



Real-time Volatility Estimation Under Zero Intelligence

Jim Gatheral
“Celebrating Derivatives”
ING, Amsterdam
June 1, 2006

The opinions expressed in this presentation are those of the author alone, and do not necessarily reflect the views of of Merrill Lynch, its subsidiaries or affiliates.

Outline of this talk

- A non-intelligent model of the continuous double auction and its time series properties
- Uses of volatility forecasts
- Market microstructure bias
- A survey of estimation and forecasting algorithms
- Experimental results

I gratefully acknowledge insightful comments from Roel Oomen. All errors are mine.

Motivation

- We often need to estimate in-sample volatility. Typically we need a good volatility estimate to reduce errors in estimating parameters of
 - Market impact models
 - Limit order fill models
 - Volatility forecasting models: without good in-sample volatility estimates, how can we even assess the quality of volatility forecasts?
- From the theoretical perspective, we would like to understand
 - the nature of market microstructure noise
 - how to find an efficient estimator of integrated volatility

Uses of volatility forecasts

- Option valuation
- Risk estimation
- Order fill probability

For the first two of these, we need to estimate the width of the distribution of relatively long-timescale returns For the order fill probability, we need a volatility to estimate the first passage time density

Given that the underlying stochastic process is not Brownian motion, there is no a priori reason why the volatility numbers required for these quite different computations should be the same.

In what follows, we will focus on measuring the width of the distribution of returns.

Our approach

- One way to measure the performance of various volatility estimators is to simulate data and see directly how they perform.
- This has been done before:
 - Zhang, Mykland and Aït-Sahalia perform a numerical simulation of the Heston model with small additional iid Gaussian microstructure noise.
- Oomen computes results in closed-form in the context of a pure jump model again with an additional iid Gaussian microstructure noise.
- In each of these cases, there is a strong assumption on the form of the microstructure noise.
- In reality, we don't even believe in the artificial split between true price and noise processes.
- So, we proceed by simulating an artificial zero-intelligence market described and analyzed in detail by Smith, Farmer et al.

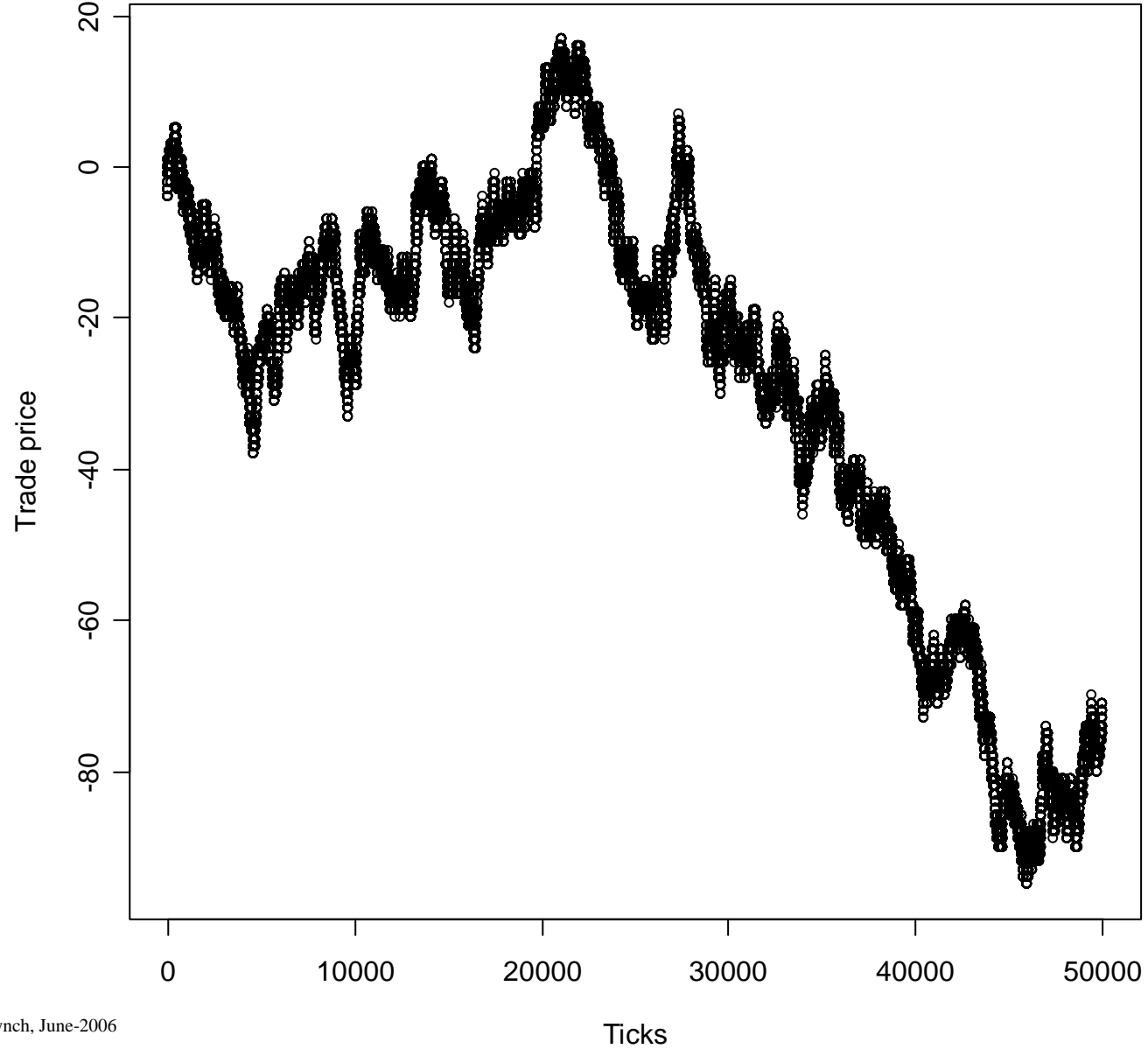
A non-intelligent (or zero-intelligence) model

Consider a model where market orders arrive randomly at rate μ , limit orders (per price level) arrive at rate α and a proportion δ per unit time of existing limit orders is canceled.

The behavior induced by such a model is rather complex and has been shown to mimic the behavior of real markets in some respects such as the distribution of spreads (Bollerslev et al. (1997)) and the relationship between volatility and volume (Farmer et al. (2005)).

Such a model is termed *non-intelligent* because order flow is random. In a real market, intelligent agents respond to orders with counter orders of their own.

A sample path



Parameters and sample properties

Smith Farmer zero-intelligence parameters are

$$\mu = 0.1$$

$$\alpha = 1$$

$$\delta = 0.2$$

Noise-to-signal ratio (see later for definition) $\gamma = 2.94$.

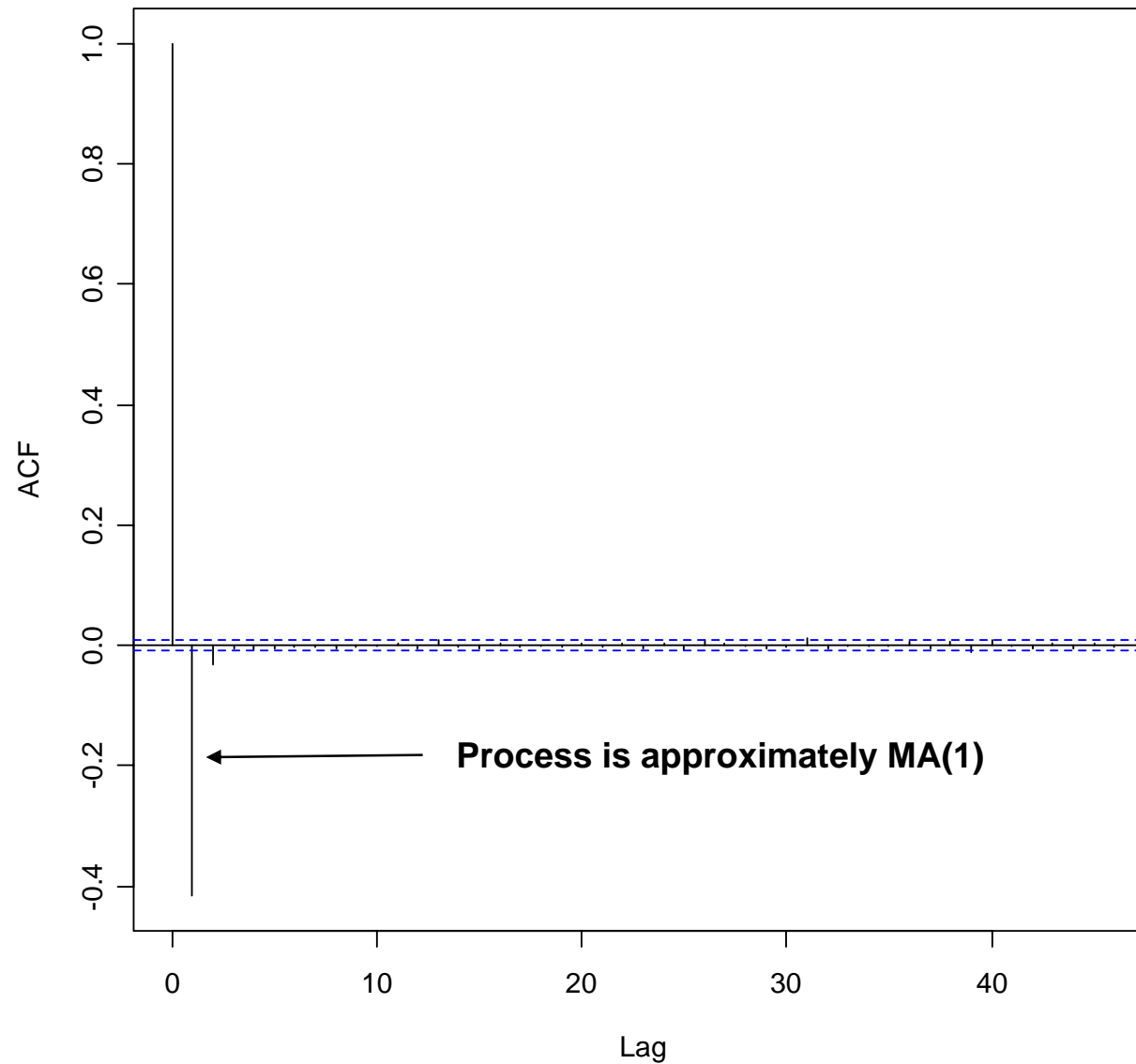
First order autocorrelation coefficient $\rho_1 = -0.43$

Parameters are chosen so that these last two are like real data.

Choice of sampling scheme

- Oomen has pointed out that it is important to sample in transaction time rather than in business time or calendar time.
 - In transaction time, empirically observed returns are MA(1)
 - In calendar time, with varying intensity, empirically observed returns are ARIMA.
- For our artificially-generated dataset where trade intensity is constant, it is most natural to sample each event, or in transaction time.
 - Our simulated process turns out to be approximately MA(1) just like real stock returns.

Autocorrelation plot



How should we compute historical (realized) volatility?

Given a set of tick data, how can we measure variance? The naïve answer would be to compute the statistic

$$\frac{1}{T} \sum_i^T \ln(S_i/S_{i-1})^2$$

where the S_i are successive prices in the dataset.

This estimator is known as Realized Variance (or RV).

Microstructure bias

In the limit of very high sampling frequency, RV picks up only market microstructure noise. To see this, suppose that the observed price Y_t is given by

$$Y_t = X_t + \epsilon_t$$

where X_t is the value of the underlying (log-)price process of interest and ϵ_t is a random market microstructure-related noise term. Suppose there are $n + 1$ ticks (so there are n price changes) in the time interval T . Then as $n \rightarrow \infty$,

$$[Y, Y] := \sum_i^T (Y_i - Y_{i-1})^2 \sim [X, X] + 2n \text{Var}[\epsilon] = n\sigma^2 + 2n \text{Var}[\epsilon]$$

where we assume that the *per tick* variance σ^2 is constant.

Oomen's noise-to-signal ratio

Following Oomen, we define the noise-to-signal ratio

$$\gamma = \frac{Var[\epsilon]}{\sigma^2}$$

and note in passing that $Var[\epsilon]$ may be efficiently estimated using

$$Var[\epsilon] \sim - \sum_i^T (Y_{i+1} - Y_i) (Y_i - Y_{i-1})$$

The conventional solution

- The conventional solution is to sample at most every five minutes or so.
- Suppose there are 78,000 trades per day (CSCO for example). Then there are roughly 1,000 trades every 5 minutes. Sampling only every 5 minutes corresponds to throwing out 99.9% of the points!
- To quote Zhang, Mykland and Aït-Sahalia, “It is difficult to accept that throwing away data, especially in such quantities, can be an optimal solution.”
- From a more practical perspective, if we believe that volatility is time-varying, it makes sense to try and measure it from recent data over the last few minutes rather than from a whole day of trading.

The multiple grid estimator

Denote by Y_{jk} ($1 \leq j \leq k$), the subsample of Y obtained by sampling every k ticks from the j th tick on. There are clearly k such non-overlapping subsamples in the dataset, the first one beginning with the first tick, the second with the second tick and so on. Then, for each Y_{jk} , we have

$$[Y_{jk}, Y_{jk}] \sim [X_{jk}, X_{jk}] + 2 n_{jk} \text{Var}[\epsilon]$$

Now define the average subsampled realized variance

$$[Y, Y]_k := \frac{1}{k} \sum_j^k [Y_{jk}, Y_{jk}]$$

and suppose further that the *per tick* variance σ^2 is constant. Then

$$[Y, Y]_k \sim \bar{n}_k k \sigma^2 + 2 \bar{n}_k \text{Var}[\epsilon]$$

where \bar{n}_k is the average number of ticks in each subsample. We obtain the estimator

$$v_k^{\text{Realized}} = \frac{[Y, Y]_k}{\bar{n}_k k}$$

The ZMA estimator

With our earlier assumption of iid noise, we can eliminate bias for each k by forming

$$[Y, Y]_k - \frac{\bar{n}_k}{n} [Y, Y] \sim \bar{n}_k k \sigma^2 - \bar{n}_k \sigma^2$$

Thus we obtain the Zhang-Mykland-Ait-Sahalia (ZMA) bias-corrected estimator of σ^2 :

$$v_k^{ZMA} := \frac{1}{\bar{n}_k (k-1)} \left\{ [Y, Y]_k - \frac{\bar{n}_k}{n} [Y, Y] \right\}$$

The Zhou estimator

Define

$$[Y, Y]^Z := \sum_i^T (Y_i - Y_{i-1})^2 + \sum_i^T (Y_i - Y_{i-1})(Y_{i-1} - Y_{i-2}) + \sum_i^T (Y_i - Y_{i-1})(Y_{i+1} - Y_i)$$

It's easy to check that under the assumption of serially uncorrelated noise independent of returns,

$$\mathbb{E} \left[[Y, Y]^Z \right] = \langle X \rangle_T$$

It follows that

$$[Y_{jk}, Y_{jk}]^Z \sim (n_{jk} - 2) \sigma^2$$

As suggested by Zhou himself, we may compute his estimator from subsamples of the data obtaining

$$v_k^{Zhou} := \frac{1}{\bar{n}_k - 2} [Y, Y]_k^Z$$

GARCH

For each subsample Y_{1k} , we use MLE to estimate the parameters of the GARCH process

$$\sigma_i^2 = a_0 + a_1 r_{i-1}^2 + b_1 \sigma_{i-1}^2$$

We take as our variance estimate, the last estimated value of $\sigma^2 \sim \sigma_{\bar{n}_k}^2$.

Note that we graph a kernel-smoothed version of the GARCH estimator.

The drift-adjusted high-low-open estimator

$$\sigma^2 n = \frac{1}{2} \mathbb{E} \left[\log^2 \left(\frac{high}{open} \right) + \log^2 \left(\frac{low}{open} \right) \right] - \frac{1}{2} \mathbb{E} \left[\log \left(\frac{high}{open} \right) + \log \left(\frac{low}{open} \right) \right]^2$$

Once again, this estimator may be computed for each subsample and averaged.

Discreteness adjustment for open-high-low estimator

Suppose we have k ticks altogether and we sample every tick. Then if \tilde{x} is the log of the true maximum and \tilde{x}_k is the log of the discrete maximum, we should have

$$\tilde{x} = \tilde{x}_k + \delta_k$$

where δ_k is a non-negative random variable. Then from Magdon-Ismail and Atiya (2003) (who in turn quote Rogers and Satchell (1991)), to leading order in the time δt between ticks,

$$\alpha \sigma := \mathbb{E} [\delta_k] = \sqrt{2\pi} \left(\frac{1}{4} + \frac{\sqrt{2}-1}{6} \right) \sigma + o(\delta t)$$

$$\beta \sigma^2 = \mathbb{E} [\delta_k^2] \approx \left(\frac{1}{12} + \frac{\pi}{16} \right) \sigma^2 + o(\delta t)$$

where σ is the volatility *per tick*.

