

On Market Microstructure Noise and Realized Volatility¹

Francis X. Diebold²
University of Pennsylvania
and NBER

The Hansen-Lunde (HL) research program is generally first-rate, displaying a rare blend of theoretical prowess and applied sense. The present paper is no exception. In a major theoretical advance, HL allow for correlation between microstructure (MS) noise and latent price. (I prefer “latent price” to terms such as “efficient price” or “true price,” which carry lots of excess baggage.) In a parallel major substantive advance, HL provide a pioneering empirical investigation of the *nature* of the correlation between MS noise and latent price, documenting a negative correlation at high frequencies. My admiration of the paper hinges on the contributions noted above and is indeed most genuine. Nevertheless, much of what follows is rather critical of the extant literature, including certain key elements of the HL approach. My intention is that the criticism be constructive, promoting and hastening additional progress.

1. On the Dynamics of Latent Price

HL work in the framework:

$$p(t) = p^*(t) + v(t) \tag{1}$$

$$dp^*(t) = \mu dt + \sigma(t)dW(t), \tag{2}$$

where $p(t)$ is observed (log) price, $p^*(t)$ is the latent (log) price, $v(t)$ is MS noise, μ is a fixed expected return (actually HL go even farther and restrict $\mu=0$), and dW_t is an increment of standard Brownian motion. Without additional assumptions equation (1) is tautological, simply defining MS noise as $v(t) = p(t) - p^*(t)$. Hence everything hinges on the assumed specifications of $p^*(t)$ and $v(t)$ – neither of which is observable! – and assumptions regarding their interaction.

¹ This paper is an invited comment on P. Hansen and A. Lunde, “Realized Variance and Market Microstructure Noise,” *Journal of Business and Economic Statistics*, 2006, forthcoming. For useful discussion I thank Torben Andersen, Tim Bollerslev, Doug Steigerwald, Clara Vega, and especially Georg Strasser, who provided outstanding research assistance.

² Correspondence to: F.X. Diebold, Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104-6297; fdiebold@sas.upenn.edu.

Prior to HL, the literature effectively focused on specifications with latent price assumed uncorrelated with the MS noise,

$$\text{corr}(v(t), dW(t)) = 0. \quad (3)$$

HL progress by allowing instead,

$$\text{corr}(v(t), dW(t)) = \rho. \quad (4)$$

Importantly, their allowance for $\text{corr}(v(t), dW(t)) \neq 0$ accords with both MS theory (more on this later) and with empirical fact (as HL emphasize).

However, the HL specification (1), (2), (4) remains quite limited relative to one allowing for *time-varying expected returns* and *jumps*, as in:

$$dp^*(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dq(t), \quad (5)$$

where $\mu(t)$ is the time-varying expected return, $\kappa(t) = p(t) - p(t-)$ is jump size, and $q(t)$ is a counting process with possibly time-varying intensity $\lambda(t)$ such that $P[dq(t)=1] = \lambda(t)dt$.

First consider the possibility of time-varying expected returns. Note that $p^*(t)$ is a real-world price, not a risk-neutral price, so there is no reason for $p^*(t)$ to follow a martingale. Hence allowance for time-varying expected returns is important in principle. In practice one might argue that, at least in high-frequency environments (e.g., hourly returns used to construct daily realized volatility (RV)), time-variation in expected returns is likely to be negligible and can therefore be safely ignored. Fair enough, but at least three caveats are in order. First, and obviously, interest sometimes centers *not* on high-frequencies, but rather on lower frequencies such as annual RV constructed from underlying monthly returns, particularly in historical asset market studies covering long calendar spans. Second, there is evidence of non-martingale behavior not only at long horizons, but also at short horizons (e.g., Lo and MacKinlay, 2001). Hence even if interest *does* center on high-frequency returns, it is not obvious that time-varying expected returns can be safely ignored. Finally, Elliott's (1998) well-known work establishes that cointegration methods are not robust even to *slight* deviations from I(1) behavior in the underlying variables. In the present case (and moving to discrete time to match the standard cointegration framework), the required I(1) behavior is for p_t^* and hence p_t , which is apparently not guaranteed when time-variation in expected returns is allowed. This concern is particularly relevant to the present paper, a large part of which is devoted to cointegration analysis.

Now consider jumps. Jumps are an important feature of empirical reality, and frameworks that ignore them do so at their own peril. This insight arises repeatedly in many studies presenting estimates of parametric jump-diffusion models for asset returns, and it is reinforced and amplified by recent

nonparametric volatility analyses (see, e.g., Andersen, Bollerslev and Diebold, 2005, and the references therein) and examinations of the financial market reaction to macroeconomic news (e.g., Andersen, Bollerslev, Diebold and Vega, 2003). Indeed, given that incorporation of MS noise and jumps are widely acknowledged as two of the most pressing items on the RV research agenda, it is unfortunate that although the RV-jumps literature acknowledges MS noise (e.g., Barndorff-Nielsen and Shephard, 2004, 2005a, 2005b; Andersen, Bollerslev and Diebold, 2005), the latest significant advance of the RV-MS literature (namely HL) does not acknowledge jumps. I look forward to additional HL progress rectifying that situation.

