Estimating and Forecasting Volatility using High Frequency Data

Asger Lunde

CREATES and University of Aarhus, Denmark

July 29, 2010
What is High Frequency Data

- Different resolutions

\[ r_{t-\Delta_o,\Delta_c} \]

\[ r_{t,i}, \; i = 1, \ldots, N_t \]

\[ r_t \]

overnight return

intra-day returns

Interday return
What is High Frequency Data is Intra-day Data

- Second by second transaction prices, bid/ask quotes, depth, trading volume, etc.
Easy access to high frequency (HF) financial data has spurred much activity in financial economics.

The introduction of empirical estimators of the quadratic variation to measure the ex post variation of asset prices is a prime example.

There is now a range of volatility estimators that are computed from high frequency data, it has become common to refer to such as realized measures.

For the problem of forecasting volatility, high frequency data have been valuable in a number of ways.

The main reason that high-frequency data can greatly improve the forecast accuracy is simply that volatility is highly persistent, so that a more accurate measure of current volatility, which high frequency data provide, is valuable for forecasting future volatility.
Six ways that HF data have improved volatility forecasting

1. HF data improve our understanding of the dynamic properties of volatility which is key for forecasting.

2. Realized measures are valuable predictors of future volatility in reduced form models.

3. Realized measures have enabled the development of new volatility models that provide more accurate forecasts.

4. HF data have improved the evaluation of volatility forecasts in important ways.

5. Realized measures can facilitate and improve the estimation of complex volatility models (e.g. continuous time models). Less estimation error improve predictions based on such models.

6. HF data have improved our understanding of the driving forces of volatility. E.g., HF data have enabled a detailed analysis of news announcements and their effect on the financial markets.
Intro
Outline of Talk

- HF Data and Realized Measures of Volatility
  - The underlying continuous time model
  - Quadratic Variation and Realized Variance
  - Connecting the Parts
  - Realized Measures of Volatility

- Inference about Latent Volatility from Noisy Proxies
  - Instrumental Variable Estimators
  - Empirical Analysis of Realized Measures
  - The ACF of A Latent Time Series

- Reduced Form Volatility Forecasts
  - Distributed Lag, ARFIMA and HAR Models

- Model based Volatility Forecasts
  - GARCH and GARCH-X Models
  - Completing the GARCH-X: Realized GARCH

- Financial Application: Realized Beta GARCH

- Conclusion
Talk is based on

Forecasting Volatility using High Frequency Data

by

Peter R. Hansen and Asger Lunde

Chapter 19 in OUP Handbook of Economic Forecasting

Forthcoming at Oxford University Press
HF Data and Realized Measures of Volatility
Notational Framework

- Log prices: $Y(t)$
- Daily returns: $y_t = Y(t) - Y(t - 1)$,
- Intraday returns: $y_{t,i} = Y(\tau_i) - Y(\tau_{i-1})$, for $i = 1, \ldots, N_t$.

Denote the conditional mean and variance of $y_t$ by

$$
\mu_t = E[y_t|\mathcal{F}_{t-1}], \quad \text{and} \quad \sigma_t^2 = \text{var}[y_t|\mathcal{F}_{t-1}].
$$

The richness of $\mathcal{F}_t$ is key for defining $\sigma_t^2$ and the forecasting problem.

In the classical approach to volatility forecasting $\mathcal{F}_t$ is typically generated by sparse daily information, such as opening or closing prices.

Recent improvements is where $\mathcal{F}_t$ also comprises HF information (e.g. intraday transaction prices, quotes, trading volume and depth).
HF Data and Realized Measures of Volatility
The underlying continuous time model

- Let us make the connection to the underlying continuous time model (Lto process), that is
  \[ dY(t) = \mu(t)dt + \sigma(t)dW(t), \]

- This is also known as a stochastic volatility (SV) model.

