Predictive Inference for
Integrated Volatility

Valentina Corradi
University of Warwick

Walter Distaso
Imperial College, London

Norman R. Swanson
Rutgers University

January 2007
Motivation

- From a risk management perspective, it is important to construct predictive conditional densities. Superiority of this approach than just focusing on conditional second moments.

- Financial volatility as a class of assets. Derivatives written on financial volatility are now traded on several over-the-counter markets. Counter cyclicality of stock market volatility has prompted the adoption of volatility exposure in portfolios.

- For the purpose of risk management of composite portfolios and for hedging positions taken on volatility derivatives, it is important to construct the conditional density forecast of volatility. Challenging, since volatility is a latent variable.
• Realized measure approach.

• Using a realized measure for predicting integrated volatility has also been adopted by Andersen, Bollerslev, Diebold and Labys (2003), Andersen, Bollerslev and Meddahi (2004, 2005).

These papers deal with pointwise prediction of integrated volatility, using ARMA models based on the log of realized volatility. Andersen, Bollerslev and Meddahi (2004, 2005) also investigate the important issue of evaluating the loss of efficiency associated with the use of a realized volatility measure, relative to the optimal (unfeasible) forecast (based on the entire path of volatility).
Results of the paper

1. Based on realized estimators of integrated volatility, we construct predictive densities and confidence intervals of daily volatility, conditional on observed market information.

2. The proposed estimators are consistent and asymptotically normally distributed, under mild assumptions on the underlying diffusion process.

3. No parametric assumption on either the functional form of the estimated densities, or the specification of the diffusion process driving the asset price. Nevertheless, the diffusive part of the log-price process is Brownian; in this sense, semiparametric approach.
Outline of the talk

1. Model
2. Kernel conditional densities
3. Limit theory
4. Empirical Application
Part 1

The model
Class of Brownian semimartingale processes plus jumps.

\[ Y_t = \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s + \sum_{i=1}^{N_t} c_i. \]

From a risk management perspective, it is of interest to separate the discontinuous (due to jumps) part of \( Y \), denoted by \( Y^d \), from the continuous Brownian component, \( Y^c \). Then, it is well known that

\[ \langle Y \rangle_t = \langle Y^c \rangle_t + \langle Y^d \rangle_t \]

and

\[ \langle Y^c \rangle_t = \int_0^t \sigma^2_s \, ds \quad \text{and} \quad \langle Y^d \rangle_t = \sum_{i=1}^{N_t} c_i^2. \]

A special case is obtained when \( N \equiv 0 \) (Brownian semimartingales).
We introduce market frictions in the following way. We assume that transaction data are contaminated by measurement errors, so that the observed process is given by

\[ X = Y + \epsilon. \]

We assume that we have a total of \( MT \) equispaced observations from the process \( X \), consisting of \( M \) intradaily observations for \( T \) days. That is

\[ X_{t+j/M} = Y_{t+j/M} + \epsilon_{t+j/M}, \quad t = 1, \ldots, T \quad \text{and} \quad j = 1, \ldots, M, \]

where \( \epsilon_{t+j/M} \) is geometrically mixing, so that, for each \( s \),

\[ \text{cov}(\epsilon_{t+j/M}, \epsilon_{t+(j+|s|)/M}) = \mathbb{E}(\epsilon_{t+j/M}\epsilon_{t+(j+|s|)/M}) \approx \rho^s, \quad |\rho| < 1. \]

Finally, we assume independence between \( Y \) and \( \epsilon \).
Our object of interest, daily integrated volatility, is defined as

\[ IV_t = \int_{t-1}^{t} \sigma_s^2 ds \]

at day \( t \). Since \( IV_t \) is not observable, different realized measures are used. The realized measure, \( RM_{t,M} \), is a noisy measure of integrated volatility:

\[ RM_{t,M} = IV_t + N_{t,M}. \]

In the sequel, we

1. derive kernel estimators of conditional densities and conditional confidence intervals based on a generic realized volatility measure.

2. Provide sufficient conditions on the measurement error \( N_{t,M} \), ensuring that the distributions of kernel density and confidence interval estimators
based on realized measures and the corresponding ones based on the “true” (but latent) daily volatility process are asymptotically equivalent.