2. On the Inadequacy of Linear/Gaussian Methods

For simplicity of exposition, assume that both the observed and latent prices evolve at transaction times, and that transaction times are equally spaced.

The HL framework corresponds to a linear discrete-time state space system:

$$p_t = p_t^* + \varepsilon_t \quad (6)$$

$$p_t^* = p_{t-1}^* + \sigma_t u_t \quad (7)$$

$$\begin{pmatrix} \varepsilon_t \\ u_t \end{pmatrix} \stackrel{iid}{\sim} N \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon u} \\ \sigma_{\varepsilon u} & 1 \end{pmatrix}. \quad (8)$$

The system may or may not be conditionally Gaussian, depending on the dynamics of σ_t . If, for example, σ_t has GARCH structure such as

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2, \quad (9)$$

then the system is conditionally Gaussian, but if σ_t has stochastic volatility structure such as

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \eta_t \quad (10)$$

$$\eta_t \stackrel{iid}{\sim} (0, \sigma_\eta^2),$$

then the system is not conditionally Gaussian. Of course the realized volatility framework makes only minimal assumptions on σ_t and is compatible with both GARCH and stochastic volatility, among many others.

Quite apart from such details, however, the intrinsic mechanics of the leading MS noise candidates (and the ones explicitly used by HL to motivate their approach), namely bid/ask bounce and discrete price quotes, induce fundamental violations of the linear/Gaussian state space framework that appear to have been ignored in the RV-MS literature thus far. To see this, dispense for the moment with volatility dynamics, because including them only complicates matters without changing the basic point.

Consider in particular a MS model in the tradition of Hasbrouck (1999a,b), incorporating both bid/ask bounce and discrete pricing:

$$p_t = \begin{cases} \text{floor}(q_t^b, \text{ticksize}) & \text{if } \pi_t = 0 \\ \text{ceiling}(q_t^a, \text{ticksize}) & \text{if } \pi_t = 1 \end{cases} \quad (11)$$

$$\begin{aligned} q_t^b &= p_t^* - c_t \\ q_t^a &= p_t^* + c_t \end{aligned} \quad (12)$$

$$p_t^* = p_{t-1}^* + u_t \quad (13)$$

$$\pi_t \stackrel{iid}{\sim} \text{Bernoulli}(0.5) \quad (14)$$

$$c_t \stackrel{iid}{\sim} N(\mu_c, \sigma_c^2) \quad (15)$$

$$u_t \stackrel{iid}{\sim} N(0, \sigma_u^2), \quad (16)$$

where q_t^b is the bid price, q_t^a is the ask price, $2c_t$ is the bid/ask spread that represents the positive stochastic cost of dealer quote exposure, π_t is a bid/ask indicator variable, $\text{floor}(\bullet, d)$ rounds its argument down to the closest multiple of ticksize and $\text{ceiling}(\bullet, d)$ rounds up, and u_t , c_t and π_t are contemporaneously and serially independent.

The MS model (11)-(16) also constitutes a state space system, relating the observed p_t to the latent p_t^* , but the bid/ask bounce and rounding render it intrinsically *non-Gaussian*. Against this background, I

worry about the adequacy of the linear/Gaussian tools on which HL rely: HAC estimation based on sample autocorrelations, Gaussian cointegrated VARs, etc.

In closing this section, let me also note that the state space framework raises the possibility of direct filtering or smoothing of MS noise from observed returns, *before* proceeding to compute RV. This is of independent interest, as one might want to use MS-corrected returns for a variety of purposes beyond construction of MS-corrected RV. For the reasons discussed above, however, the optimal filter will generally be non-linear, and hence the Kalman filter will be suboptimal. Diebold and Vega (2002) explore optimal nonlinear MS noise filtering, and Owens and Steigerwald (2005) independently explore Kalman MS noise filtering. An interesting and open issue concerns the goodness of approximation of the Kalman filter to the fully optimal filter.

Alternatively, if interest centers exclusively on volatility, one can first construct realized volatility and then filter to reduce MS noise (as well as estimation error due to incomplete convergence of RV to underlying integrated volatility). This is the approach implicitly taken by HL, although they do not attempt to construct an optimal nonlinear filter tailored to the structure of the MS noise. It is also the approach taken explicitly in Andersen, Bollerslev, Diebold and Wu (2005, 2006), who analyze explicit state-space systems but use only a Kalman filter.

