular} \\ &= \sum_{i=1}^{1/h} \left\{ \left(\frac{a_i + b_i + c_i}{3} - \frac{a_{i-1} + b_{i-1} + c_{i-1}}{3}\right)^2 + v_i + v_{i-1} - 2Min\left(\sqrt{\frac{v_{i-1}}{v_i}}, 1\right) \sqrt{v_i}\sqrt{v_{i-1}} \right\}, \end{aligned} \tag{42}$$

where

$$v_i = \frac{b_i^2 + a_i^2 + c_i^2 - b_i a_i - b_i c_i - c_i a_i}{18}. \tag{43}$$

Observe that the non-centred triangular distribution assumption for the price implies that the most probable value for the frictionless price depends on the quoted depths variations. Incorporating the bid volume and ask volume into the mode expression connects inventory control to volatility estimation. The variations of the quoted depths measure how severe the friction is.

The signature plot of $RV^{depths.Triangular}$ does not beat the traditional realized variance RV^{all} at high frequencies, as shown in Figures 1 and 5. We find that the volume information improves the forecasting ability of the univariate triangular-based measure. Indeed, $RV^{depths.Triangular}$ has higher R^2 than $RV^{triangular}$ for all horizons and both in-sample and out-of-sample, as shown in Table 1. In addition, the improvement in the 5-day out-of-sample performance is less important than that in the one-day. However, the bivariate triangular-based measure $RV^{corr.Triangular}$ beats both $RV^{depths.Triangular}$ and $RV^{triangular}$. Therefore, we introduce in the next section a measure that is based on a bivariate triangular distribution, also incorporating the volume information.

For the option-trading exercise, Table 2 shows that $RV^{depths.Triangular}$ is better at the long horizon because it causes losses at the short horizon. Moreover, the profit that the trader using $RV^{depths.Triangular}$ achieves is less than the one for the trader using $RV^{depths.weighted}$ at the 5-day horizon. Both $RV^{triangular}$ and $RV^{depths.Triangular}$ cause losses at the one-day horizon, but at the long horizon the depths information incorporated in $RV^{depths.Triangular}$ makes it achieve a positive gain to the trader using that measure, compared to the one using $RV^{depths.Triangular}$ who endures losses.

In Table 13, we report the option-trading results for the restricted game where only the agents using the traditional measures and $RV^{depths.Triangular}$ play. At the long horizon, this new measure achieves a profit, whereas the other measures result in losses. But at the short horizon, the agent using $RV^{depths.Triangular}$ endures a loss.

5.3 The bivariate triangular distribution

In this section, we use the volume information and impose a bivariate structure for successive prices. We assume the following triangular distribution for intraday prices:

$$\begin{aligned}
&p_i \mid I \sim \ Triangular\{[b_i, a_i]; c_i\}, \\
&p_{i-1} \mid I \sim \ Triangular\{[b_{i-1}, a_{i-1}]; c_{i-1}\}, \\
&Cov[p_i, p_{i-1} \mid I] = \frac{\rho(h)}{\sqrt{a_i - b_i}\sqrt{a_{i-1} - b_{i-1}}}[\\
&\frac{1}{18}((c_i - b_i)^{3/2}(c_{i-1} - b_{i-1})^{3/2} + (a_i - c_i)^{3/2}(a_{i-1} - c_{i-1})^{3/2}) \\
&+ (\frac{\pi}{8} + \frac{4}{9})((c_i - b_i)^{3/2}(a_{i-1} - c_{i-1})^{3/2} + (a_i - c_i)^{3/2}(c_{i-1} - b_{i-1})^{3/2})],
\end{aligned} \tag{44}$$

where the mode expression is given in (41).
Using equations (20) and (22), we define $RV^{depths.corr.Triangular}$ as

$$\begin{aligned}
&RV^{depths.corr.Triangular} \\
&= \sum_{i=1}^{1/h} \left(\frac{a_i + b_i + c_i}{3} - \frac{a_{i-1} + b_{i-1} + c_{i-1}}{3} \right)^2 + v_i + v_{i-1} - 2Cov[p_i, p_{i-1} \mid I],
\end{aligned} \tag{45}$$