Final formula adjusted for discreteness

$$\sigma \sqrt{k - 2\alpha^2 + \beta} = \sqrt{A - B + \frac{\alpha^2 C^2}{4(k - 2\alpha^2 + \beta)}} + \frac{\alpha C}{2\sqrt{k - 2\alpha^2 + \beta}}$$

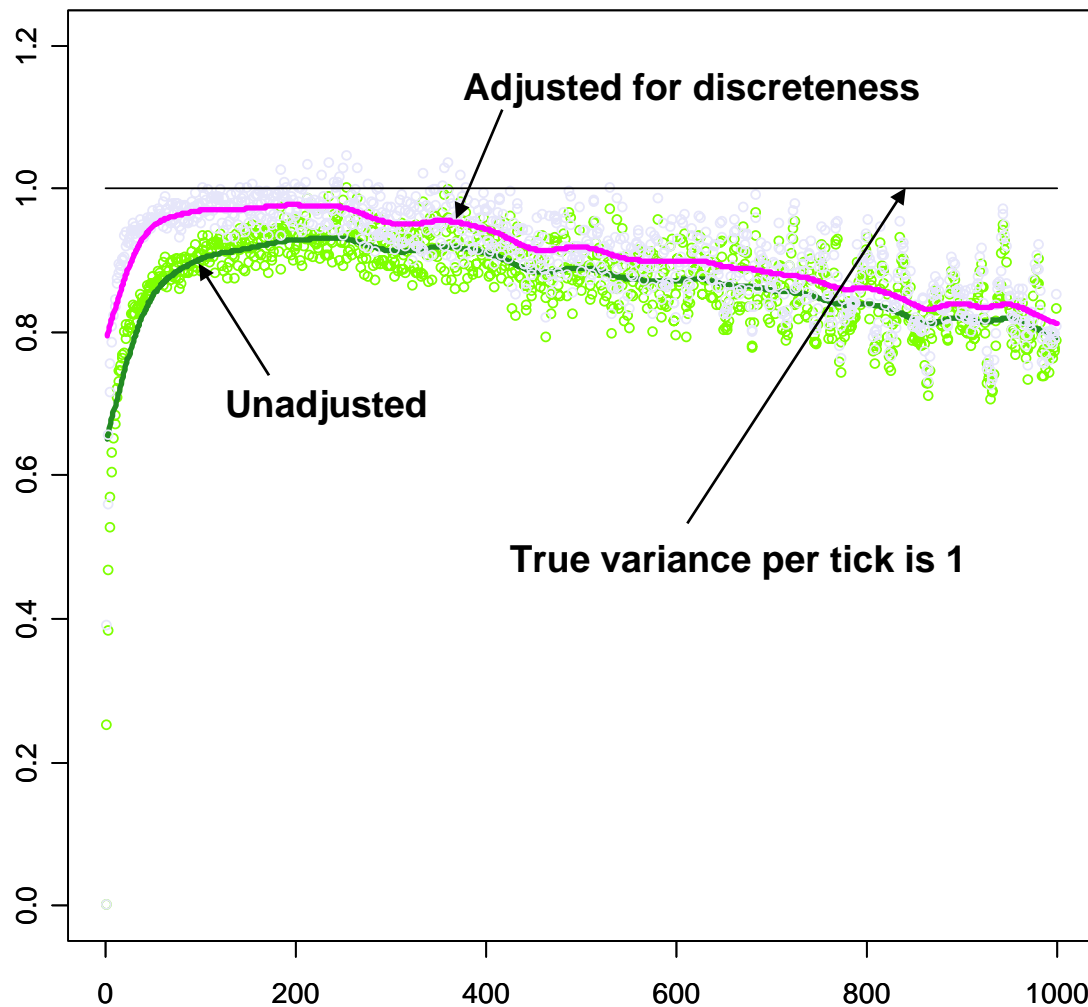
where

$$A := \frac{1}{2} \mathbb{E} [\tilde{x}_k^2 + \underline{x}_k^2]$$

$$B := \frac{1}{2} \mathbb{E} [\tilde{x}_k + \underline{x}_k]^2$$

$$C := \mathbb{E} [\tilde{x}_k - \underline{x}_k]$$

Effect of discreteness adjustment on Open-High-Low estimator (artificial normal random walk data)

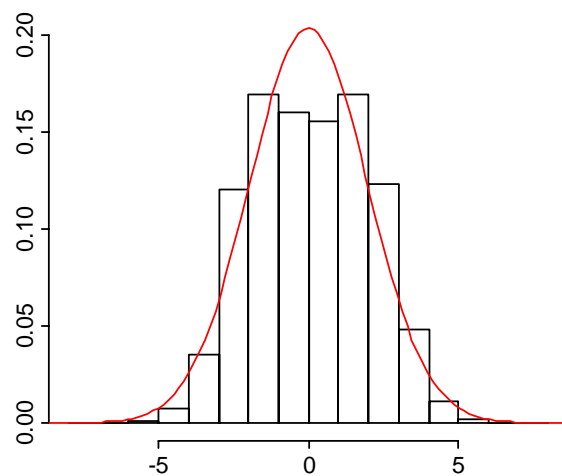


Measuring the true variance of ZI data

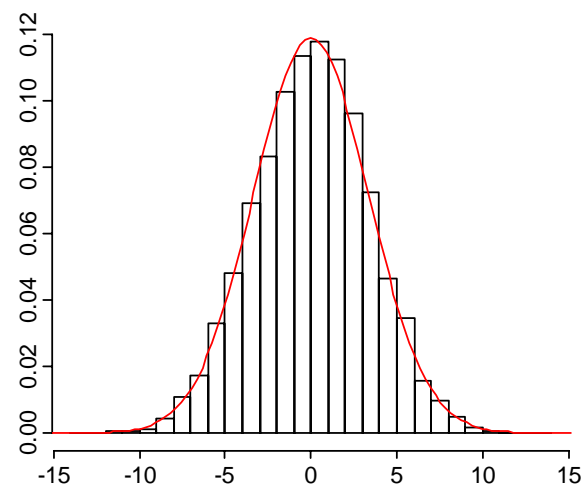
- We generate 10,000 paths of the process with 1,000 timesteps.
- For each time slice, we generate a distribution of returns and compute its variance.

Return distributions

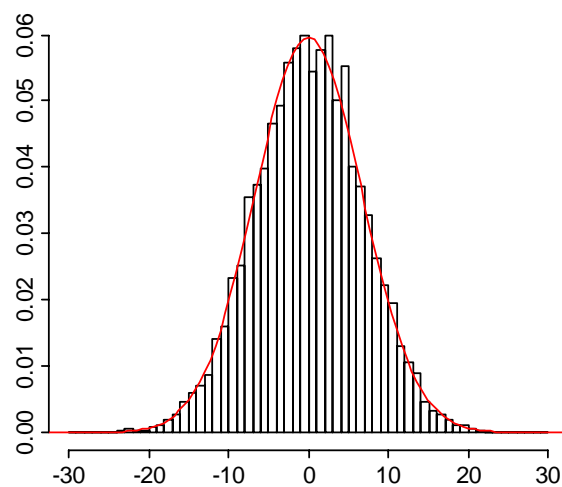
After 10 trades



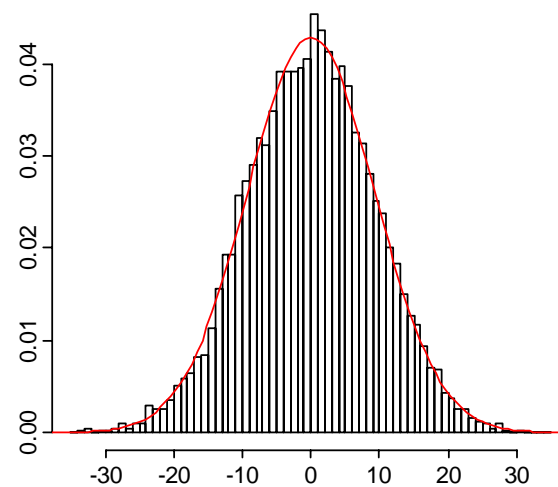
After 100 trades



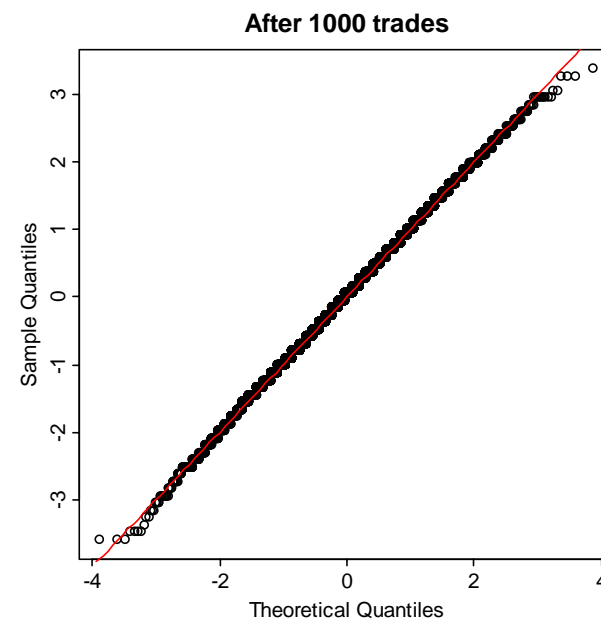
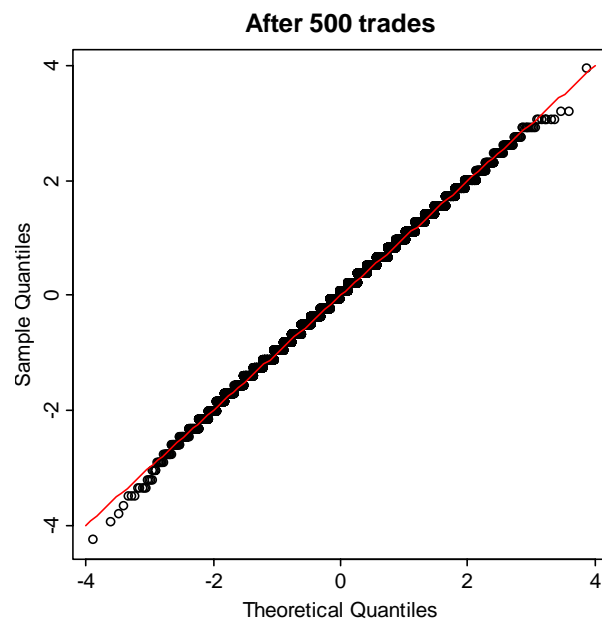
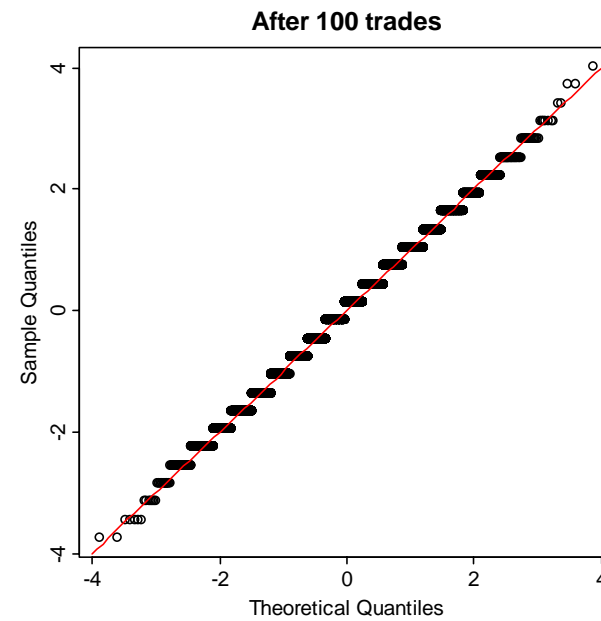
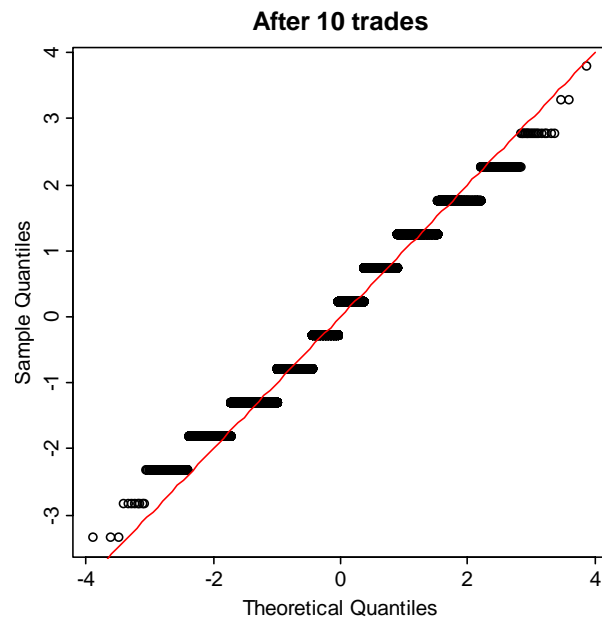
After 500 trades



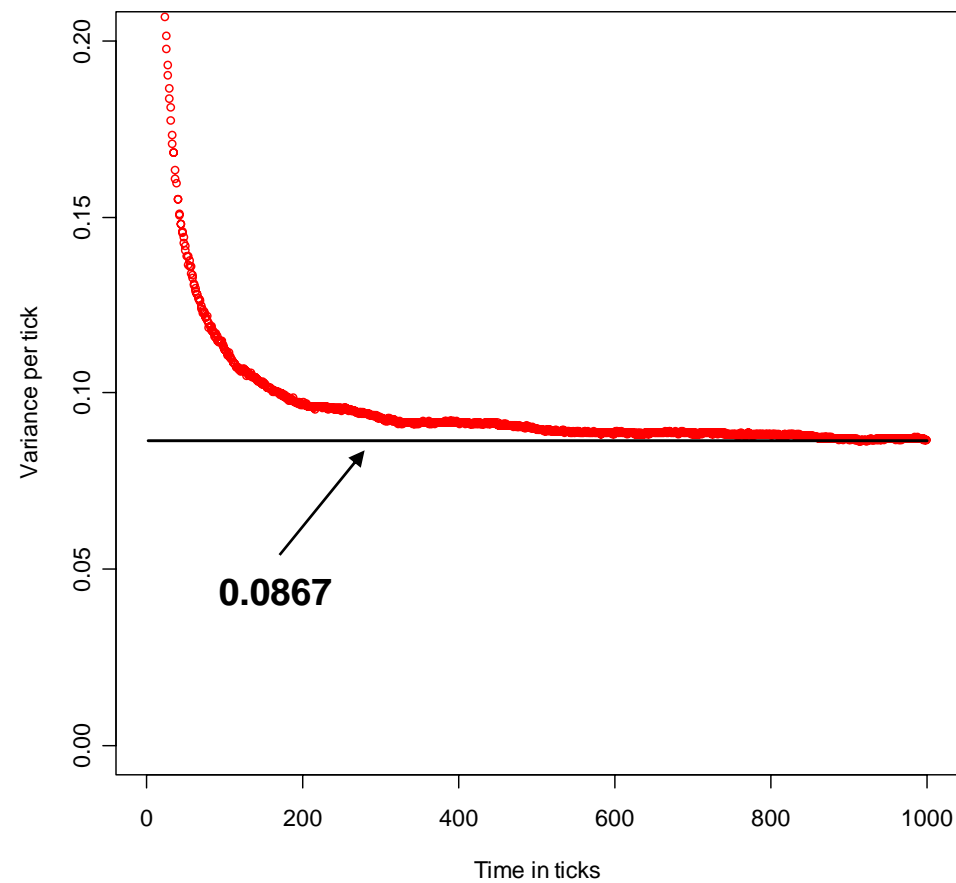
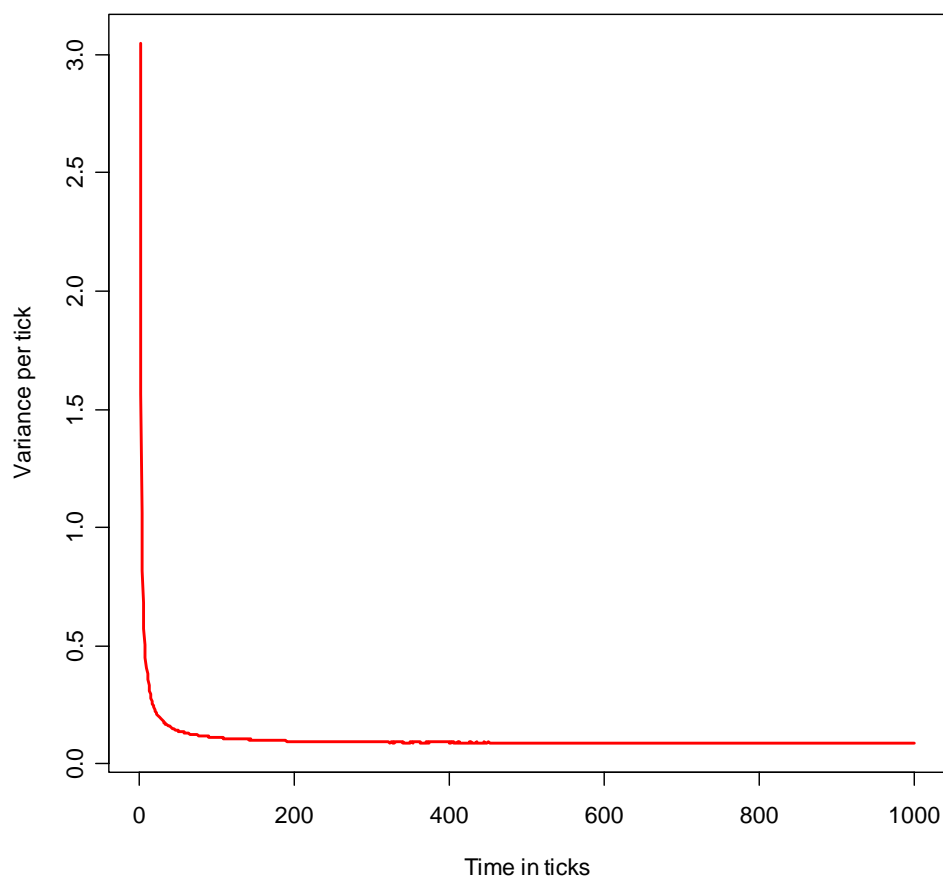
After 1000 trades