3. Finally, we adapt results on $N_{t,M}$ to four realized measures: namely,

(a) realized volatility

(b) normalized bipower variation

(c) two scale realized volatility measure, $\hat{RV}_{t,l,M}$ (Zhang, Mykland and Aït-Sahalia, 2005).

(d) multi scale realized volatility, $\tilde{RV}_{t,e,M}$, proposed by Zhang (2004), Aït-Sahalia, Mykland and Zhang (2005b) (see also Barndorff-Nielsen, Hansen, Lunde and Shephard, 2005).
Realized Volatility Measures - 1

(i) Realized Volatility:

\[ RV_{t,M} = \sum_{j=1}^{M-1} \left( X_{t+(j+1)/M} - X_{t+j/M} \right)^2. \]

(ii) Normalized Bipower Variation:

\[ (\mu_1)^{-2} BV_{t,M} = (\mu_1)^{-2} \sum_{j=2}^{M-1} \left| X_{t+(j+1)/M} - X_{t+j/M} \right| \left| X_{t+j/M} - X_{t+(j-1)/M} \right|, \]

where \( \mu_1 = \text{E} |Z| \) and \( Z \) is a standard normal random variable.
Microstructure Noise
Constructing Estimators Robust to
Realized Volatility Measures - 2:

Predictive Inference for Integrated Volatility

January 2007
Part 2

Conditional Density Estimators
We construct a nonparametric estimator of the density of integrated volatility at time $T + 1$, conditional on a given realized volatility measure actually observed at times $T, \ldots, T - d$. Let

$$RM_{t,M}^{(d)} = (RM_{t,M}, \ldots, RM_{t-(d-1),M}),$$

and

$$IV_{t}^{(d)} = (IV_{t}, \ldots, IV_{t-(d-1)}).$$

The conditional density is defined as the ratio of the joint and marginal densities, and the latter is assumed to be bounded away from zero.
The feasible and unfeasible conditional kernel density estimators are

\[
\hat{f}_{RM_{T+1},M|RMT_{T,M}}(x|RMT_{T,M}) = \frac{1}{T} \sum_{t=1}^{T} K \left( \frac{RM_{t,M}^{(d)} - RMT_{T,M}^{(d)}}{\xi_1} \right) \frac{1}{\xi_1^d} K \left( \frac{RM_{t+1,M}^{(d)} - x}{\xi_2} \right) \frac{1}{\xi_2}
\]

and

\[
\hat{f}_{IV_{T+1}|IV_{T}}^{(d)}(x|RMT_{T,M}) = \frac{1}{T} \sum_{t=1}^{T} K \left( \frac{IV_{t}^{(d)} - RMT_{T,M}^{(d)}}{\xi_1} \right) \frac{1}{\xi_1^d} K \left( \frac{IV_{t+1} - x}{\xi_2} \right) \frac{1}{\xi_2}
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} K \left( \frac{IV_{t}^{(d)} - RMT_{T,M}^{(d)}}{\xi_1} \right) \frac{1}{\xi_1^d}.
\]
We provide conditions on the relative rate of growth of \( M, T \) (as \( M, T \to \infty \)) and \( \xi_1, \xi_2 \) (as \( \xi_1, \xi_2 \to 0 \)), under which the limiting distribution of

\[
\sqrt{T^{d} \xi_1 \xi_2} \left( \hat{f}_{RM_{T+1,M}|RM_T,M}^{(d)}(x|R_{MT,M}) - \hat{f}_{IV_{T+1}|IV_T}^{(d)}(x|R_{MT,M}) \right)
\]

is the same as that of

\[
\sqrt{T^{d} \xi_1 \xi_2} \left( \hat{f}_{IV_{T+1}|IV_T}^{(d)}(x|R_{MT,M}) - f_{IV_{T+1}|IV_T}^{(d)}(x|R_{MT,M}) \right).
\]

• Connection with literature on kernel density estimators with measurement error.
Part 3

Asymptotic Theory
Assumptions

Assumption A1, A2: $\mu_s$ and $\sigma_s$ satisfy regularity conditions such that

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

is a strictly stationary $\alpha-$mixing process.