- It has some very convenient properties, in particular,
  \[ y_t|\mu_t, IV_t \sim N(\mu_t, IV_t), \]

where
  \[ \mu_t \equiv \int_{t-1}^t \mu(s)ds \quad \text{and} \quad IV_t \equiv \int_{t-1}^t \sigma^2(s)ds. \]

- So in this framework the integrated variance, \( IV_t \), is the population measure of actual return variance.
\[ Y(t) \] could look like this over a unit interval:

The quadratic variation of a stochastic process over \([t - 1, t]\) is given by

\[
QV_t \equiv \lim_{N \to \infty} \sum_{j=1}^{N} \{ Y(\tau_j) - Y(\tau_{j-1}) \}^2,
\]

where \( \max_j |\tau_j - \tau_{j-1}| \to 0 \) as \( N \to \infty \).
The empirical counterpart of $QV_t$ is called the realized variance which is simply the sum of the squared observed intraday returns,

$$RV_t \equiv \sum_{j=1}^{N_t} y_{t,j}^2.$$ 

$RV_t$ consistently estimates the corresponding $QV_t$ for all semimartingales, and for SV models it holds that $QV_t = IV_t$.

So for SV models the population measure, $IV_t$, is also consistently estimated by $RV_t$. 
We have the following relation

\[ \sigma_t^2 = \text{var} \left[ r_t \mid \mathcal{F}_{t-1} \right] \approx \mathbb{E} \left[ RV_t \mid \mathcal{F}_{t-1} \right] \approx \mathbb{E} \left[ IV_t \mid \mathcal{F}_{t-1} \right]. \]

so we can build models for volatility forecasting directly on the time series of realized variances.

The literature has evolved to a higher level of generality than this.

Due to various measurement issues, known as market microstructure noise, several alternative estimators of \( IV_t \) have been suggested.

We refer to such estimators as realized measures, \( RM_t \).

These measures are key for obtaining more precise forecasts and for performance evaluation among different forecasting methods.
Merton (1980); in a noiseless environment the variance of returns can be estimated much more precisely from realized returns than the expected return.

Barndorff-Nielsen & Shephard (2002) extend this insight to the SV model and show that $RV_t$ convergence stably in law, "$\rightarrow$",

\[
n^{1/2}(RV_n - IV) \overset{Ls}{\rightarrow} MN(0, 2IQ),
\]

implying a joint convergence, so that

\[
\frac{n^{1/2}(RV_n - IV)}{\sqrt{2IQ}} \overset{L}{\rightarrow} N(0, 1).
\]

Merton (1980) also wrote: “in practice, the choice of an even-shorter observation interval introduces another type of error which will ‘swamp’ the benefit [...] long before the continuous limit is reached”.
Market microstructure effects cause the observed price to deviate from the efficient price (that has the semi-martingale property).

So if $Y(t)$ is the latent true price process. Then for estimating $IV_t$ we will observe:

$$X(\tau_j) = Y(\tau_j) + U(\tau_j).$$

We think of $U$ as noise that can be due to, for example, liquidity effects, bid/ask bounce and misrecordings.


See Hansen & Lunde (2006) for a comprehensive analysis of the statistical properties of $U$. 
Several methods have been developed for estimating the IV and QV in the presence of noise.


For instance the realized kernel estimator of BNHLS (2008) takes the form

$$ RK_t = \sum_{h=-H}^{H} k \left( \frac{h}{H} \right) \gamma_h, \quad \gamma_h = \sum_{j=|h|+1}^{n} x_j x_{j-|h|}, \quad x_j = X(\tau_j) - X(\tau_{j-1}) $$

where \( k(x) \) is a kernel weight function, such as the Parzen kernel.

BNHLS (2009); details about implementation.

BNHLS (2010); extension to multivariate price processes.
Observed volatility time series are naturally viewed as proxies for the underlying population quantities.

Consider a *time series of daily realized variances*.

Each element of this time series can be viewed as a noisy estimate of the latent volatility.

Much progress has recently been made in estimating financial volatility from high-frequency data using realized measures, such as the realized variance.

Despite this progress, it is *important to discriminate between the realized measure of volatility* and the underlying *population quantity*.
Even for the most accurate estimators of daily volatility the standard error for a single estimate is rarely less than 10% of the point estimate, see Barndorff-Nielsen, Hansen, Lunde & Shephard (2008).
Inference about Latent Volatility from Noisy Proxies

- The analysis draws on the paper:

- Consider for simplicity an univariate ARMA\(_p,q\) specification:

\[
\varphi(L)(y_t - \delta) = \theta(L)\varepsilon_t, \quad x_t = y_t + \zeta + \eta_t,
\]

where

\[
\exp(y_t) = \text{is the latent volatility}
\]

\[
\exp(x_t) = RA_t \text{ is the observed volatility}
\]

\[
\eta_t \text{ is the measurement error}
\]