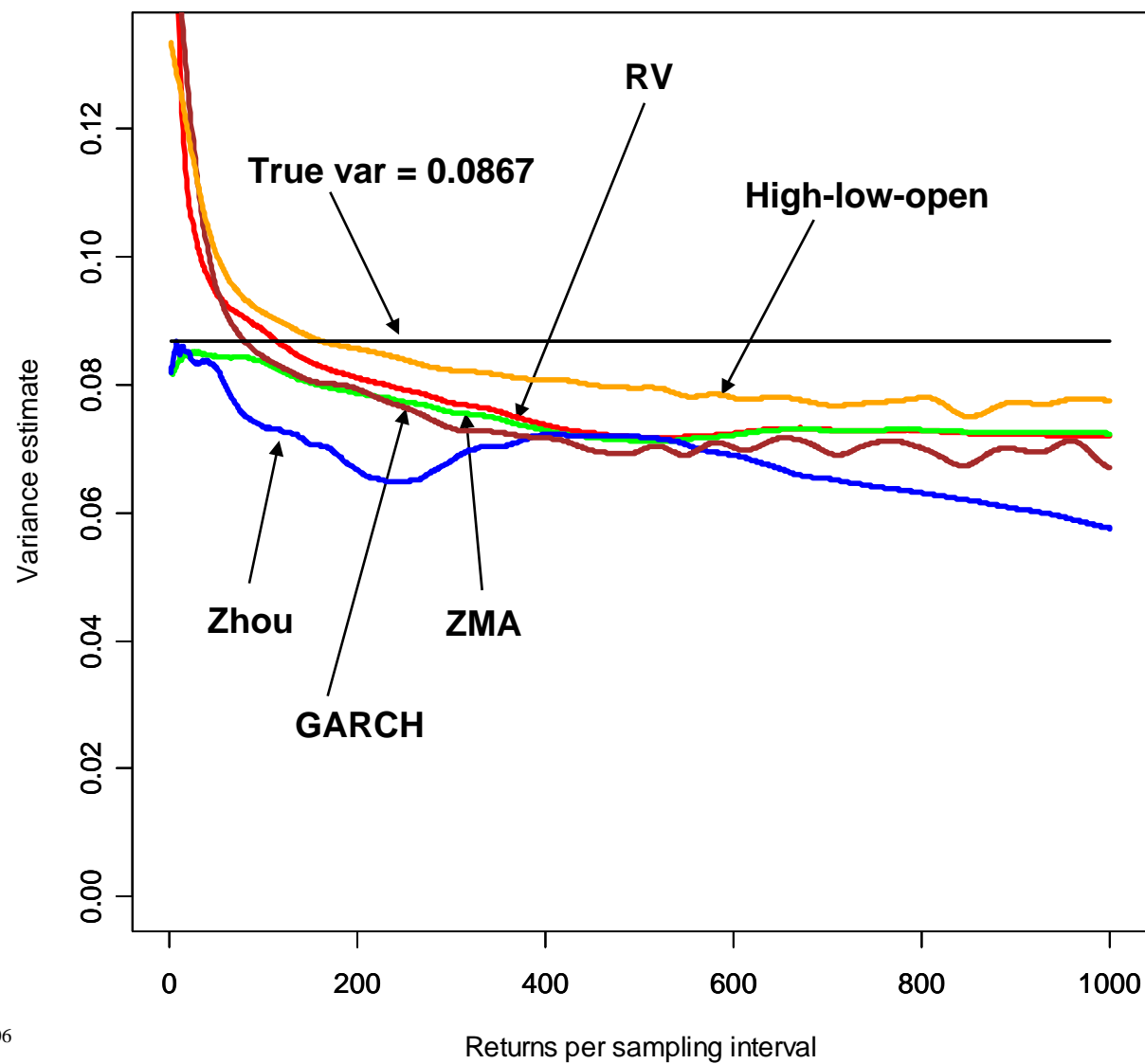
QQ plots



True variance vs cumulative number of trades

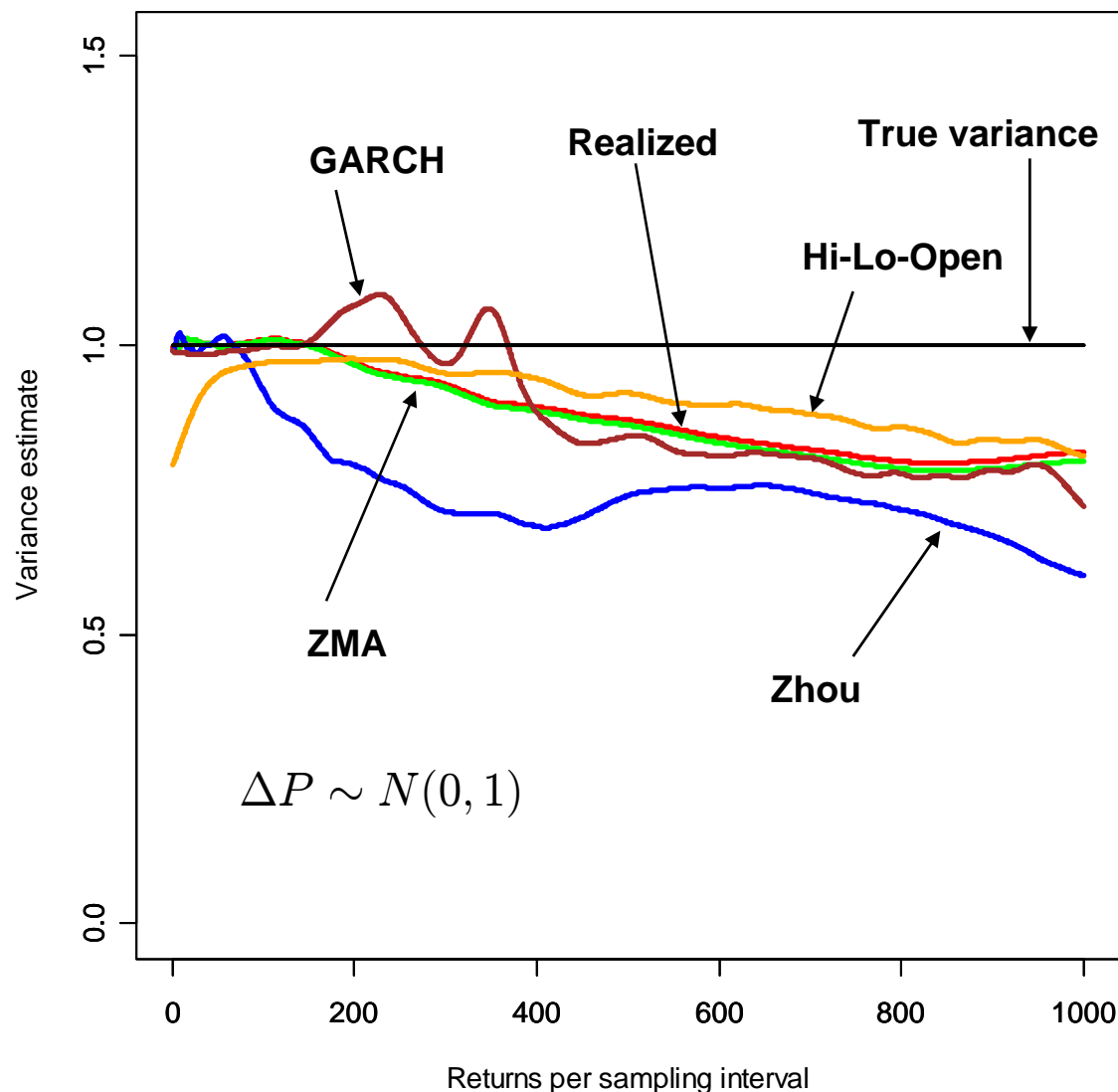


Performance of various estimators on ZI data

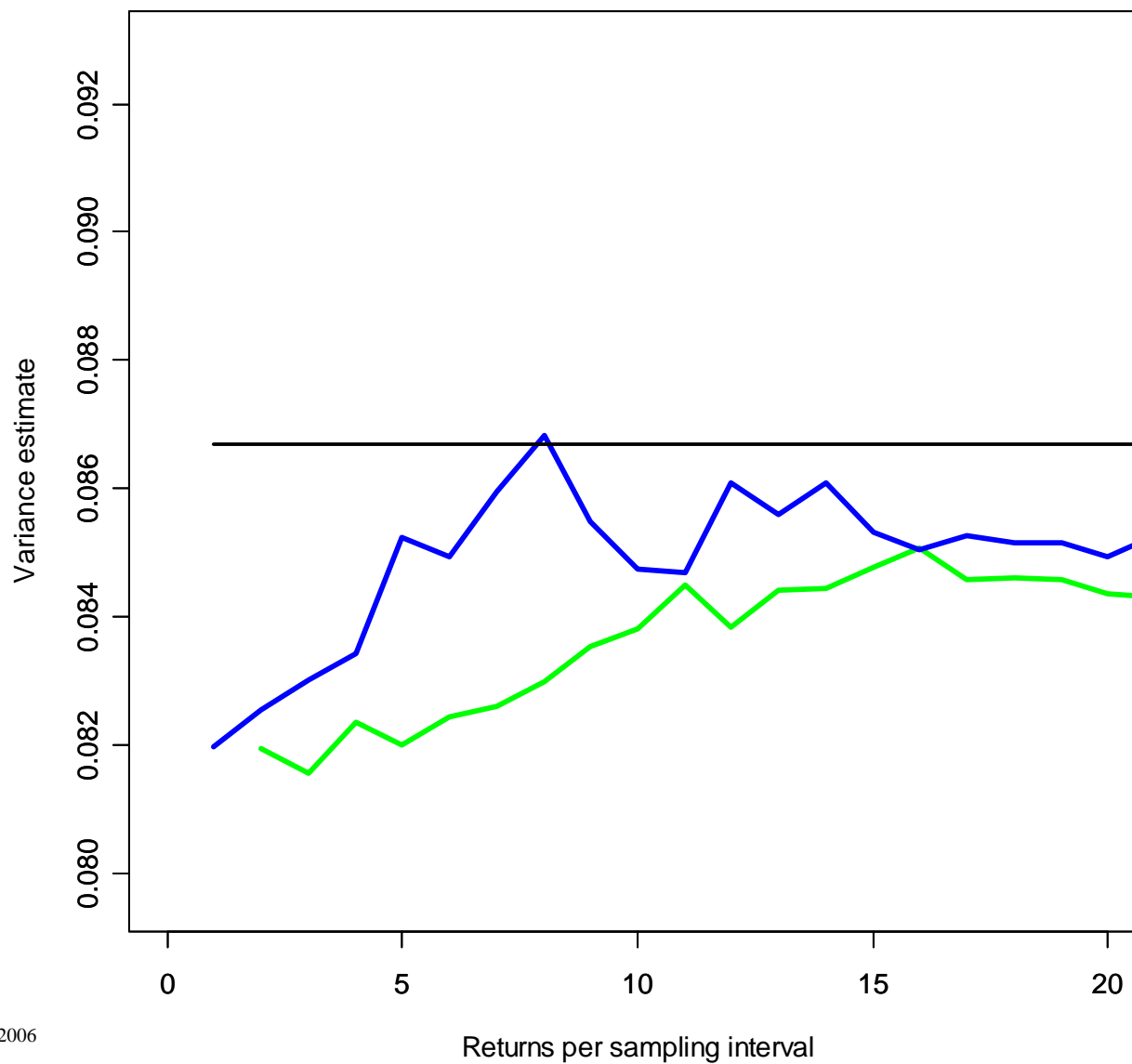


Are the computations correct: a check

- All estimators except hi-lo-open give an estimated variance per tick of one at the highest sampling frequency.
- The hi-lo-open estimator approaches one at about 100 sampling points per interval.



Detail of plot of Zhou and ZMA performance



Interim conclusion

- Both RV and GARCH are very biased at high frequency
- GARCH is seen to be very similar to RV
 - From its definition, GARCH is like exponentially smoothed RV
- The high-low-open estimator is seen to underperform.

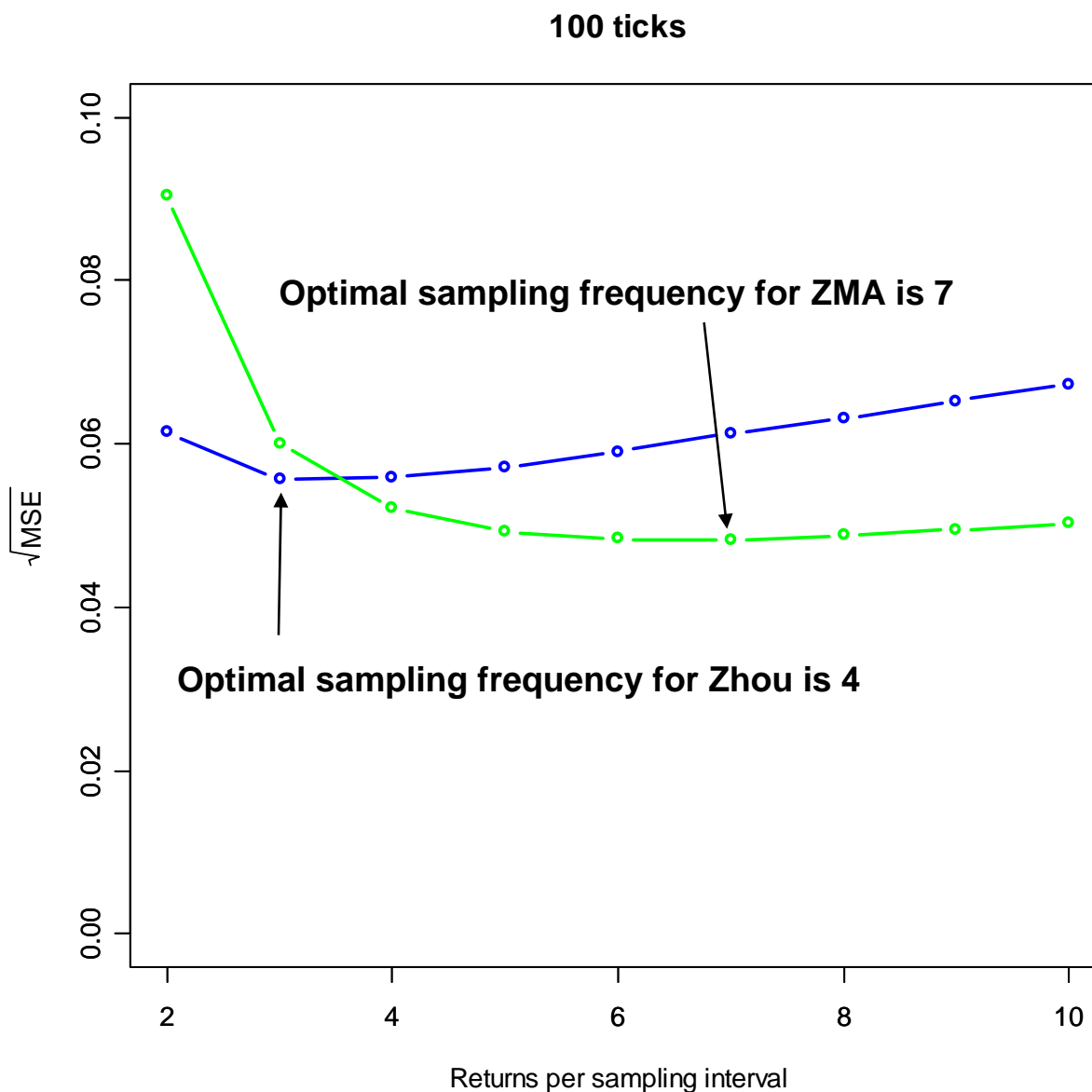
- The only estimators worth considering at are Zhou and ZMA.
 - On this path, the Zhou estimator definitely performed better.
- It seems from the plot that
 - The Zhou estimator may have lower bias but greater variance
 - The ZMA estimator may have greater bias but lower variance