where

$$v_i = \frac{b_i^2 + a_i^2 + c_i^2 - b_i a_i - b_i c_i - c_i a_i}{18},$$

and the covariance is given in (44).
Empirically, the trader using $RV^{depths.corr.Triangular}$ has the best profit among the overall realized measures of this paper (see Table 2) at the long horizon. Therefore, it is important to exploit the volume information as well as a correlated structure of successive intraday prices. However, at the short horizon, the trader using $RV^{depths.corr.Triangular}$ endures a loss. As expected for the forecasting results in Table 1, $RV^{depths.corr.Triangular}$ has the highest R^2 among all the triangular-based measures $RV^{corr.Triangular}$, $RV^{Triangular}$ and $RV^{depths.Triangular}$, whether in-sample or out-of-sample and at short or long horizons. Finally, the signature plot of $RV^{depths.corr.Triangular}$ in Figure 5 shows a more noisy measure at high frequencies than the traditional RV^{all} of Figure 1.
In Table 14, we report the option-trading results for the game where the agents using the four traditional measures and $RV^{depths.corr.Triangular}$ only play. This new measure achieves a large profit at the long horizon compared to the traditional measures that endure losses. However, at the short horizon, $RV^{depths.corr.Triangular}$ does not perform better than the other measures.

6 Conclusion

In this paper, we make distributional assumptions on the frictionless price and obtain new realized measures that incorporate the spread and the quoted depths information. To assess the performance of the new realized measures, we empirically compare their forecasting ability using the Mincer-Zarnowitz regression and an option-trading game. We also measure the noise magnitude at each sampling frequency using the signature plot. For an Alcoa data sample covering January 2009 – March 2011, we find that the new realized measures beat in many cases the common robust-to-noise realized measures. However, there is no clear conclusion regarding the superior performance of the new volatility measures among the competing measures with respect to all the aforementioned criteria.

In practice, it could be that the transaction price lies outside the interval bounded by the bid and the ask prices. In the future, it would be interesting to consider the case where the frictionless price lies outside the bid-ask interval and thus relax the main assumption of this paper.

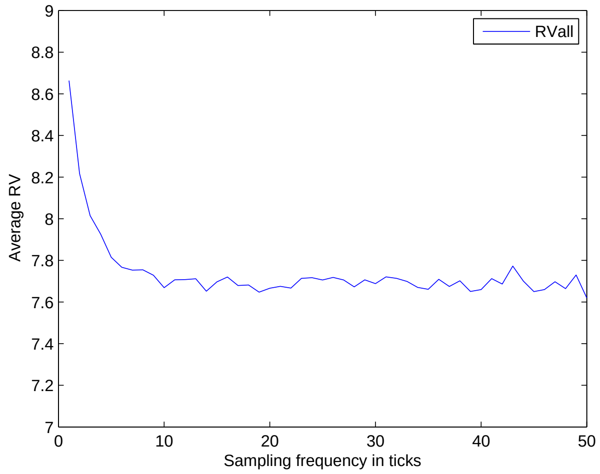

Figure 1: Signature plot for the realized measure RV^{all}. We use the expression given in (16). The intraday returns are computed over the frequencies of the x-axis. The sample period covers 01/2009–03/2011 for Alcoa stock. The reported average is over the business days of our sample.

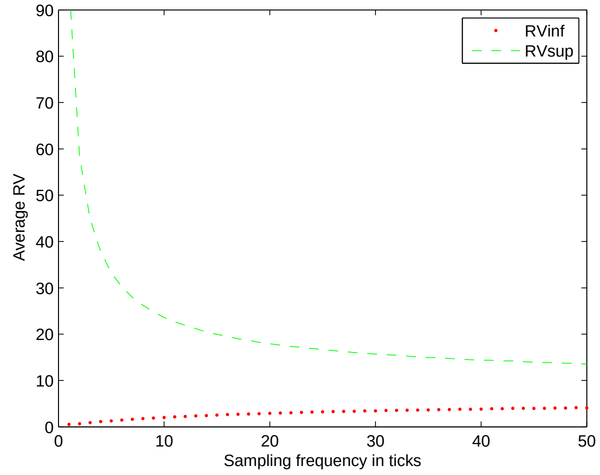

Figure 2: Signature plot for the realized measures RV^{inf} and RV^{sup}. We use the expressions given in (14). The intraday increments are computed over the frequencies of the x-axis. The sample period covers 01/2009–03/2011 for Alcoa stock. The reported average is over the business days of our sample.