- Then

\[
\varphi(L) (x_t - \delta - \zeta) = \theta(L)\varepsilon_t + \varphi(L)\eta_t.
\]

- So \(x_t\) is an ARMA process with the exact same autoregressive polynomial, (noted by Barndorff-Nielsen & Shephard (2002) and Meddahi (2003) for the ARMA\(_{1,1}\) case).
A key parameter is the *persistence parameter* ($z_1^*, \ldots, z_p^*$ are the roots of $\varphi(z)$):

$$\pi = \max_{i=1, \ldots, p} \frac{1}{|z_i^*|},$$

For $\pi = 1$ when $\varphi(z)$ has a unit root and for persistent processes we have $\pi \approx \varphi = \varphi_1 + \cdots + \varphi_p$.

We have a classical errors-in-variable problem. So in general the estimates, $\hat{\varphi}_i$, we get from estimating

$$x_t = \sum_{i=1}^{p} \varphi_i x_{t-i} + \varepsilon_t$$

will be biased and inconsistent for autoregressive parameters of the latent volatility process.
Inference about Latent Volatility from Noisy Proxies
Instrumental Variable Estimators

- IV estimators of the persistence parameter $\pi$, that have the form

$$\hat{\pi}_{IVz} = \frac{\sum_{t=1}^{n} Z_t x_{t+1}}{\sum_{t=1}^{n} Z_t x_t},$$

where $z_t$ is the instrument. Fx lagged values of $x_t$.

- Includes the 2SLS estimator, which is based on,

$$\tilde{Z}_t = (x_{t-J_1} - \bar{x}_{J_1}, \ldots, x_{t-J_2} - \bar{x}_{J_2})' \quad \text{with} \quad 0 \leq J_1 \leq J_2.$$

with

$$z_t = \tilde{Z}_t' \hat{\alpha}_{TSL}, \quad \text{where} \quad \hat{\alpha}_{TSL} = \left( \sum_{t=1}^{n} \tilde{Z}_t \tilde{Z}'_t \right)^{-1} \sum_{t=1}^{n} \tilde{Z}_t x_t.$$
Inference about Latent Volatility from Noisy Proxies
The Empirical and Latent Acfs: Volatility for BA

- Stocks in the Dow (DJIA). Realized variance and the realized kernel.
  sample period: Jan 3, 2002 to Jul 31, 2009 gives 1,907 obs. per stock.

- Conventional ACFs are quite different ... the two ACF* s are in agreement.
Inference about Latent Volatility from Noisy Proxies
The Empirical and Latent Acfs: Volatility for MRK

Again the two $ACF^*$-estimates are in agreement.
Empirical Analysis
Squared or Absolute Returns

Squared returns

Absolute returns

log(absolute returns)

ACF-LS
ACF-2SLS
Empirical Analysis
Squared or Absolute Returns

The new estimation of the autocorrelation of the underlying process, $\text{ACF}_x^*$, reveals that the choice between long memory and unit root is less clear-cut than is suggested by the conventional autocorrelation function for the observed process, $\text{ACF}_x$. 
Methods are quite similar in terms of the forecasts they produce.

- can be viewed as flexible extensions of the simple exponential smoothing forecast.

Not surprising because volatility is known to be highly persistent (close to unit root) and a realized measure is simply a noisy measurement of the underlying population volatility.

Since, the MSE optimal forecast within the local level model (random walk with noise) is given by exponential smoothing, we should not be surprised to see the reduced form forecasts yield point forecasts that closely resembles that of exponential smoothing.
Generalized in Ghysels, Santa-Clara & Valkanov (2006) that consider volatility (MIDAS) regressions such as

\[ RV_{t,h} = \mu + \phi \sum_{k=0}^{K} b(k, \theta) \tilde{X}_{t-k} + \varepsilon_t, \]

where \( RV_{t,h} = RV_t + \cdots + RV_{t+h} \).

Idea: \( \tilde{X}_{t-k} \), can be sampled at any frequency that seems relevant for predicting \( IV_{t,h} \).

\( \tilde{X}_{t-k} \): daily RV, 5 minute or daily squared/absolute returns, the daily range and the daily sum of intraday absolute returns.
The idea of fitting ARFIMA models to realized measures was put forth in 1999 and later published as Andersen, Bollerslev, Diebold & Labys (2003).