- So we take an experimental look at Mean Squared Error (MSE) for these estimators.

$$\text{MSE} = \text{Bias}^2 + \text{Variance}(v)$$

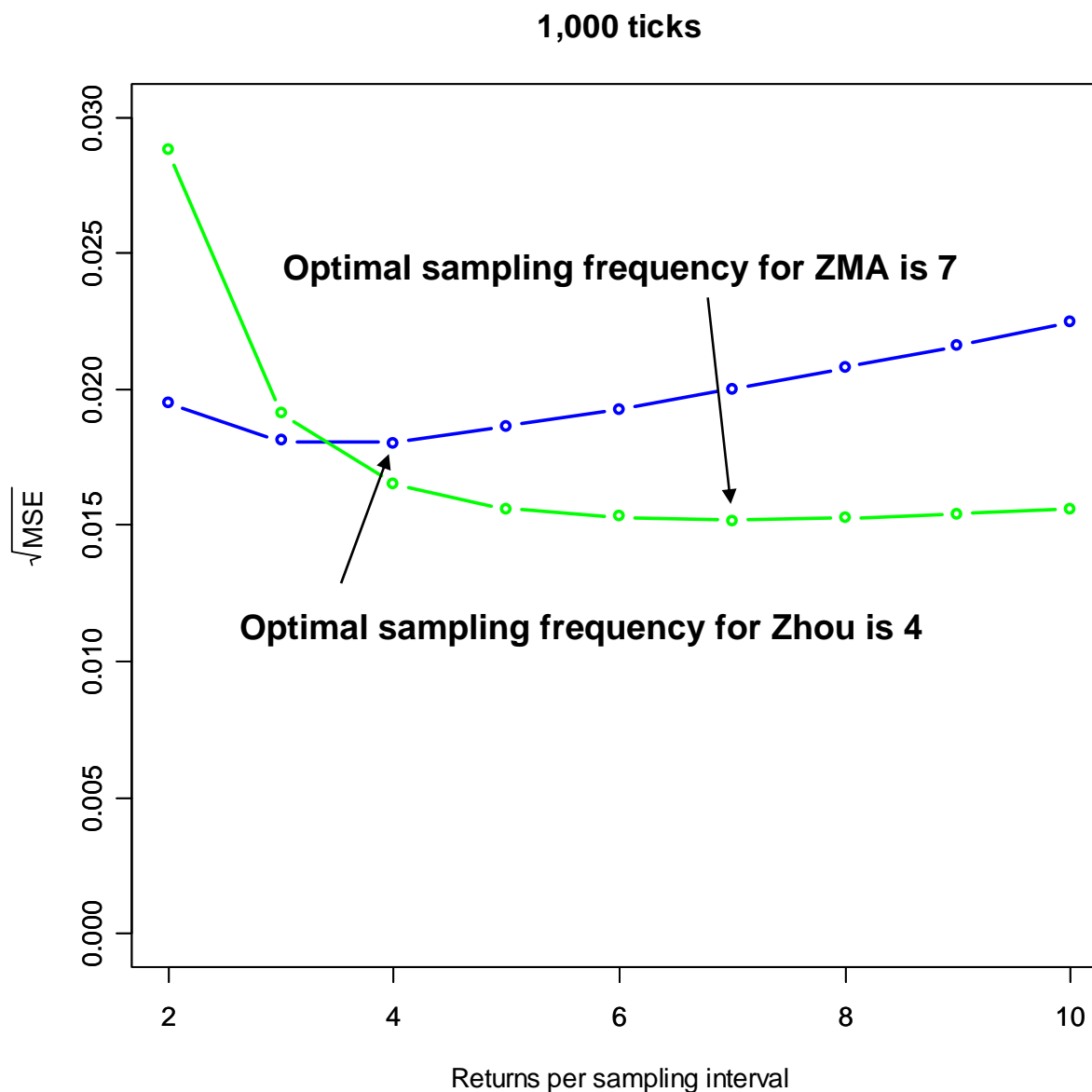
Mean Squared Error (MSE) vs sampling frequency

- As we guessed, ZMA is more biased but Zhou has higher variance.
- For a sample size of 100 trades, the winner is ZMA with optimal sampling frequency of 7.
- An illiquid stock may trade around 100 times in 10 minutes.



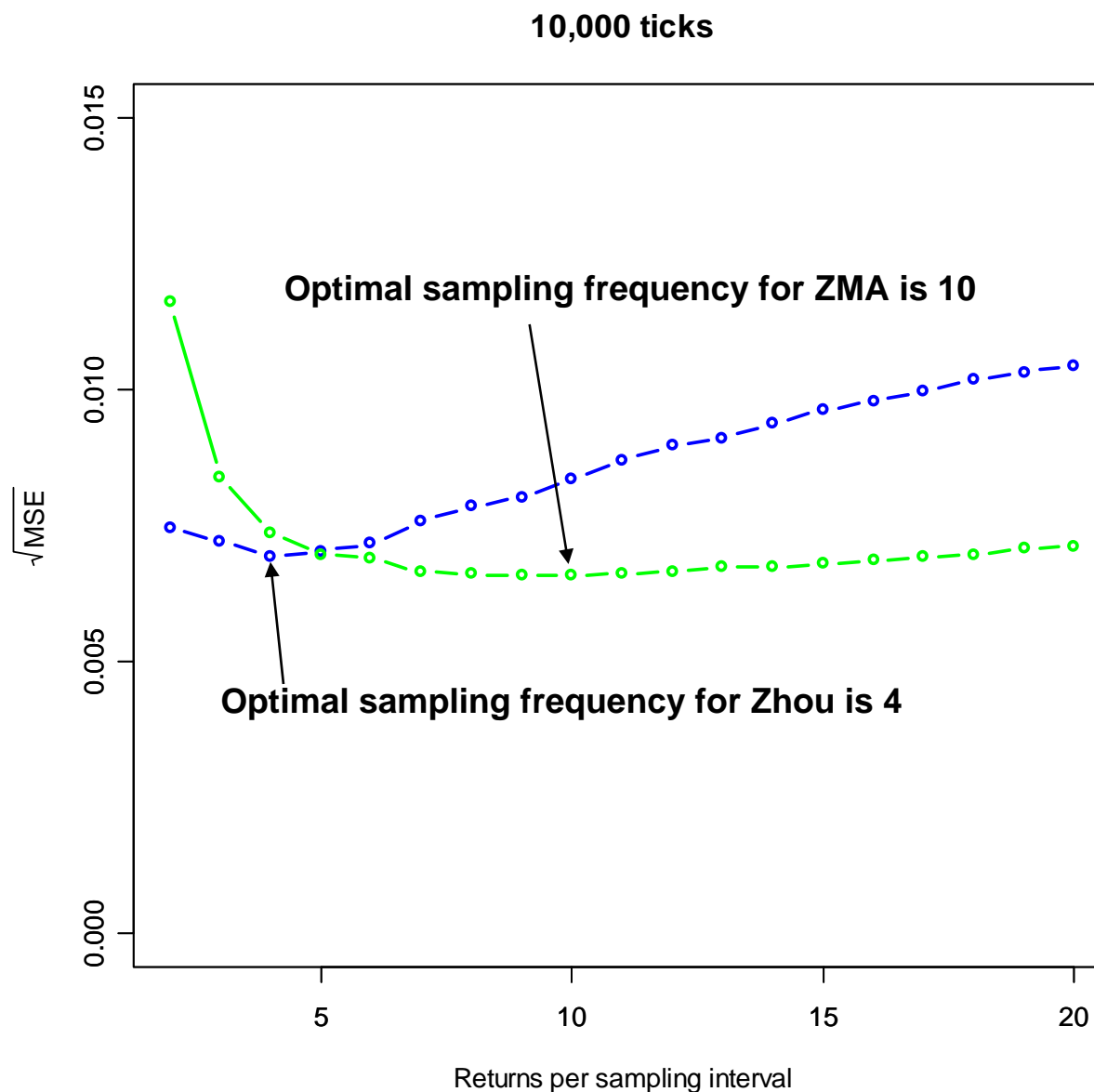
Mean Squared Error (MSE) vs sampling frequency

- For a sample size of 1,000 trades, the winner is ZMA with optimal sampling frequency of 7.
- A liquid stock can have 1,000 trades in 5 minutes.



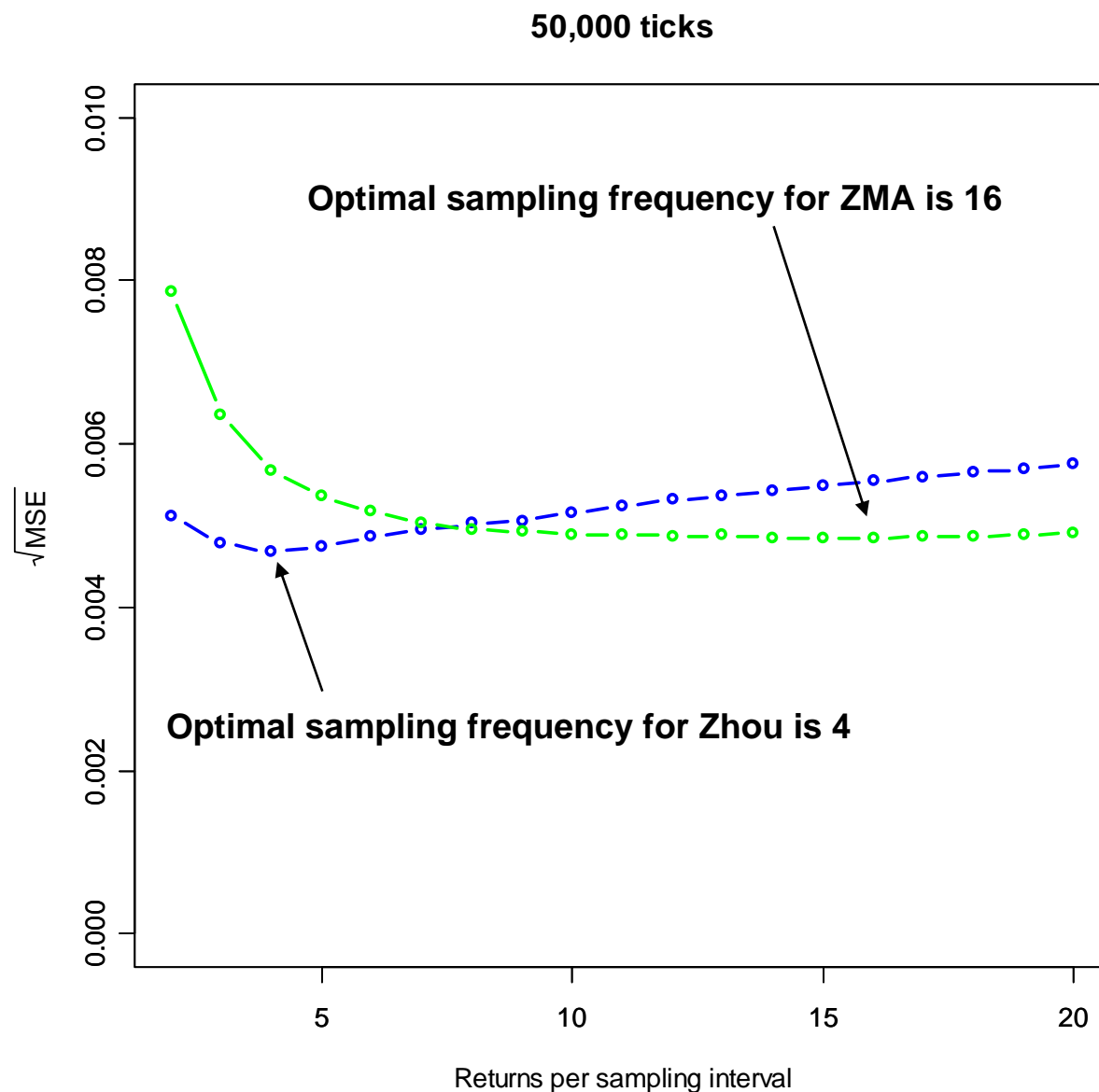
Mean Squared Error (MSE) vs sampling frequency

- For a sample size of 10,000 trades, the winner is Zhou with optimal sampling frequency of 4.
- A reasonably liquid stock can have 1,000 trades in one day.

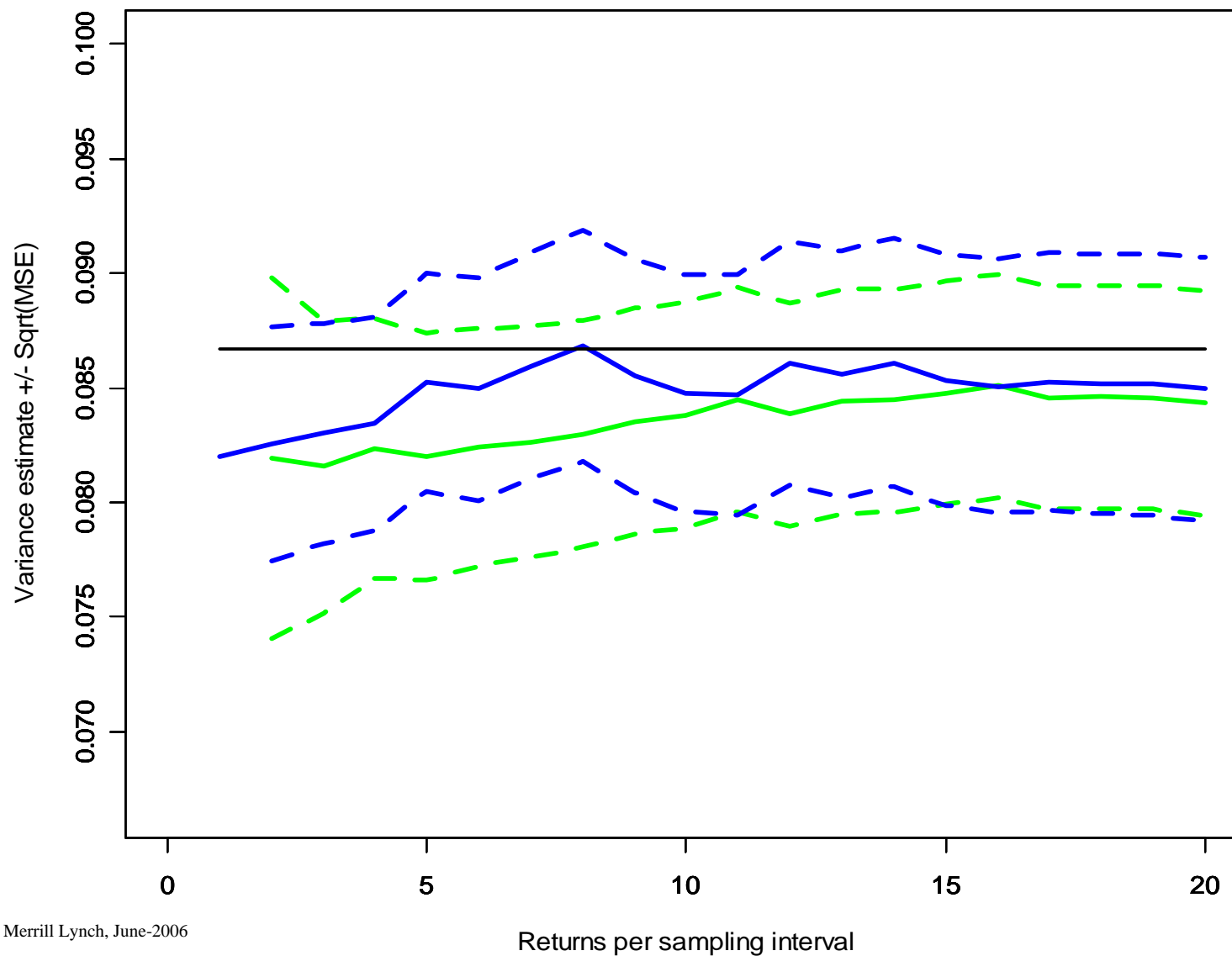


Mean Squared Error (MSE) vs sampling frequency

- For a sample size of 50,000 trades, the winner is Zhou with optimal sampling frequency of 4.
- A liquid stock may have 50,000 trades in one day.