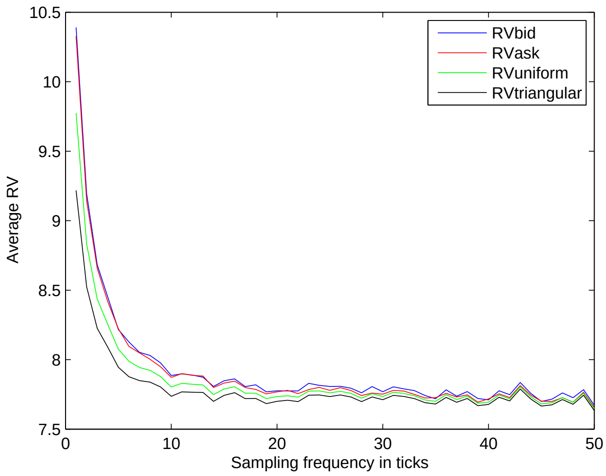

Figure 3: Signature plot for the univariate realized measures RV^{bid}, RV^{ask}, $RV^{Uniform}$, $RV^{Triangular}$. We use the expressions given in (17), (18), (25) and (28). The intraday increments are computed over the frequencies of the x-axis. The sample period covers 01/2009–03/2011 for Alcoa stock. The reported average is over the business days of our sample.

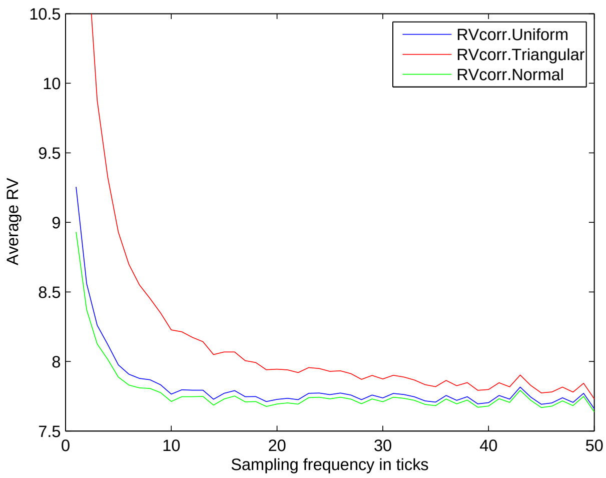

Figure 4: Signature plot for the bivariate realized measures $RV^{corr.Normal}$, $RV^{corr.Uniform}$, $RV^{corr.Triangular}$. We use the expressions given in (33), (34) and (36). The intraday increments are computed over the frequencies of the x-axis. The sample period covers 01/2009–03/2011 for Alcoa stock. The reported average is over the business days of our sample.

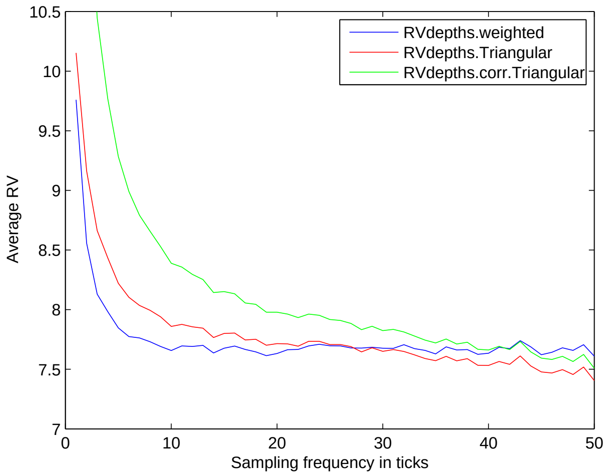

Figure 5: Signature plot for the realized measures based on the depths $RV^{depths.weighted}$, $RV^{depths.Triangular}$ and $RV^{depths.corr.Triangular}$. We use the expression given in (38), (42) and (45). The intraday increments are computed over the frequencies of the x-axis. The sample period covers 01/2009–03/2011 for Alcoa stock. The reported average is over the business days of our sample.