They fit a trivariate long-memory Gaussian VAR:

\[ \Phi(L)(1 - L)^d(y_t - \mu) = \epsilon_t \]

The dependent variable, \( y_t \), is a vector with \( \log(RV_t)/2 \) components for DM/$, ¥/$ and ¥/DM exchange rates. They found \( \hat{d} = 0.401 \).

A general finding of empirical studies with ARFIMA models in this context, is that this framework produces volatility forecasts that dominate those of conventional GARCH models that are based on daily returns.
Corsi (2009) proposes the a (HAR-RV) model of different volatility components designed to mimic the actions of different types of market participants.

It is a predictive model for the daily integrated volatility

$$\sqrt{RV_{t+1}} = c + \beta^{(d)} \sqrt{RV_t} + \beta^{(w)} \sqrt{RV_{t-5,5}} + \beta^{(m)} \sqrt{RV_{t-22,22}} + \varepsilon_{t+1}.$$ 

Predicts future volatility using a daily, a weekly and a monthly component.

The HAR model is found to be very successful in practice, and is comparable to the ARFIMA models.
Model based Volatility Forecasts
GARCH and GARCH-X Models

- Recall
  \[ \mu_t = E [r_t | \mathcal{F}_{t-1}] , \quad \text{and} \quad \sigma_t^2 = \text{var} [r_t | \mathcal{F}_{t-1}] . \]

- where \( \mathcal{F}_t \) comprises all variables that are available for prediction at times \( t \).

- Conventional GARCH models provide a specification for a conditional variance that is defined with a simpler filtration, which is generated exclusively by past returns, \( \mathcal{F}_t^r = \sigma(r_t, r_{t-1}, \ldots) \).

- Realized measures are useful additions to GARCH models because
  \[ \text{var}(r_t | \mathcal{F}_{t-1}^r) \neq \text{var}(r_t | \mathcal{F}_{t-1}^{r,RM}) , \]

- where \( \mathcal{F}_t^{r,RM} = \sigma(RM_t, r_t, RM_{t-1}, r_{t-1}, \ldots) \).

- The difference tends to be more pronounced after a sudden change in the conditional variance.
Illustrating $\operatorname{var}(r_t | \mathcal{F}_{t-1})$ and $\operatorname{var}(r_t | \mathcal{F}_{t-1}, R_{t-1})$, 

\begin{align*}
\text{overnight return} & \\
\text{intra-day returns} & \\
\text{interday return} & \\
\text{Realized variance based on intra-day returns} & \\
\text{Time} & \\
\end{align*}
Model based Volatility Forecasts
GARCH and GARCH-X Models

- Distinguish between the true conditional variance,
  \[ \sigma_t^2 = \text{var}(r_t | \mathcal{F}_{t-1}) \]
  and the model-based equivalent
  \[ h_t^G = \text{var}(r_t | \mathcal{F}_{t-1}^r), \quad \text{and} \quad h_t^{GX} = \text{var}(r_t | \mathcal{F}_{t-1}^{r, RM}), \]

- The simplest GARCH(1,1) model is given by
  \[ r_t = \sqrt{h_t^G} z_t, \]
  \[ h_t^G = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}^G, \]
  where \( z_t \sim \text{iid}(0, 1) \).

- So squared returns, \( r_{t-1}^2 \), defines the dynamic of the conditional variance

- \( \beta \approx 0.95 \) and \( \alpha \approx 0.05 \) in practice.
Model based Volatility Forecasts
GARCH and GARCH-X Model

- \( r_{t-1}^2 \) can be viewed as a noisy measure of volatility

Realized measures, such as the Realized Variance, provide more accurate measurements of volatility.

- Engle (2002), and many others

\[
h_t^{Gx} = \omega + \alpha r_{t-1}^2 + \beta h_{t-1}^{Gx} + \gamma R_{t-1}.
\]

- \( R_{t-1} \) is a realized measure of volatility (e.g. RV)

Typically
- \( \hat{\gamma} \approx 0.4 \)
- \( \hat{\alpha} \approx 0 \) (ARCH parameter becomes insignificant)

Huge improvement in the empirical fit.
Value of introducing realized measures of volatility into GARCH models?

- How well does $h_t^G$ or $h_t^{GX}$ approximate $\sigma_t^2$ in various circumstances?