Plot of Zhou and ZMA performance with error bars



Now we analyze real data...

Oomen's data cleaning rule

We eliminate trades with non-blank condition codes; this has the side-benefit of eliminating all trades after the close. We further adopt Oomen's cleaning procedure where outliers are removed if there is an instantaneous price reversal as follows.

Let P_k denote the logarithmic price at which the k th transaction is executed. For P_k to be removed by the filter, the following conditions need to be satisfied:

$$|P_k - P_{k-1}| > c$$

and

$$\frac{\Delta P_{k+1}}{\Delta P_k} \in [-1 - w, -1 + w] \text{ for } 0 < w < 1$$

Again following Oomen, we set $w = 0.25$ and $c = 8\sigma$ where σ is some reasonable estimate of standard deviation of tick-by-tick returns (the sample standard deviation in this case).

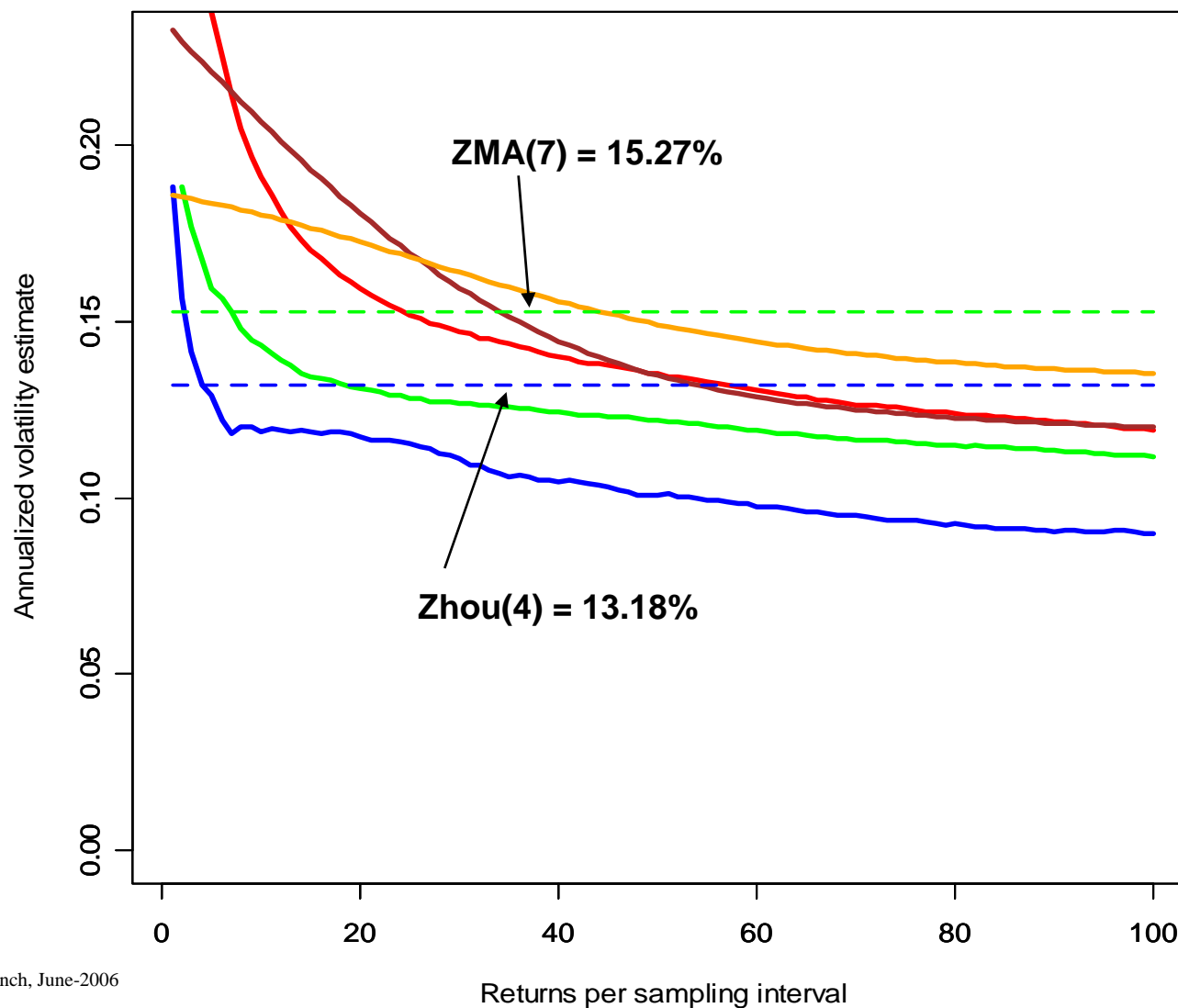
Data

- Average number of trades over the 69-day period from 03-Jan-2006 to 07-Apr-2006 is as follows:
 - IBM: 9,227
 - CSCO: 72,471
 - PFE: 25,515
 - COCO: 3,610

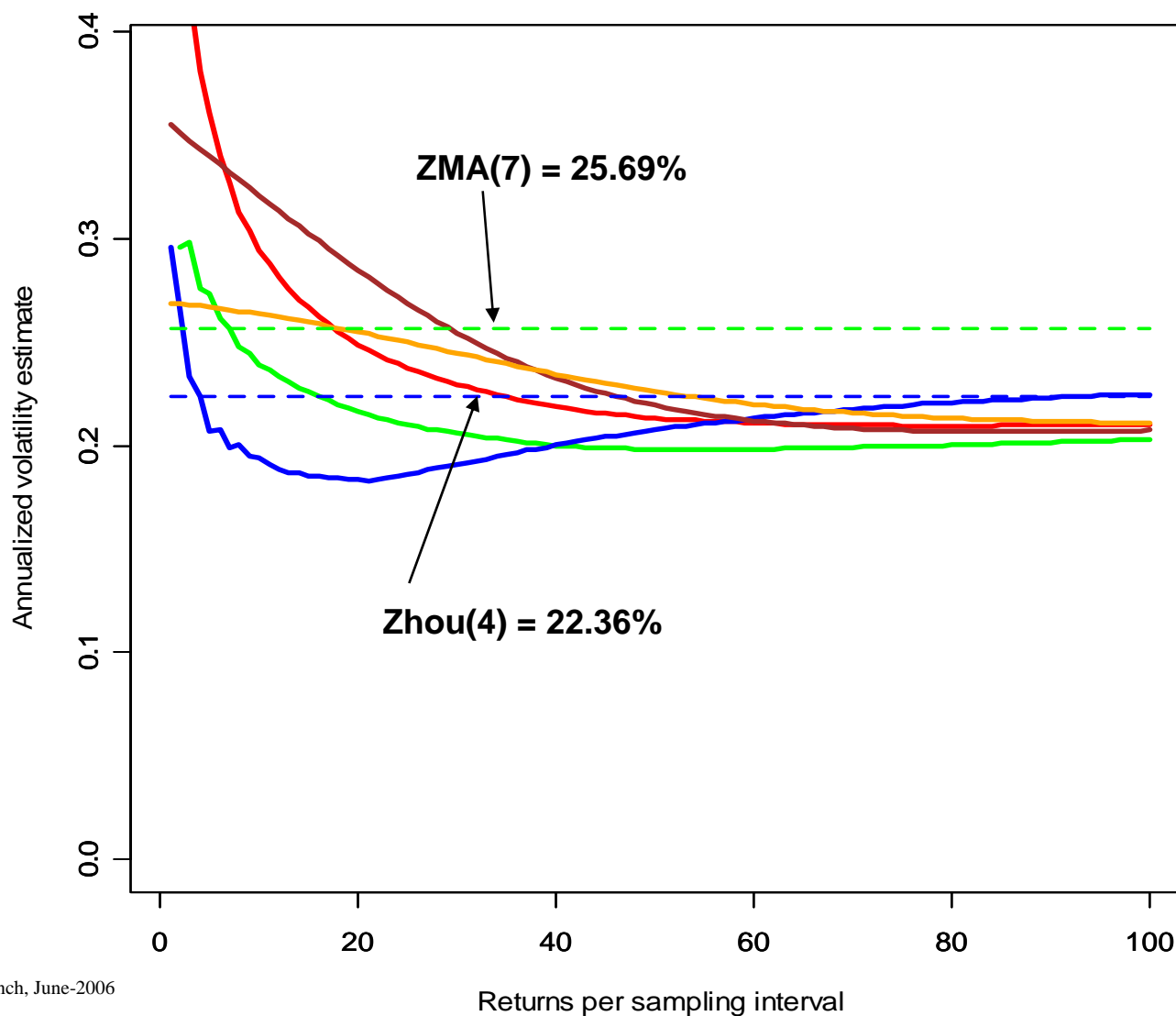
- After cleaning, number of trades in these stocks on 06-Apr-2006 was:
 - IBM: 6,729
 - CSCO: 81,483
 - PFE: 18,943
 - COCO: 2,975

- So Thursday 06-Apr-2006 was a typical day.

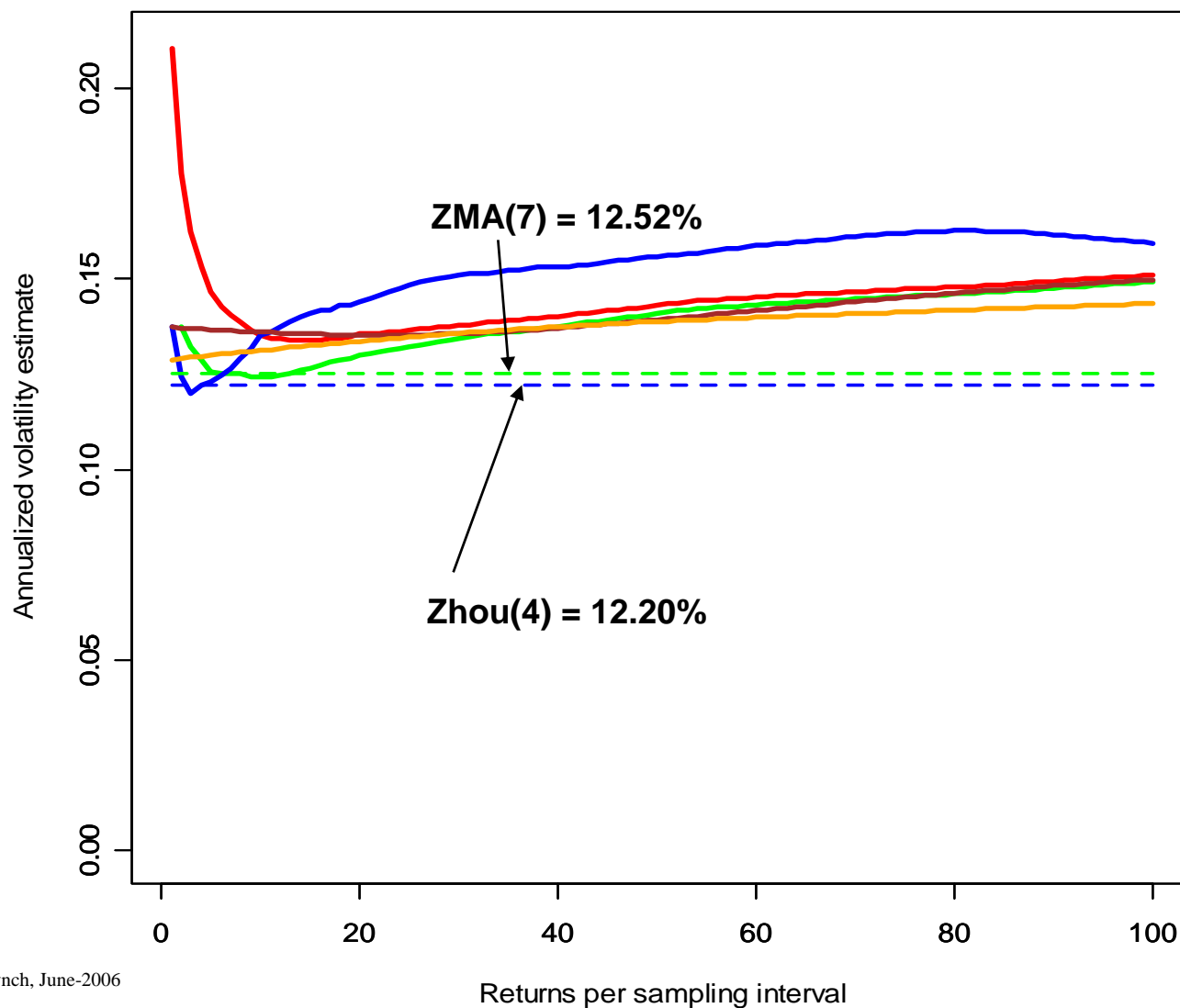
Comparison of estimates for PFE (06-Apr-2006)



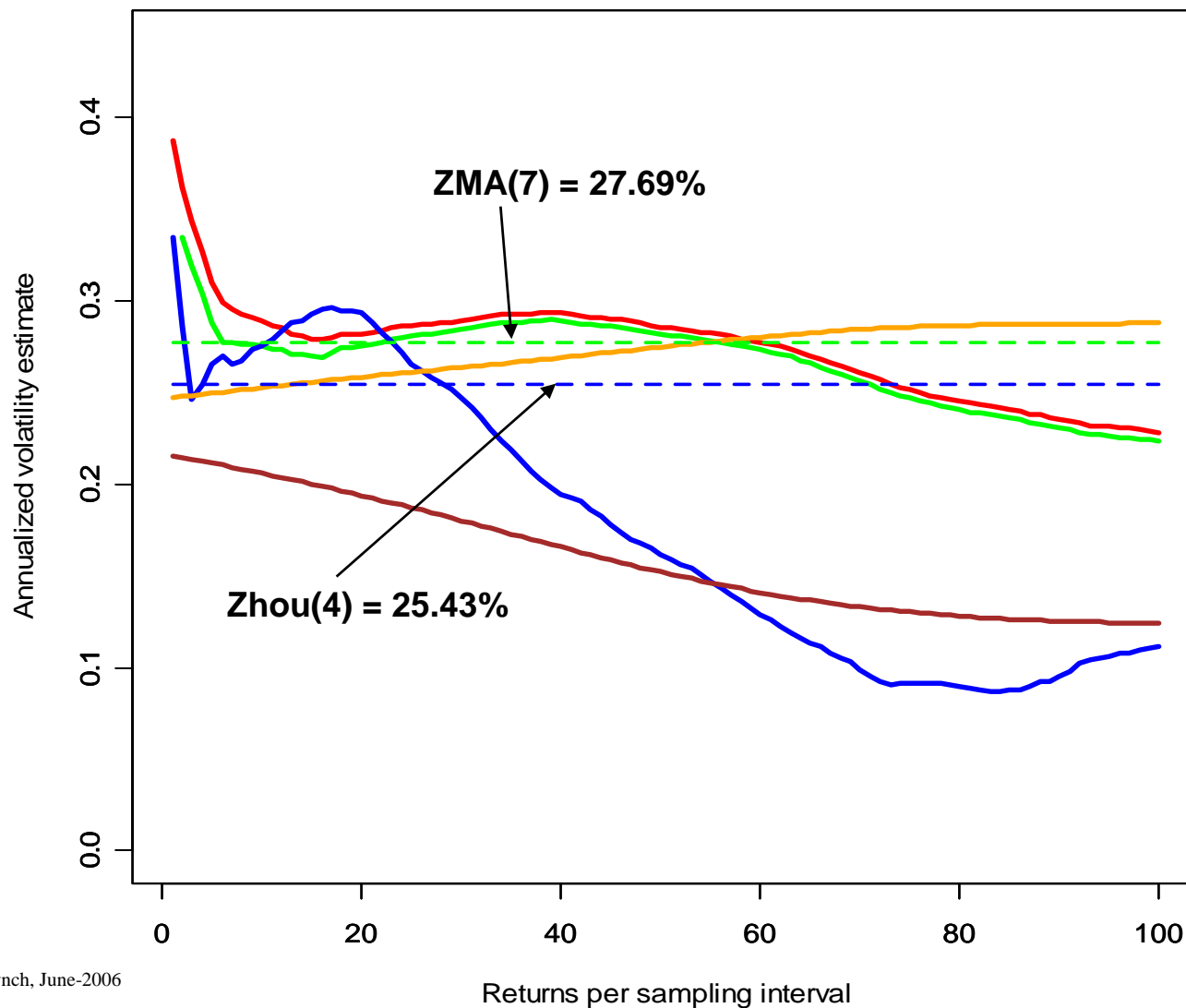
Comparison of estimates for CSCO (06-Apr-2006)