Alcoa	1 day		5 days	
R^2	In	Out	In	Out
RV^{all}	0.4767	0.4746	0.5105	0.4965
RV^{TS}	0.4800	0.4873	0.5133	0.5006
RV^{kernel}	0.4870	0.4934	0.5123	0.4997
RV^{pre}	0.4927	0.4959	0.5025	0.4910
RV^{inf}	0.4159	0.4101	0.5275	0.5108
RV^{sup}	0.5121	0.5095	0.5432	0.5313
RV^{bid}	0.4851	0.4827	0.5056	0.4927
RV^{ask}	0.4754	0.4728	0.5101	0.4954
$RV^{Uniform}$	0.4804	0.4781	0.5088	0.4950
$RV^{Triangular}$	0.4787	0.4765	0.5097	0.4957
$RV^{corr.Uniform}$	0.4790	0.4768	0.5096	0.4957
$RV^{corr.Triangular}$	0.5019	0.4997	0.5238	0.5116
$RV^{corr.Normal}$	0.4778	0.4756	0.5101	0.4961
$RV^{depths.weighted}$	0.4901	0.4883	0.5179	0.5045
$RV^{depths.Triangular}$	0.4862	0.4839	0.5120	0.4983
$RV^{depths.corr.Triangular}$	0.5072	0.5049	0.5275	0.5151

Table 1: In-sample and out-of-sample forecasting R^2. The table provides the R^2 of the Mincer-Zarnowitz regression (12) related to forecasts of IV_{t+1} for 1 and 5 days ahead. We report in-sample and out-of-sample results with a 100-days rolling window. The forecasting model is an AR(3) with RV^{low} as the dependent variable. The sample period covers 01/2009–03/2011 for Alcoa stock.

Profits	H=1	H=5
RV^{all}	-0.0635	-5.8989
RV^{TS}	-0.2562	-2.3097
RV^{kernel}	-0.2520	-3.4331
RV^{pre}	-0.5743	-1.7204
RV^{inf}	-1.5702	-10.6035
RV^{sup}	-0.4364	5.7356
RV^{bid}	-0.2634	1.3735
RV^{ask}	-0.3421	1.3403
$RV^{Uniform}$	-0.0533	0.3004
$RV^{Triangular}$	-0.0816	-2.8749
$RV^{corr.Uniform}$	0.0418	-2.0785
$RV^{corr.Triangular}$	-0.1139	5.9221
$RV^{corr.Normal}$	-0.0653	-4.5714
$RV^{depths.weighted}$	0.2051	3.0643
$RV^{depths.Triangular}$	-0.0609	2.6296
$RV^{depths.corr.Triangular}$	-0.2759	7.0041

Table 2: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre}, RV^{inf}, RV^{sup}, RV^{bid}, RV^{ask}, $RV^{Uniform}$, $RV^{Triangular}$, $RV^{corr.Uniform}$, $RV^{corr.Triangular}$, $RV^{corr.Normal}$, $RV^{depths.weighted}$, $RV^{depths.Triangular}$ and $RV^{depths.corr.Triangular}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0831	0.5507
RV^{TS}	-0.0186	0.6674
RV^{kernel}	-0.0258	0.1889
RV^{pre}	-0.1331	0.6702
RV^{inf}	-0.4006	-2.7671

Table 3: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and RV^{inf} – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	-0.0001	-0.5381
RV^{TS}	-0.0244	-0.3058
RV^{kernel}	-0.0455	-0.8169
RV^{pre}	-0.1576	-0.3348
RV^{sup}	-0.1863	1.4011

Table 4: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and RV^{sup} – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0368	-0.5341
RV^{TS}	-0.0513	-0.1807
RV^{kernel}	-0.0673	-0.7073
RV^{pre}	-0.1748	-0.1111
RV^{bid}	-0.0277	1.1772

Table 5: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and RV^{bid} – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0329	-0.5269
RV^{TS}	-0.0470	-0.0926
RV^{kernel}	-0.0624	-0.6736
RV^{pre}	-0.1614	-0.1219
RV^{ask}	-0.0396	1.0494

Table 6: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and RV^{ask} – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0104	-0.6048
RV^{TS}	-0.0201	-0.0873
RV^{kernel}	-0.0425	-0.6494
RV^{pre}	-0.1598	-0.0652
$RV^{Uniform}$	-0.0513	1.0695

Table 7: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{Uniform}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0172	-0.6111
RV^{TS}	-0.0555	0.0218
RV^{kernel}	-0.0425	-0.5595
RV^{pre}	-0.1660	0.0485
$RV^{Triangular}$	-0.0068	0.7792

Table 8: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{Triangular}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0196	-0.5738
RV^{TS}	-0.0673	0.0488
RV^{kernel}	-0.0539	-0.4825
RV^{pre}	-0.1764	0.0660
$RV^{corr.Normal}$	0.0291	0.6283