- Simple example (Hansen, Huang & Shek (2009)). Suppose that volatility is $\sigma_t = 20\%$ for $t < T$ and then jumps to $\sigma_t = 40\%$ for $t \geq T$,
Model based Volatility Forecasts
GARCH is Slow

- Suppose that $r_t^2$ and $RM_t$ are both unbiased estimates of $\sigma_t^2$.
- Let $\omega = 0$, $\beta = 0.95$, and $\alpha = 0.05$ as typical for a GARCH(1,1)
  - The implication is that for $k \geq 0$
    \[
    E(h_{T+k}^G) = (1 - \beta^k)(40\%)^2 + \beta^k(20\%)^2
    \]
Model based Volatility Forecasts
GARCH-X is Fast

\[ \omega = \alpha = 0, \beta = 0.5, \text{ and } \gamma = 0.5, \]
are in line with typical GARCH-X estimates:

\[ h_t^G \text{ to get slightly above } 39\% \text{ (on average)} \]
\[ \text{It takes just four days for the GARCHX model to bring } h_t^{GX} \text{ up to date.} \]
The GARCH-X is an incomplete model in the sense that it makes no attempt to model the realized measure, $RM_t$.

For predicting volatility one-period ahead there is not necessary to specify a model for $RM_t$.

However, for predicting the volatility a longer horizons a specification for $RM_t$ is needed to “complete” the model.

Engle & Gallo (2006) and Shephard & Shephard (2010) among others use parallel GARCH structures where $RM_t$ is effectively being modeled analogously to the way $r_t^2$ is modeled in the GARCHX model.
The Realized GARCH model by Hansen, Huang & Shek (2009) introduces a measurement equation that ties $RM_t$ to $h_t$.

For instance, the measurement equation could take the form $(z_t = r_t / h_t)$

$$RM_t = \zeta + \phi h_t + \tau(z_t) + u_t,$$

where $u_t \sim iid(0, \sigma^2_u)$

$\tau(z_t)$: dependence between shocks to returns and shocks to volatility (leverage)

it should be ensured that $E[\tau(z_t)] = 0$ whenever $E[z_t] = 0$ and $Var[z_t] = 1$.

An implication of the measurement equation is that

$$E(RM_t|\mathcal{F}_{t-1}) = \zeta + \phi h_t,$$

since $\phi$ may be less than one, we can have $RM_t$ computed over a period that is shorter than the period that the return, $r_t$, spans.

Important in practice: For modelling of close-to-close returns, the availability of HF data is often limited to the open-to-close period.
Financial Application: Realized Beta GARCH
Notation and Modelling Strategy

Based on Hansen, Lunde and Voev (2010) ... work in progress

- $r_{0,t}$: denotes the market return.
- $x_{0,t}$: realized measure of market volatility.
- $r_{1,t}$: denotes the return on the individual asset.
- $x_{1,t}$: realized measure of the individual asset volatility.
- $\tilde{x}_{01,t}$: realized covariance measure
- $y_{1,t} = \frac{\tilde{x}_{01,t}}{\sqrt{x_{0,t}x_{1,t}}}$: realized correlation. From a psd $2 \times 2$ realized covariance matrix, ensuring $y_{1,t} \in (-1, 1)$.

The information set is thus given by

$$\mathcal{F}_t = \sigma(r_{0,t}, r_{1,t}, x_{0,t}, x_{1,t}, y_{1,t}, r_{0,t-1}, r_{1,t-1}, x_{0,t-1}, x_{1,t-1}, y_{1,t-1}, \ldots).$$

In our modelling we utilize the following decomposition of the joint conditional density,

$$f(r_{0,t}, x_{0,t}, r_{1,t}, x_{1,t}, y_{1,t}|\mathcal{F}_{t-1}) = f(r_{0,t}, x_{0,t}|\mathcal{F}_{t-1}) f(r_{1,t}, x_{1,t}, y_{1,t}|r_{0,t}, x_{0,t}, \mathcal{F}_{t-1}).$$
A Realized EGARCH model for market returns and realized measures of volatility \(f(r_{0,t}, x_{0,t} | \mathcal{F}_{t-1})\) is used:

\[
  r_{0,t} = \mu_0 + \sqrt{h_{0,t}}z_{0,t}, \\
  \log h_{0,t} = a_0 + b_0 \log h_{0,t-1} + c_0 \log x_{0,t-1} + \tau_0(z_{0,t-1}), \\
  \log x_{0,t} = \xi_0 + \varphi_0 \log h_{0,t} + \delta_0(z_{0,t}) + u_{0,t},
\]

where we model \(z_{0,t} \sim \text{iid} \mathcal{N}(0, 1)\), \(u_{0,t} \sim \text{iid} \mathcal{N}(0, \sigma_u^2)\).

