# Comments on Hansen and Lunde

**Eric Ghysels*** **Arthur Sinko†**

First Draft: September 5, 2005
This Draft: October 14, 2005

*Department of Finance, Kenan-Flagler School of Business and Department of Economics University of North Carolina, Gardner Hall CB 3305,Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, phone: (919) 966-5325, e-mail: eghysels@unc.edu.

†Department of Economics University of North Carolina, Gardner Hall CB 3305,Gardner Hall CB 3305, Chapel Hill, NC 27599-3305, e-mail: sinko@email.unc.edu.

Hansen and Lunde have written an impressive paper on estimation of volatility using high frequency financial data and the presence of microstructure noise. In these comments we will make two arguments: (1) as far as *predicting* future volatility is concerned it appears that corrections for microstructure noise do not matter very much, and (2) power variation *uncorrected* for microstructure noise still remains by far the best predictor. The comments will be structured in three sections, the first dealing with decomposition issues, the second with volatility forecasting and the third deals with volatility measures other than increments in quadratic variation.

# 1 Decompostions and Theory: A déjà vu issue?

It may sound very strange, but the topic of the Hansen and Lunde (2005) paper has similarities with the at first sight totally unrelated topic of seasonality in economic time series. The similarity is not because financial asset returns features intra-daily seasonal fluctuations. The similarity goes along the following lines: a seasonal time series, say $x_t$ is decomposed into a nonseasonal component $x_t^{NS}$ and a seasonal component $x_t^S$. Both components are unobserved and therefore can only be identified via a set of assumptions. First, it is assumed that the two components are orthogonal. Second, the seasonal component is identified via its temporal dependence structure featuring only so called seasonal autocorrelation. These assumptions can easily be criticized and indeed it has been argued by Ghysels (1988), among others, that 'economic theory' does not yield the decomposition used to seasonally adjust economic time series. Ghysels (1988) in fact showed that standard economic models do not yield orthogonal decompositions. Hansen and Lunde also explain that microstructure noise cannot be assumed to be independent from fundamental price shifts.

Despite the incompatibility of the identifying assumptions with theory one still overwhelmingly seasonally adjusts economic time series with filters based on orthogonal decomposition models. One of the interesting findings in Hansen and Lunde is that for sampling frequencies around 20 to 30 minutes the independence assumption seems to be a reasonable approximation.

The seasonality literature also takes us to the next logical step. The perennial question regarding seasonality is: Why do we seasonality adjust economic time series? Arguments are abundant on both sides, namely those who oppose it and those who agree (see e.g. Ghysels

(1996) and Ghysels and Osborn (2001) for further discussion and references). Do any of these arguments apply to adjustments for microstructure noise? Or, to put it differently: why do we adjust high frequency volatility measures for microstructure noise? The answer is obviously that we want to have a better measurement of the volatility that 'matters', which is identified by the component of discrete intra-daily squared returns that features long term autocorrelations (like the business cycle component 'matters' for macroeconomists). Implicitly, it is therefore argued that better measurement of realized volatility matters for something and that is most likely ... forecasting future volatility. At this point, the seasonality literature and microstructure noise issues diverge. Seasonal fluctuations dwarf business cycle movements in amplitude. The majority of economic time series are overwhelmingly seasonal. Hansen and Lunde look at the 30 Dow Jones stock. These are very liquid stocks and the magnitude of microstructure noise is rather small. One therefore wonders what difference microstructure noise adjustments make.

## 2 What difference does it make?

In this section we examine whether the corrections suggested by Hansen and Lunde improve, when compared to uncorrected realized volatilities, the prediction of future volatility. We are consider the following alternative volatility estimates: $RV^{5min}$, $RV^{30min}$, $RV^{30min}_{AC_1}$ and $RV^{1tick}_{ACNW_{30}}$. The former two are unadjusted, whereas the latter two are adjusted for microstructure noise. To assess forecasting performance, we follow the recent work of Ghysels, Santa-Clara, and Valkanov (2003) who use MIDAS regressions to predict realized volatility at weekly, two-weeks, three-weeks, and monthly horizons. In the context of forecasting the increments in quadratic variation, denoted $RV^x_y(t+H,t)$ for horizon $H$ with $x$ and $y$ taking the values above - for example $x = 30min$ and $y = AC_1$ for the Zhou corrected RV estimates. For the various measures we consider the following regressors:

$$RV^x_y(t+H,t) = \mu^Q_H + \phi^Q_H \sum_{k=0}^{k^{max}} b^Q_H(k,\theta) RV^x_y(t-k,t-k-1) + \varepsilon^Q_{Ht} \tag{1}$$