Table 9: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{corr.Normal}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0196	-0.5795
RV^{TS}	-0.0555	0.0252
RV^{kernel}	-0.0479	-0.5576
RV^{pre}	-0.1673	0.0467
$RV^{corr.Uniform}$	-0.0032	0.7433

Table 10: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{corr.Uniform}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0370	-0.7064
RV^{TS}	-0.0422	-0.3570
RV^{kernel}	-0.0763	-0.8816
RV^{pre}	-0.1682	-0.2951
$RV^{corr.Triangular}$	-0.0577	1.8229

Table 11: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{corr.Triangular}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	-0.0093	-0.7469
RV^{TS}	-0.0487	-0.2222
RV^{kernel}	-0.0928	-0.7560
RV^{pre}	-0.1704	-0.1632
$RV^{depths.weighted}$	0.0590	1.5445

Table 12: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{depths.weighted}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0499	-0.6673
RV^{TS}	-0.0353	-0.1466
RV^{kernel}	-0.0623	-0.7222
RV^{pre}	-0.1675	-0.1466
$RV^{depths.Triangular}$	-0.0520	1.3344

Table 13: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{depths.Triangular}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

Profits	H=1	H=5
RV^{all}	0.0552	-0.6968
RV^{TS}	-0.0560	-0.3869
RV^{kernel}	-0.0736	-0.8670
RV^{pre}	-0.1615	-0.3082
$RV^{depths.corr.Triangular}$	-0.0870	1.8126

Table 14: Alcoa profits from option trading. Using the alternative forecasts – RV^{all}, RV^{TS}, RV^{kernel}, RV^{pre} and $RV^{depths.corr.Triangular}$ – agents price options on Alcoa stock before trading with each other at average prices. H refers to the horizon of the agents' forecasts in days. The average profits from option trading are reported in cents. See section 2.3 for further details. The sample period covers 01/2009–03/2011.

References

Aït-Sahalia, Y. and L. Mancini (2008), "Out of Sample Forecasts of Quadratic Variation," *Journal of Econometrics*, 147, 17-33.

Andersen, T., T. Bollerslev, F.X. Diebold and P. Labys (1999), "(Understanding, Optimizing, Using and Forecasting) Realized Volatility and Correlation," *Working paper*. Published in revised form as "Great Realizations," *Risk*, March 2000, 105-108.

Andersen, T., T. Bollerslev, F.X. Diebold and P. Labys (2003), "Modeling and Forecasting Realized Volatility," *Econometrica*, 71 2, 579-625.

Bandi, F., J. Russell and C. Yang (2008), "Realized volatility forecasting and option pricing," *Journal of Econometrics*, 147, 34-46.

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

Barndorff-Nielsen, O.E. and N. Shephard (2002), "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 64, 2, 253-280.

Engle, R.F. and C.H. Hong (1990), "Valuation of variance forecasts with simulated option markets," *NBER Working Paper*, No: 3350.

Gatheral, J. and R.C.A. Oomen (2010), "Zero-intelligence realized variance estimation," *Finance and Stochastics*, 12 2, 249-283.

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

Hasbrouck, J. (1999), "The dynamics of discrete bid and ask quotes," *Journal of Finance*, 54, 2109-2142.

Jacod, J. (1994), "Limit of random measures associated with the increments of a Brownian semimartingale," *Tech. Rep. Université Paris VI.*

Jacod, J., Y. Li, P. Mykland, M. Podolskij and M. Vetter (2009), "Microstructure noise in the continuous case: The pre-averaging approach," *Stochastic Processes and their Applications*, 119, 2249-2276.

Kavajecz, K.A. (1999), "A Specialist's Quoted Depth and the Limit Order Book," *Journal of Finance*, 54, 2, 747-771.

Manski, C. F. (2003), "Partial Identification of Probability Distributions," New York: Springer-Verlag.

Mincer, J. and V. Zarnowitz (1969), "The Evaluation of Economic Forecasts," J. Mincer (ed.), *Economic Forecasts and Expectations*, New York: National Bureau of Economic Research.

Tamer, E. (2010), "Partial Identification in Econometrics," *Annual Review of Economics*, 167-195.

Zhang, L. (2006), "Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach," *Journal of the American Statistical Association*, 100, 1394-1411.

Zhang, L., P.A. Mykland and Y. Aït-Sahalia (2005), "A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394-1411.