Note that we do not follow the conventional GARCH notation, because we want to reserve the notation “\(\beta\)” for

\[
  \beta_t = \text{cov}(r_{1,t}, r_{0,t} | \mathcal{F}_{t-1}) / \text{var}(r_{0,t} | \mathcal{F}_{t-1}),
\]

We are particularly interested in the dynamic properties of \(\beta_{1,t}\).
Model \( f(r_{1,t}, x_{1,t}, y_{1,t} | r_{0,t}, x_{0,t}, \mathcal{F}_{t-1}) \) ... conditional on contemporary “market” variables.

Further decomposition

\[
f(r_{1,t}, x_{1,t}, y_{1,t} | r_{0,t}, x_{0,t}, \mathcal{F}_{t-1}) = f(r_{1,t} | r_{0,t}, x_{0,t}, \mathcal{F}_{t-1}) f(x_{1,t}, y_{1,t} | r_{1,t}, r_{0,t}, x_{0,t}, \mathcal{F}_{t-1}).
\]

The first two have an Realized EGARCH structure for the individual asset,

\[
\begin{align*}
r_{1,t} &= \mu_1 + \sqrt{h_{1,t}} z_{1,t}, \\
\log h_{1,t} &= a_1 + b_1 \log h_{1,t-1} + c_1 \log x_{1,t-1} + d_1 \log h_{0,t} + \tau_1(z_{1,t-1}).
\end{align*}
\]
Financial Application: Realized Beta GARCH
Individual Asset Returns, Volatility, and Covolatility

- A conditional covariance capture the dependence between market and individual returns
  \[ \rho_t = \text{cov}(z_{0,t}, z_{1,t} | \mathcal{F}_{t-1}), \]
  note that \( \rho_t \) is the conditional correlation between \( r_{0,t} \) and \( r_{1,t} \).

- Need to specify the dynamic properties of \( \rho_t \)
  \[ F(\rho_t) = a_{01} + b_{01} F(\rho_t) + c_{01} F(y_{1t}). \]
  Fisher transformation, \( F(\rho) = \frac{1}{2} \log \frac{1+\rho}{1-\rho} \), mapping \((-1, 1)\) into \( \mathbb{R} \).

- Finally, the model is completed using more two measurement equations:
  \[ \log x_{1,t} = \xi_1 + \varphi_1 \log h_{1,t} + \delta_1(z_{1,t}) + u_{1,t}, \]
  and
  \[ F(y_{1t}) = \xi_{01} + \varphi_{01} F(\rho_t) + v_{1t}. \]

- Estimation using quasi maximum likelihood, for details see Hansen, Lunde and Voev (2010).
The implied beta is now given by

$$\beta_t = \rho_t \sqrt{h_{1,t} h_{0,t}} = \rho_t \sqrt{\frac{h_{1,t}}{h_{0,t}}}.$$ 

Because volatility is usually rather persistent, we expect $\beta_t$ to “inherit” this persistence.
Financial Application: Realized Beta GARCH
A look at the Data
Financial Application: Realized Beta GARCH

Estimation Results: Latent Volatilities

Graphs showing the realized beta GARCH estimation results for RK CVX and RK SPY for the years 2007 to 2010.
Financial Application: Realized Beta GARCH
Estimation Results: Latent Correlation and Beta

- Realized Correlation
- Model Correlation

- Realized Beta
- Model Beta
Conclusion
Seven ways that HF data have improved volatility forecasting

1. HF data improve our understanding of the dynamic properties of volatility which is key for forecasting.

2. Realized measures are valuable predictors of future volatility in reduced form models.

3. Realized measures have enabled the development of new volatility models that provide more accurate forecasts.

4. HF data have improved the evaluation of volatility forecasts in important ways.

5. Realized measures can facilitate and improve the estimation of complex volatility models (e.g. continuous time models). Less estimation error improve predictions based on such models.

6. HF data have improved our understanding of the driving forces of volatility.

7. Deliver much more precise estimates and forecast of an assets beta.