Hence, we compare how correcting for microstructure noise improves the forecasts of future corrected increments and consider $H$ equal to one week. Note that we consider uncorrected and corrected measures of quadratic variation on both sides of equation (1). The specifications of the polynomials are further discussed in Ghysels, Sinko, and Valkanov

(2003). They discuss several alternative parametric specifications for the polynomials and we use their so called beta polynomial particularly suitable for the application at hand. The $MSFT$ stock is used as an empirical example. Figure 1 displays the daily volatility dynamics using the $RV_{AC_1}^{5min}$ volatility measure for the sample considered by Hansen and Lunde. The time series plot clearly demonstrates that volatility dynamics of the first part of the sample is quite different from the dynamics of the second one. There is evidence of a structural change or regime switch, and this leads us to study not only the entire sample but also two subsamples, respectively three and two years long. The sample mean of the daily series for the first three years (first 753 observations) is 6.15 whereas for the last two years (last 501 observations) is 1.47.

Our analysis covers two sample sizes and two measures of stock returns. We start with the entire sample, i.e. from January 3, 2000 to December 31, 2004. Next, we consider the subsample from January 3, 2000 – December 31, 2002. The returns are computed using mid-quotes prices and trading prices. The results covering both definitions of returns and covering both samples appear in Table 1, where each row corresponds to the same left hand side variable discussed above but with different explanatory variables and sample sizes.

In this section we will discuss the explanatory power of the different $RV_y^x$ measures of the quadratic variation, i.e. first four columns for each sample in the table. The main finding is that there is no significant difference between unadjusted and adjusted predictors. Moreover, the unadjusted $RV^{5min}$ measure has the best explanatory power across all models and samples except for the whole sample where the model with $RV_{AC_1}^{5min}$ does marginally better (the difference only being 1.1%). The results are robust for different data samples and different return measures. Hence, form $MSFT$, we find that noise-corrected volatility measures performs on average worse than unadjusted five minutes volatility measure. One explanation is that the signal-to-noise ratio is high and, in terms of the MSE, the uncorrected measure is better than the corrected ones. Another explanation is that the MIDAS regression is extremely effective in extracting the signal from the daily realized volatility measures and the noise correction only reduces explanatory power of the regression.

# 3 Why quadratic variation?

Realized volatility is not the only measure of volatility. Ghysels, Santa-Clara, and Valkanov (2003) find that volatility can be forecasted using daily regressors other than squared returns. They show that better in- and out-of-sample estimates of the volatility are obtained when the predictors on the right-hand side are daily absolute returns, daily realized volatilities, daily ranges, and daily realized powers. The exact definitions of these predictors are provided below. The daily realized volatility, daily ranges, and daily realized powers are obtained from intra-daily ($5 - minute$) data of the Dow Jones index returns over the period from April 1993 to October 2003. Ghysels, Santa-Clara, and Valkanov (2003) show that the best overall predictor of conditional volatility is the realized power and that, not surprisingly, better forecasts are obtained at shorter (weekly) horizons. This empirical evidence was corroborated by theoretical and further empirical work by Forsberg and Ghysels (2004) who use the theoretical framework of Barndorff-Nielsen and Shephard (2001) involving a non-Gaussian Ornstein-Uhlenbeck (OU) volatility process with non-negative Lévy increments. Within this diffusion framework Forsberg and Ghysels show that realized power *is expected to be* the best predictor. To appraise how microstructure corrected quadratic variation compares to the measures considered by Ghysels, Santa-Clara, and Valkanov (2003) we examine:

$$RV_y^x(t+H,t) = \mu_H^p + \phi_H^p \sum_{k=0}^{k^{max}} b_H^P(k,\theta)\widetilde{PV}^{(m)}(t-k,t-k-1) + \varepsilon_{Ht}^p \tag{2}$$

where $\widetilde{PV}^{(m)}$ is the power variation or the sum of intra-daily absolute returns.

We use the same data as in the previous section. $PV$ is obtained using daily aggregation of five minutes returns. The results for the power variation appear in the $5^{th}$ and $10^{th}$ column of the Table 1. Consistent with Ghysels, Santa-Clara, and Valkanov (2003) and Forsberg and Ghysels (2004), $PV$ appears not only to be the best predictor compared to uncorrected predictors, but also $PV$ has the best explanatory power for all models except one. The $R^2$'s for PV are most often between 1% to 4% higher for one week ahead predictions. These results become even more dramatic at long horizons (not reported here). Hence, combining this with the results from the previous section, we can conclude that, on average, noise correction does not appear to be important for volatility prediction.

Correcting $PV$ for microstructure noise is a much harder problem than correcting $QV$. By analogy with $QV$, the correction may perhaps not improve much on the performance of forecasting.