Comparison of estimates for IBM (06-Apr-2006)



Comparison of estimates for COCO (06-Apr-2006)



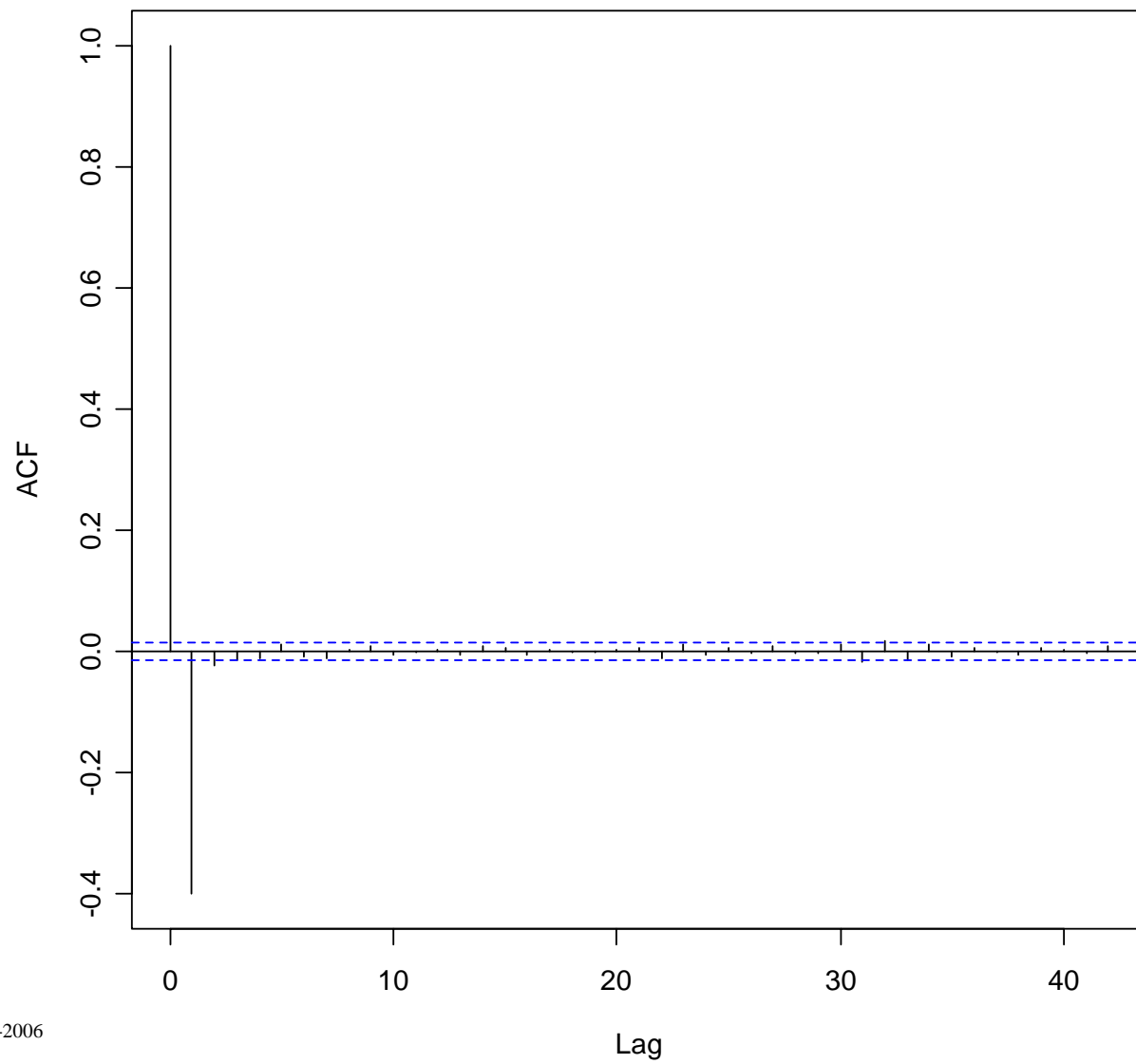
Summary of results

The following table compares estimated and implied volatilities (from Bloomberg) as of 06-Apr-2006 for the four stocks studied:

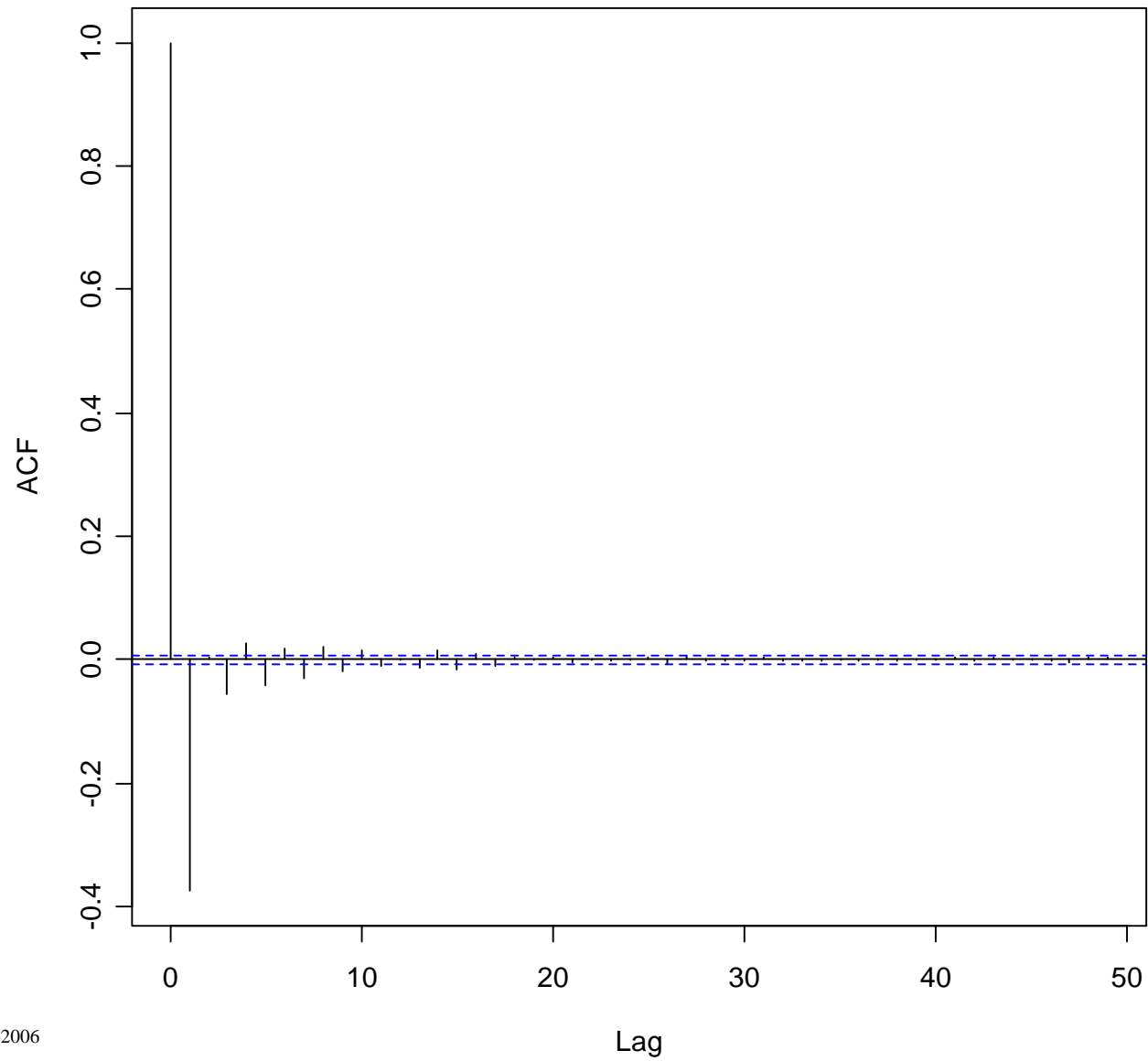
	σ_{ZMA}	σ_{Zhou}	$\sigma_{Implied}$
PFE	15.27%	13.18%	19.90%
CSCO	25.69%	22.36%	30.50%
IBM	12.52%	12.20%	16.41 %
COCO	27.69 %	25.43%	41.92 %

The bias might in large part be explained by accounting for overnight moves. Note also that the ZMA estimate is always higher than the Zhou estimate.