Assumption A3, A4: Kernel functions are symmetric, nonnegative, continuous function with bounded support, satisfying some regularity conditions.

Assumption A5: Some regularity conditions (again!!) on $f_{IV_T}^{(d)}(\cdot)$ and $f_{IV_{T+1}|IV_T}^{(d)}(x|\cdot)$. 
Theorem 1. Let A1-A5 hold. Then, pointwise in $x$,

(i) if $\xi_1, \xi_2 \to 0$, $T\xi_1\xi_2 \to \infty$,

$$\hat{f}_{RM_{t+1,M}|RM_{t,M}^{(d)}}(x|R_{T,M}^{(d)}) - \hat{f}_{IV_{t+1,IV_{t,M}^{(d)}}}(x|R_{T,M}^{(d)}) = O_P \left( b_M^{-1/2} \right).$$

(ii) If $\xi_1, \xi_2 \to 0$, $T\xi_1\xi_2 \to \infty$, and $Tb_M^{-1}\xi_1\xi_2 \to 0$,

$$\sqrt{T \xi_1\xi_2} \left( \hat{f}_{RM_{t+1,M}|RM_{t,M}^{(d)}}(x|R_{T,M}^{(d)}) - \hat{f}_{IV_{t+1,IV_{t,M}^{(d)}}}(x|R_{T,M}^{(d)}) \right) = \sqrt{T \xi_1\xi_2} \left( \hat{f}_{IV_{t+1,IV_{t,M}^{(d)}}}(x|R_{T,M}^{(d)}) - \hat{f}_{IV_{t+1,IV_{t,M}^{(d)}}}(x|R_{T,M}^{(d)}) \right) + o_P(1).$$

Part (i) reveals that the rate at which the measurement error approaches zero depends only on the number of intradaily observations.
Theorem 2. Let A1-A5 hold. If $\xi_1, \xi_2 \to 0$, $T\xi_1^d\xi_2 \to \infty$, $T\xi_1^{4+d}\xi_2 \to 0$, $T\xi_1^d\xi_2^5 \to 0$ and $Tb_0^{-1}\xi_1^d\xi_2 \to 0$, then

$$
\sqrt{T\xi_1^d\xi_2} \left( \hat{f}_{RM_{t+1,M}|RM_t^{(d)}}(x|RM_T^{(d)},M) - f_{IV_{t+1}|IV_t^{(d)}}(x|RM_T^{(d)}) \right)
\xrightarrow{d} N \left( 0, \frac{f_{IV_{t+1}|IV_t^{(d)}}(x|RM_T^{(d)})}{f_{IV_t^{(d)}}(RM_T^{(d)})} \int K^2(u)du \int K^2(v)dv \right).
$$

As the underlying conditional density is at least twice differentiable, the bias term goes to zero at a rate not slower than $\min(\xi_2^2, \xi_2^2)$. Thus, if $T\xi_1^{4+d}\xi_2 \to 0$ and $T\xi_1^d\xi_2^5 \to 0$ the bias term associated with the estimator will be asymptotically negligible.
Specific Realized Volatility Measures

Lemma 1. Let A1 and A2 hold.

If $RM_{t,M} = RV_{t,M}$, then $b_M = M$.

If $RM_{t,M} = BV_{t,M}$, then $b_M = M$.

If $RM_{t,M} = \hat{RV}_{t,l,M}$, then $b_M = M^{1/3}$.

If $RM_{t,M} = \tilde{RV}_{t,e,M}$, then $b_M = M^{1/2}$. 
Remarks

Remark 1

In practice, one can choose between selecting a relatively small $M$ and a non microstructure robust realized measure, or selecting a large $M$ and a microstructure robust realized measure.

Remark 2

In empirical work, volatility is often modelled and predicted with ARMA models constructed using logs of realized volatility. A Taylor expansion of $\log(RM_{t,M})$ around $IV_t$ gives

$$
\log(RM_{t,M}) = \log(IV_t) + \frac{1}{IV_t} N_{t,M} - \frac{1}{2 IV_t^2} N_{t,M}^2 + \frac{1}{3 IV_t^3} N_{t,M}^3 + \ldots,
$$

and our theory can be applied straightforwardly.
Part 4

Empirical Illustration
Forecasting the conditional density of daily volatility of Intel

Data description

Data are retrieved from the Trade and Quotation (TAQ) database at the New York Stock Exchange (NYSE). We base our analysis on two different sample sizes. The first one extends from January 2 to May 27, 1998; the second from January 2 to May 22, 2002. Both sample sizes consist of a total of 100 trading days.