To be fair, we have changed the objective function. Hansen and Lunde try to find the best filter that eliminates microstructure noise. Our objective function is to predict future volatility. Arguably, predicting future volatility is one of the ultimate objectives. Our findings reported here can be robustified. Namely, we have computed results for stocks other than MSFT and they yield very much the same type of results as those reported in Table 1. Nevertheless more work is needed to come to a full picture and this is done in ongoing work by Ghysels and Sinko (2005).

# 4 Acknowledgement

We are most grateful to Peter Hansen and Asger Lunde for providing us the data used in this comment.

# References


Forsberg, Lars, and Eric Ghysels, 2004, Why do absolute returns predict volatility so well?, UNC at Chapel Hill.

Ghysels, E., 1988, A study towards a dynamic theory of seasonality for economic time series, *Journal of the American Statistical Association* 85, 168–172 Reprinted in *Modelling Seasonality*, S. Hylleberg (ed.), Oxford University Press, 181-192.

———, 1996, On the Economics and Econometrics of Seasonality, in C. A. Sims, ed.: *Advances in Econometrics, Sixth World Congress* pp. 257–316. Cambridge University Press.

———, and E. Osborn, 2001, *The Econometric Analysis of Seasonal Time Series* (Cambridge University Press, Cambridge).

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov, 2003, Predicting volatility: getting the most out of return data sampled at different frequencies, *Journal of Econometrics* forthcoming.

Ghysels, Eric, and Arthur Sinko, 2005, Volatility prediction and microstructure noise, Work in progress.

———, and Rossen Valkanov, 2003, MIDAS Regressions: Further Results and New Directions, Working paper, UNC and UCLA.

Hansen, Peter R., and Asger Lunde, 2005, Realized variance and market microstructure noise, *Journal of Business and Economic Statistics.*

**Table 1: $R^2$ Comparison of MIDAS Models for One Week Horizon - MSFT Stock**

Each entry in the table corresponds to the $R^2$ for different models (1-2) and different estimation samples. The whole sample covers January 3, 2000 - December 31, 2004. Subsample 2000 - 2002 covers January 3, 2000 - December 31, 2002. The regressions are run on a weekly (5 days) data sampling scheme. The name of the variables are consistent with the notation in Hansen and Lunde paper. Every column corresponds to the explanatory power of the different $LHS$ variables for the same $RHS$ variable.

| | 2000 − 2004 | | | | | 2000 − 2002 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Model | $RV^{5min}$ | $RV^{30min}$ | $RV^{5min}_{AC_1}$ | $RV^{1tick}_{ACNW_{30}}$ | $PV$ | $RV^{5min}$ | $RV^{30min}$ | $RV^{5min}_{AC_1}$ | $RV^{1tick}_{ACNW_{30}}$ | $PV$ |
| Mid quotes | | | | | | | | | | |
| $RV^{5min}$ | 0.557 | 0.556 | 0.547 | 0.543 | 0.585 | 0.433 | 0.365 | 0.404 | 0.401 | 0.447 |
| $RV^{30min}$ | 0.603 | 0.599 | 0.593 | 0.595 | 0.618 | 0.404 | 0.345 | 0.378 | 0.370 | 0.408 |
| $RV^{5min}_{AC_1}$ | 0.552 | 0.556 | 0.563 | 0.535 | 0.556 | 0.412 | 0.352 | 0.410 | 0.371 | 0.407 |
| $RV^{1tick}_{ACNW_{30}}$ | 0.590 | 0.589 | 0.579 | 0.576 | 0.610 | 0.456 | 0.386 | 0.422 | 0.421 | 0.464 |
| Trades | | | | | | | | | | |
| $RV^{5min}$ | 0.570 | 0.569 | 0.557 | 0.554 | 0.604 | 0.447 | 0.373 | 0.408 | 0.399 | 0.466 |
| $RV^{30min}$ | 0.616 | 0.609 | 0.600 | 0.602 | 0.639 | 0.421 | 0.353 | 0.384 | 0.373 | 0.432 |
| $RV^{5min}_{AC_1}$ | 0.558 | 0.558 | 0.564 | 0.529 | 0.571 | 0.423 | 0.351 | 0.409 | 0.357 | 0.424 |
| $RV^{1tick}_{ACNW_{30}}$ | 0.596 | 0.592 | 0.573 | 0.589 | 0.634 | 0.452 | 0.374 | 0.401 | 0.418 | 0.476 |

**Figure 1: Daily $RV^{5min}_{AC_1}$ Realized Volatility**

The figure shows daily realized volatility with $AC_1$ noise-correction scheme. The $753^{rd}$ observation is 2002 end-of-year observation. Mean of the first three years is 6.18; mean of the second two years – 1.74.