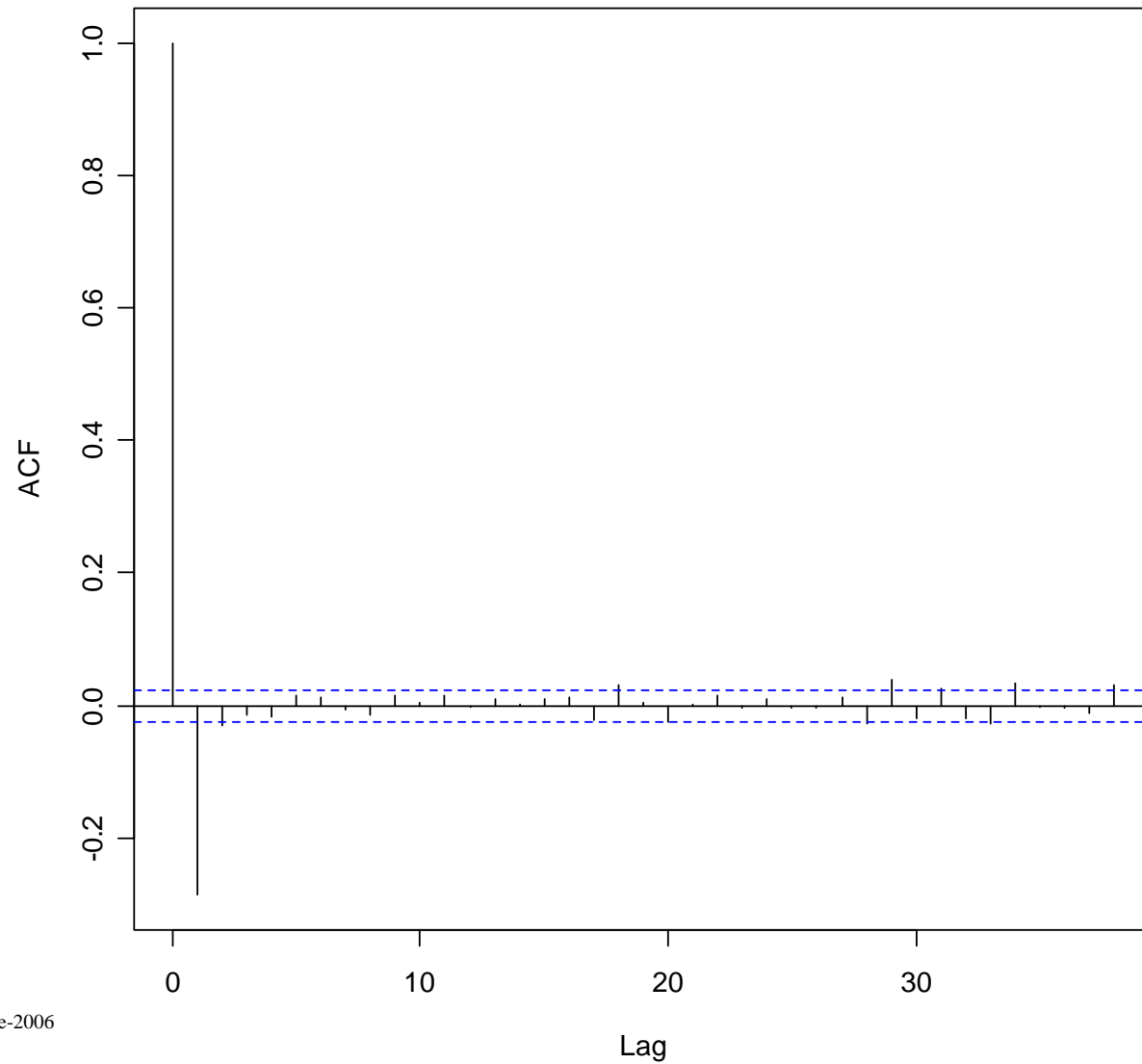
PFE autocorrelation function



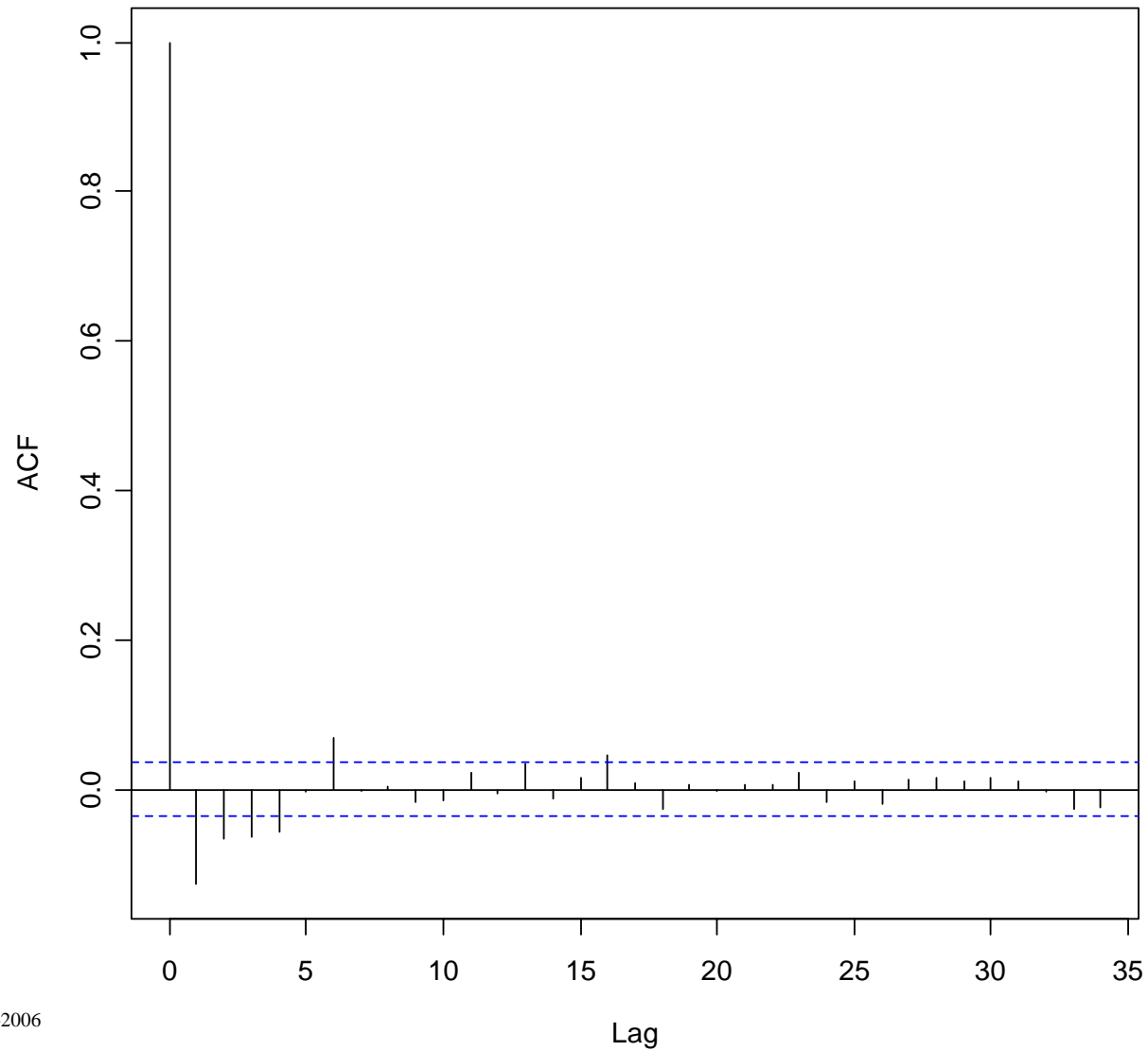
CSCO autocorrelation function



IBM autocorrelation function



COCO autocorrelation function



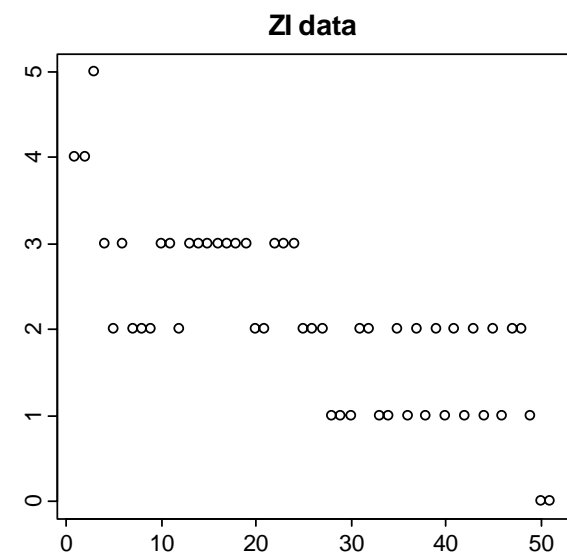
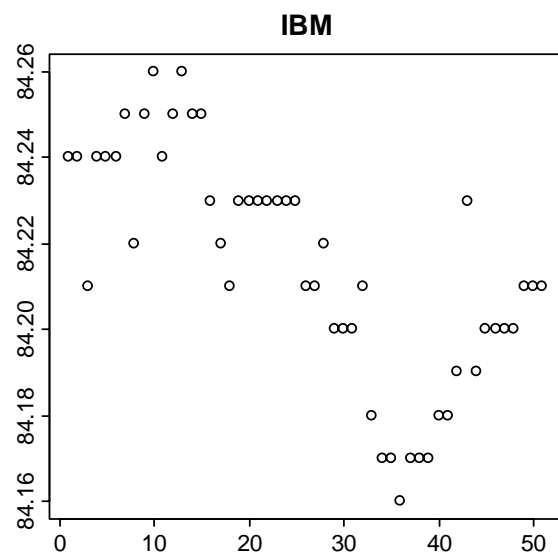
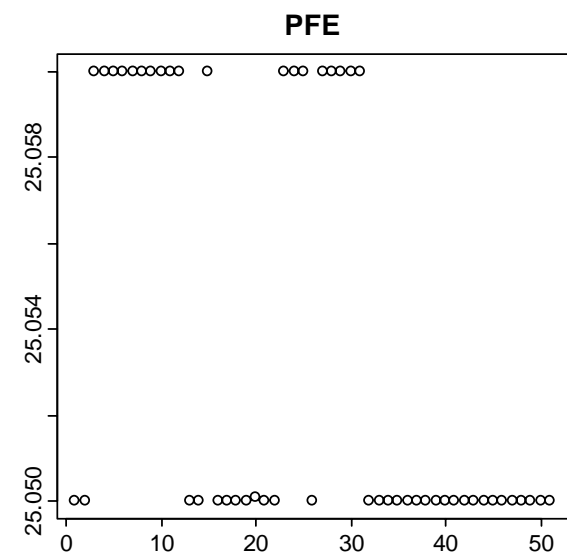
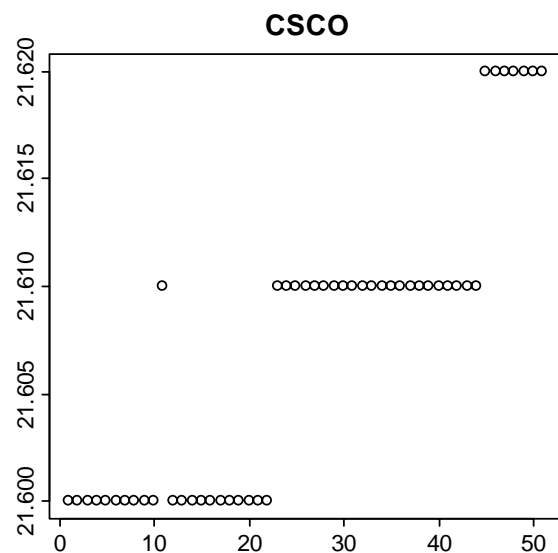
Noise-to-signal ratios (γ) and first order autocorrelations

	γ	γ^*	ρ_1	ρ_1^*
PFE	4.17	1.73	-0.40	-0.40
CSCO	2.63	2.22	-0.38	-0.34
IBM	0.85	0.18	-0.29	-0.12
COCO	0.30	0.35	-0.13	-0.17
ZI data	2.94		-0.43	

The starred quantities come from restricting the analysis to data from the primary exchange. Note the massive reduction in noise on NYSE. The combined impacts of Archipelago and regulation NMS should change that!

Samples of 50 successive trades

- Note autocorrelation of order flow in real markets.
- No such correlation in the zero-intelligence market.
- However, plots look qualitatively similar.



Recommendations

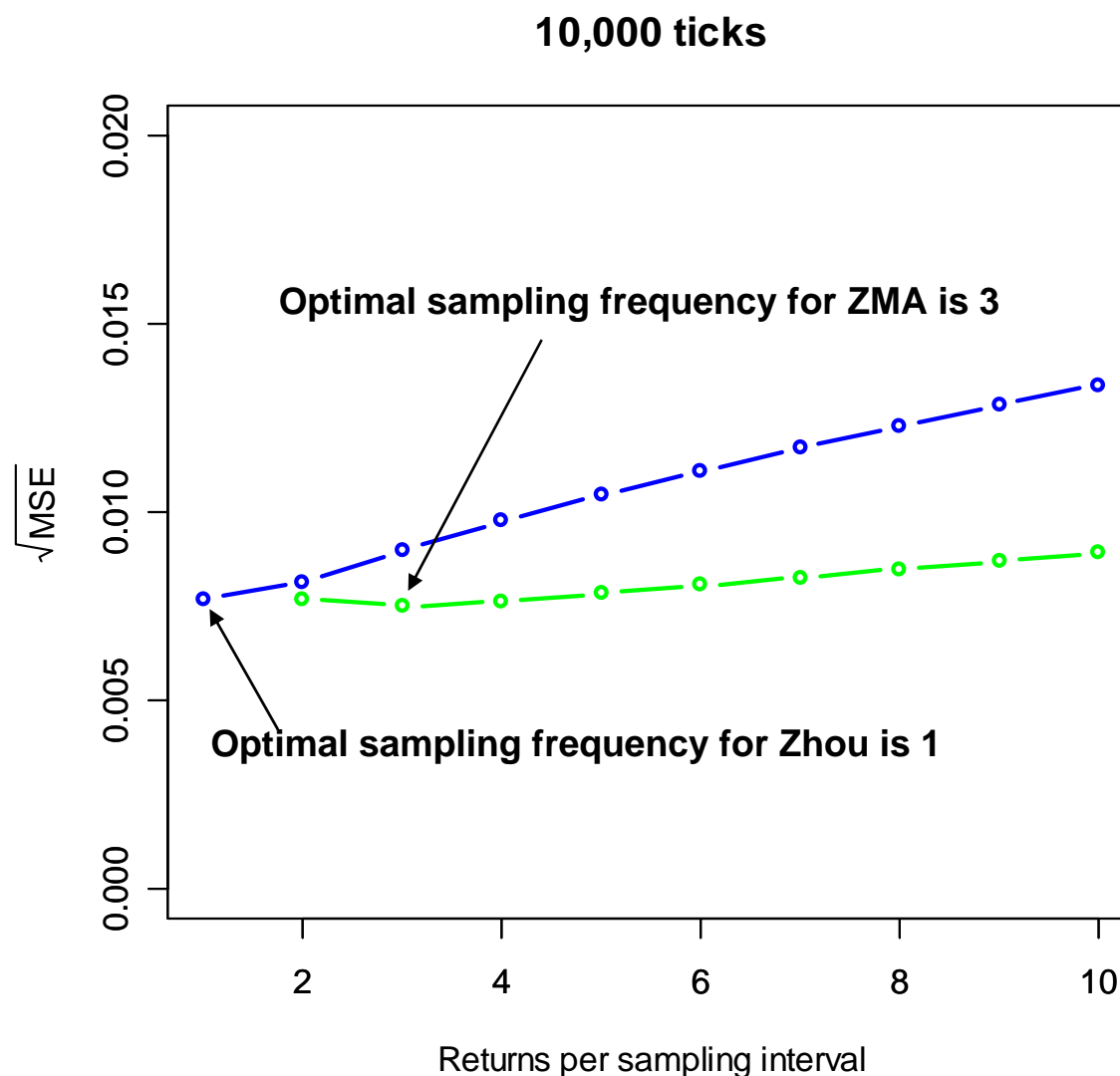
- Under zero intelligence,
 - when the sample consists of a large number of ticks, the Zhou estimator performs better and the optimal sampling frequency is 4.
 - when the sample consists of a small number of ticks (1,000 or less), the ZMA estimator performs best.
- However, closer inspection of the results of applying Zhou and ZMA to the real data suggest that better results are obtained from Zhou in all cases
 - See for example the smaller difference between Zhou and RV estimated at lower sampling frequency in all cases.
 - The probable implication is that the zero-intelligence model is underestimating microstructure bias in the estimators (though zero-intelligence still does better than assuming iid, serially uncorrelated noise).
 - Specifically, in the real world, order flow signs are highly autocorrelated; they are uncorrelated in the zero-intelligence world.
- We thus recommend using Zhou with a sampling frequency of 4 in all cases.
- Finally, if for practical reasons, the high-low-open estimator must be used, it must at least be discreteness adjusted.

What about mid-quotes?

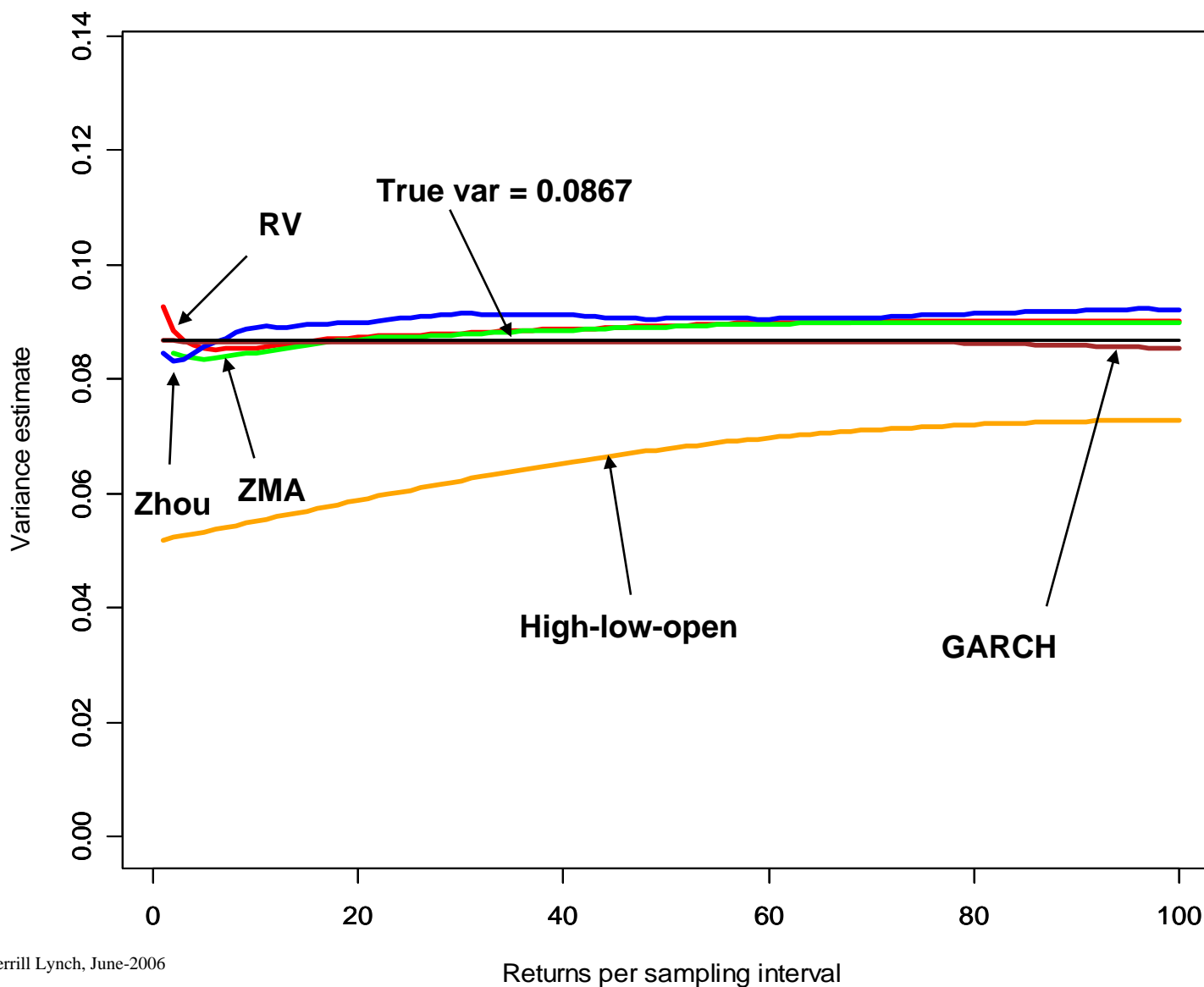
- The foregoing recommendations certainly apply if you only have trade data.
- What if you also have quote data?
- According to practitioners, using mid-quotes eliminates bid-ask bounce.
- If we use mid-quotes, do we get the same volatility estimate?
 - When should we sample the mid-quotes? Every quote change? Every second?
- According to Bouchaud, Gefen, Potters and Wyart (2004) and also Bandi and Russell (2006), it's best to sample the mid-quote just before each trade.
- We now repeat our previous artificial market experiment, logging mid-quotes rather than trade prices.

Mid-quote based MSE vs sampling frequency

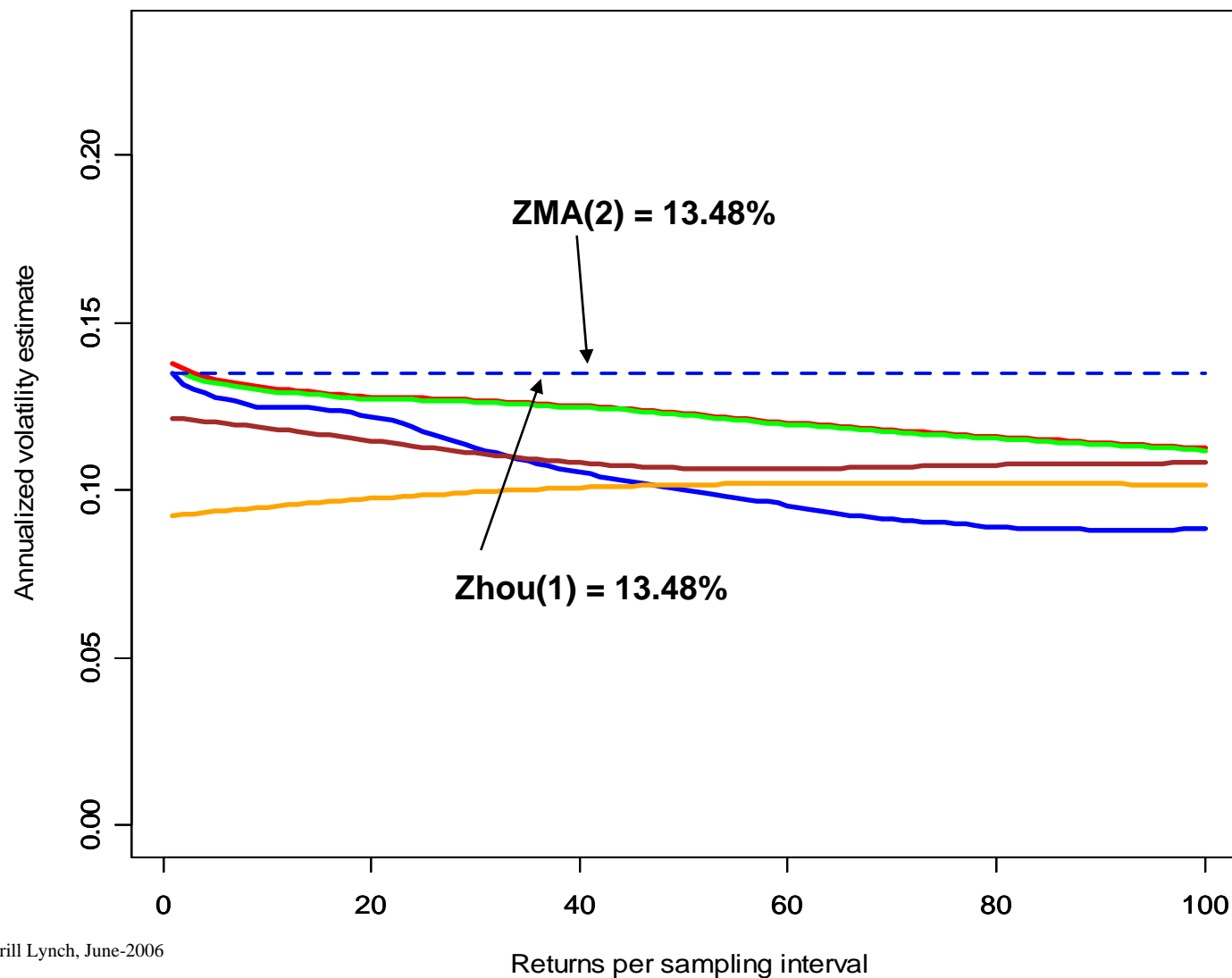
- For a sample size of 10,000 trades, minimal MSEs are almost identical for Zhou and ZMA.
- Optimal sampling frequency is *every tick* for Zhou.
- We note that minimal MSEs are similar to those from trade data.