This to analyze the effect of the decimalization of the tick size. The tick size was reduced from a sixteenth of a dollar to one cent on January 29, 2001.

We predict the conditional distribution of daily integrated volatility for Thursday May 28th, 1998 and Thursday May 23rd, 2002.
From the original data set, which includes prices recorded for every trade, we extracted 10 seconds and 5 minutes interval data. The choice of the two frequencies has been done in order to evaluate the effect of microstructure noise on the estimated densities.

From the calculated series we we have obtained 10 seconds and 5 minutes intraday returns as the difference between successive log prices. A full trading day consists of 2340 (resp. 78) intraday returns calculated over an interval of ten seconds (resp. five minutes).
Boundary Corrected Kernels

Conventional kernel functions do not produce consistent estimates when the evaluation points are close to the boundary.

To resolve this problem, we have used the boundary corrected kernel function of Müller (1991), using a locally variable bandwidth. Consider a density estimator based on the standard quartic kernel

$$\hat{f}(x) = \frac{1}{n\xi} \sum_{i=1}^{n} K\left( \frac{x - X_i}{\xi} \right),$$

where

$$K(u) = \frac{15}{16} \left( 1 - u^2 \right)^2 1_{\{|u| \leq 1\}}.$$
Denote $q = \min(x/\xi, 1)$. The boundary modified kernel estimator will have the form

$$\hat{f}_q(x) = \frac{1}{n\xi_q} \sum_{i=1}^{n} K_q\left(\frac{x - X_i}{\xi_q}\right),$$

where

$$K_q(u) = \frac{30 (1 + u^2) (q - u)^2}{(1 + q)^5} \left(1 + 7 \left(\frac{1 - q}{1 + q}\right)^2 + 14 \frac{(1 - q) u}{(1 + q)^2}\right) \mathbf{1}_{\{-1 \leq u \leq q\}}$$

and

$$\xi_q = b(q) \xi = (2 - q) \xi.$$

Notice that $K_1(u) = K(u)$. Hence, the resulting limiting distributions are not changed.
Main Results

Results are for the case of $d = 1$, using the boundary modified quartic kernel function and 1000 evaluation points. They are reported in Figures 1 to 5.

Instead Figures 6 and 7 report the estimated conditional densities and 10% confidence intervals of the log of integrated volatility. Results have been obtained using a standard quartic kernel function for 1000 evaluation points.
Figure 1: Predictive conditional density of $IV^T+1$ for Intel on May 28th, 1998, based on realized volatility.

$M = 78, T = 100$.

Figure 2: Predictive conditional density of $IV^T+1$ for Intel on May 28th, 1998, based on bipower variation.

$M = 78, T = 100$. 

Realized Volatility vs Bipower Variation
Figure 3: Predictive conditional densities of $\text{IV}^{T+1}$ for Intel on May 28th, 1998, based on realized volatility.

$\text{IV} = I_{100} \times \text{IV} = I_{78} \times \text{IV} = I_{100} \times \text{IV}$ for $T = 100$, $M = 7,8340$ (blue line), $M = 2340$ (red line).
Figure 4: Predictive conditional densities of $\text{IV}_{T+1}$ for Intel on May 23rd, 2002, based on bipower variation. $T = 100$, $M = 378$ (blue line), $M = 2340$ (red line).
Figure 5: Predictive conditional densities of $I^{T+1}$ for Intel on May 23rd, 2002, based on $\hat{RV}_{t,l,M}$ for $T=100$ and $M=78$ (blue line), $M=2340$ (red line).
Figure 7: Predictive conditional density of the log of IV$_{T+1}$ for Intel on May 23rd, 2002, based on realized volatility. $M = 78, T = 100$.

Figure 6: Predictive conditional density of the log of IV$_{T+1}$ for Intel on May 23rd, 2002, based on realized volatility. $M = 2340, T = 100$. 

Log of Integrated Volatility