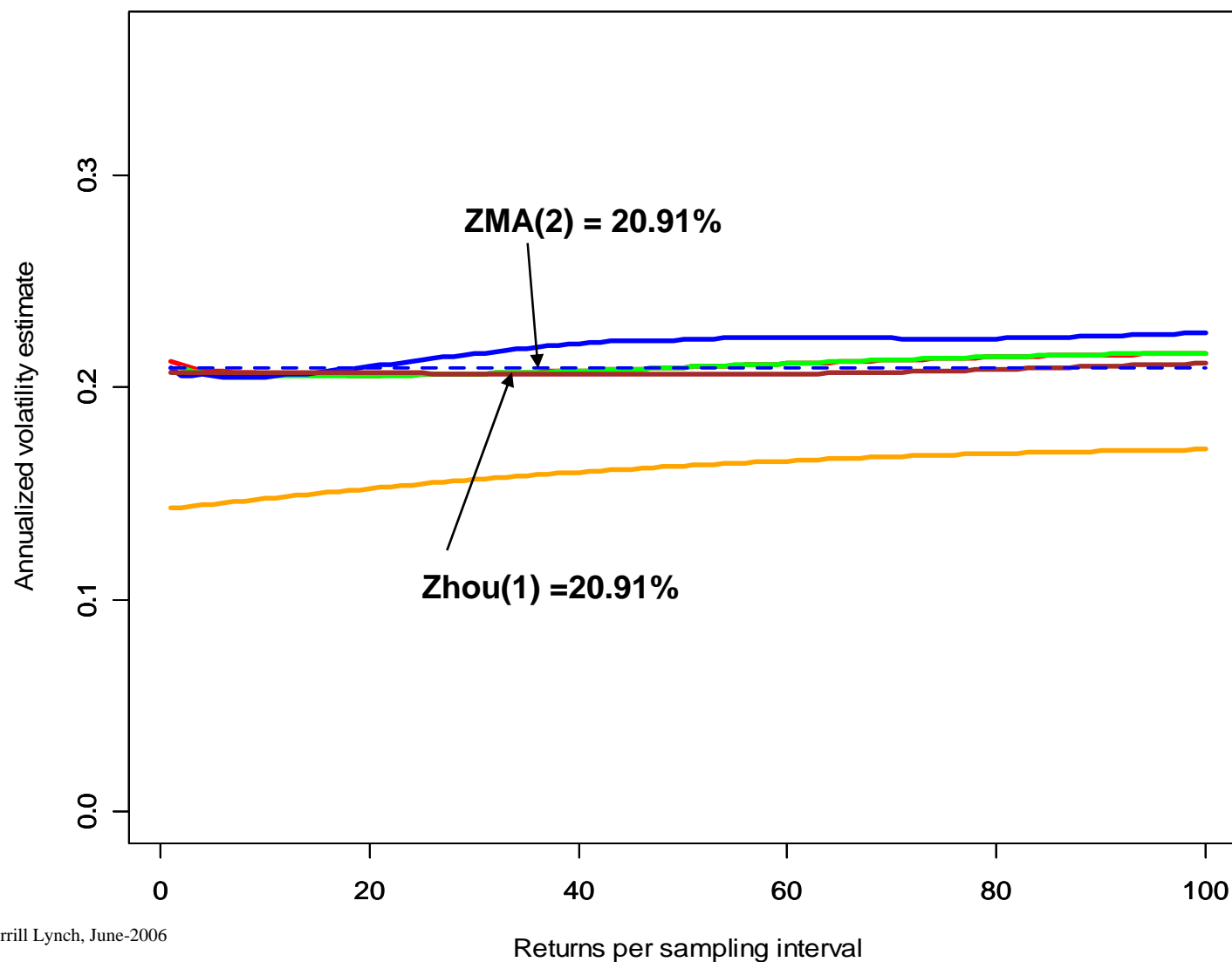
Performance of estimators on ZI mid-quote data



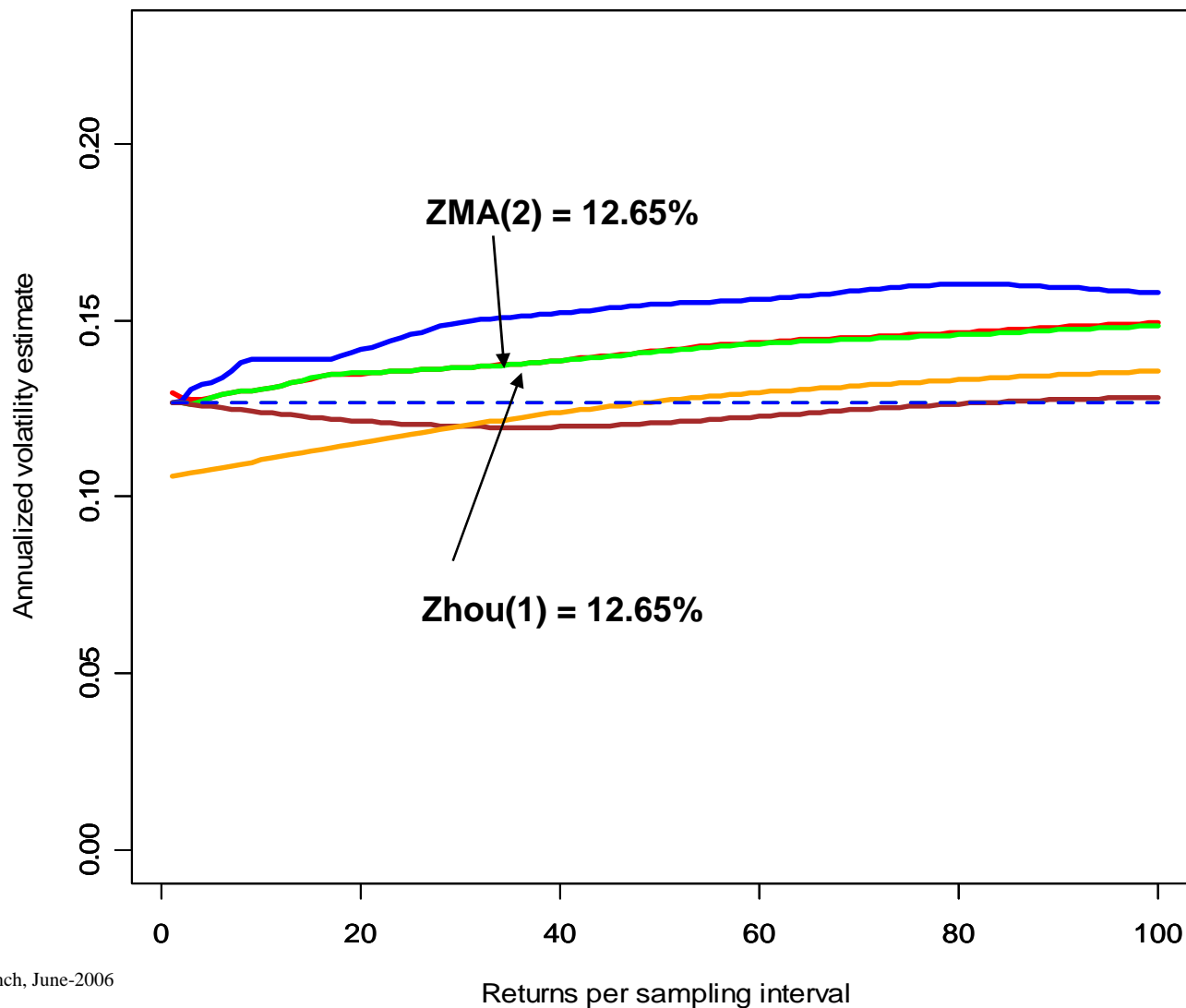
Mid-quote based estimates for PFE (06-Apr-2006)



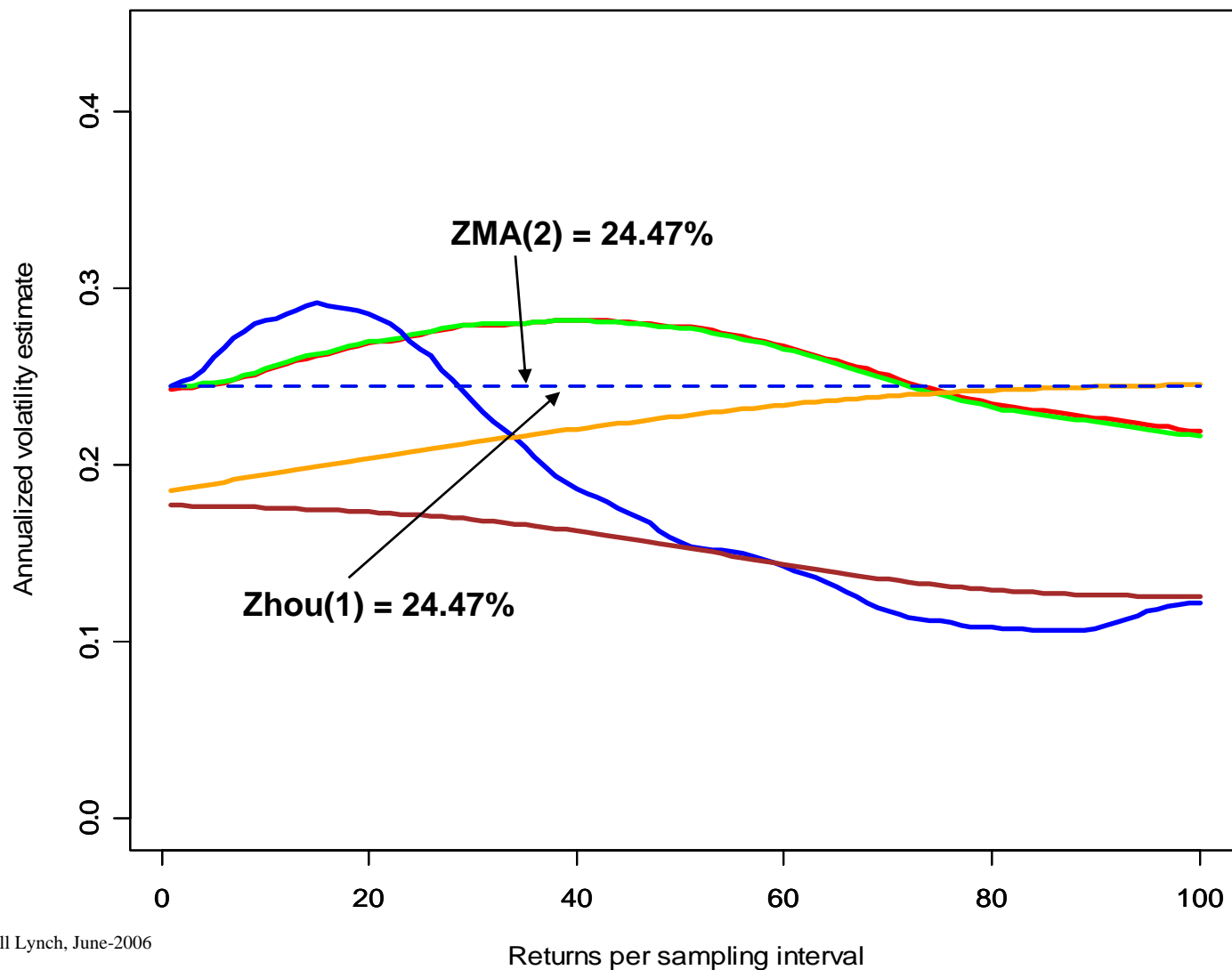
Mid-quote based estimates for CSCO (06-Apr-2006)



Mid-quote based estimates for IBM (06-Apr-2006)



Mid-quote based estimates for COCO (06-Apr-2006)



Summary of results

The following table compares estimated and implied volatilities (from Bloomberg) as of 06-Apr-2006 for the four stocks studied and includes mid-quote based estimates:

	σ_{ZMA}	σ_{Zhou}	$\sigma_{Implied}$	σ_{ZMA}^{Mid}	σ_{Zhou}^{Mid}
PFE	15.27%	13.18%	19.90%	13.48%	13.48%
CSCO	25.69%	22.36%	30.50%	20.91%	20.91%
IBM	12.52%	12.20%	16.41 %	12.65%	12.65%
COCO	27.69 %	25.43%	41.92%	24.47%	24.47%

The mid-quote based estimates are all consistent with our prior trade-based estimates and with our previous guess that the Zhou estimator is closer to the truth.

Noise-to-signal ratios (γ) and first order autocorrelations with added mid-quoted based results

	γ	γ^*	γ^{Mid}	ρ_1	ρ_1^*	ρ_1^{Mid}
PFE	4.17	1.73	0.02	-0.40	-0.40	-0.02
CSCO	2.63	2.22	0.01	-0.38	-0.34	-0.01
IBM	0.85	0.18	0.02	-0.29	-0.12	-0.02
COCO	0.30	0.35	-0.03	-0.13	-0.17	0.04
ZI data	2.94			-0.43		

We note that both noise and first order correlations in the mid-quotes are effectively zero.

Final conclusion

- We have shown that in the zero intelligence simulation, the various estimators of realized variance are also good estimators of the variance of the distribution of final returns.
- We also showed that the Zhou and ZMA give reasonable and consistent estimates of realized variance when applied to real data.
- We further showed that time series of real mid-quote changes exhibit effectively no noise or autocorrelation. The various estimators generate realized variance estimates consistent with those generated from time series of transaction price changes.

References

- Federico M. Bandi and Jeffery R. Russell (2006). Comment on JBES 2005 invited address by Peter R. Hansen and Asger Lunde. *Journal of Business & Economic Statistics* **24** (2), 167-173.
- T. Bollerslev, I. Domowitz and J. Wang (1994). Order-flow and the bid-ask spread: an empirical probability model of screen-based trading. *Journal of Economic Dynamics and Control* **21** (8-9), 1471-1491.
- Jean-Philippe Bouchaud, Yuval Gefen, Marc Potters, and Matthieu Wyart (2004) Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quantitative Finance* **4**, 176-190.
- J. Doyne Farmer, Paolo Patelli and Ilija Zovko (2005). The predictive power of zero-intelligence in financial markets. *Proceedings of the National Academy of Sciences of the USA* **102**(11), 2254-2259.
- Peter R. Hansen and Asger Lunde (2005). Realized variance and market microstructure noise. Available at SSRN: <http://ssrn.com/abstract=506542>.
- Roel C. A. Oomen (2005). Properties of bias-corrected realized variance under alternative sampling schemes. *Journal of Financial Econometrics* **3**(4), 555-577.

More references

- Roel C. A. Oomen (2006). Comment on JBES 2005 invited address by Peter R. Hansen and Asger Lunde. *Journal of Business & Economic Statistics* **24** (2), 195-202.
- Eric Smith, J Doyne Farmer, Laszlo Gillemot and Supriya Krishnamurthy (2003). Statistical theory of the continuous double auction. *Quantitative Finance* **3**, 481-514.
- Lan Zhang, Per A. Mykland and Yacine Aït-Sahalia (2005), A tale of two timescales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* **100**(472), 1394-1411.
- B. Zhou (1996), High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics* **14**(1), 45-52.