3. On the Theoretical and Empirical Correlation Between MS Noise and Latent Price

In a pioneering substantive contribution, HL discover a negative relationship between latent returns and MS noise at high frequencies. Does the negative correlation match the predictions of MS theory? HL claim rather casually that the answer is yes, arguing that “the correlation between noise and efficient price arises naturally in some models of market microstructure effects, including (a generalized version of) the bid-ask model by Roll (1984) ... and models where agents have asymmetric information, such as those by Glosten and Milgrom (1985) and Easley and O’Hara (1987, 1992).”

In fact the situation is more nuanced, with some MS considerations suggesting a positive correlation, and others suggesting negative, so that on net the correlation could be empirically positive, negative, or even zero. The issues are well-illustrated by the generalized Roll model emphasized by HL (see Roll, 1984, and Hasbrouck, 199a, 1999b, 2004), which can produce *either* positive or negative correlation, as follows.

Let us first generate a positive correlation. Focusing for simplicity only on bid/ask bounce (that is, ignoring the discreteness of price quotes), rewrite the system (11)-(16) in condensed notation as

$$p_t = p_t^* + c q_t \quad (17)$$

$$p_t^* = p_{t-1}^* + u_t,$$

where q_t indicates direction of trade (1 for a buy, and -1 for a sell). Also refine (13) by decomposing the latent price increment as

$$u_t = \lambda q_t + \omega_t, \quad (18)$$

where λ is the reaction of latent price to a trade and hence captures some aspects of private information, and ω_t denotes public information. Simple calculations then reveal that

$$\text{cov}(\Delta p_t^*, \Delta(p_t - p_t^*)) = E((\lambda q_t + \omega_t)(c(q_t - q_{t-1}))) = c\lambda E(q_t^2) = c\lambda > 0, \quad (19)$$

if q_t is independent of ω_t and independent over time, as is commonly assumed.

Now let us generate a potentially negative correlation by allowing for sluggish adjustment of transactions prices, which could arise for a variety of reasons, such as learning. We replace (17) with

$$p_t = p_{t-1}^* + c q_t, \quad (20)$$

yielding

$$\text{cov}(\Delta p_t^*, \Delta(p_t - p_t^*)) = (c - \lambda)\lambda - E(\omega_t^2), \quad (21)$$

which can be negative if c is small relative to λ and the variance of ω_t is large. It is interesting to note that MS models in the tradition of Glosten and Milgrom (1985), Kyle (1985), and Easley and O'Hara (1987, 1992) all assume similar adjustment lags and hence predict negative correlation.

The upshot: it is trivial to write down a microstructure model with either a negative or a positive contemporaneous correlation between latent returns and MS noise. The basic insight is that certain factors promote *over*reaction of observed price to movements in latent price and hence a positive correlation, while other factors promote *under*reaction of observed price to movements in latent price and hence a negative correlation. Different models (or different variants of the same model) emphasize different factors.

Empirically, moreover, several such factors may be operative simultaneously, and whether they aggregate to something negative or positive is an empirical matter. Hence the negative estimates obtained by HL do not necessarily “support” or “refute” any particular MS model.

In any event, a rich vein remains to be mined. The discussion in HL (and thus far in this comment) focuses only on the contemporaneous correlation between latent price and MS noise. The contemporaneous correlation, however, is just the tip of the iceberg. The entire cross-correlation structure associated with the transmission of latent price movements into eventual observed price movements is of great interest, and

little is known about it, whether theoretically or empirically. Diebold and Strasser (2005) provide an initial exploration.

References

- Andersen, T.G., Bollerslev, T. and Diebold, F.X. (2005), "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," Manuscript, Northwestern University, Duke University and University of Pennsylvania. Available at http://www.ssc.upenn.edu/~fdiebold/papers/paper56/abd_081605.pdf
- Andersen, T., Bollerslev, T., Diebold, F.X. and Vega, C. (2003), "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review*, 93, 38-62.
- Andersen, T.G., Bollerslev, T., Diebold, F.X. and Wu, J. (2006), "Realized Beta: Persistence and Predictability," in T. Fomby (ed.) *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger*, in press.
- Andersen, T.G., Bollerslev, T., Diebold, F.X. and Wu, J. (2005), "A Framework for Exploring the Macroeconomic Determinants of Systematic Risk," *American Economic Review*, 95, 398-404.
- Barndorff-Nielsen, O.E. and Shephard, N. (2004), "Power and Bipower Variation with Stochastic Volatility and Jumps" (with discussion), *Journal of Financial Econometrics*, 2, 1-48.
- Barndorff-Nielsen, O. E. And Shephard, N. (2005a), "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, forthcoming.
- Barndorff-Nielsen, O.E. and Shephard, N. (2005b), "Impact of Jumps on Returns and Realised Variances: Econometric Analysis of Time-deformed Levy Processes," *Journal of Econometrics*, forthcoming.
- Diebold, F.X. and Strasser, G. (2005), "On the Correlation Structure of Microstructure Noise in Theory and Practice," Manuscript in progress, University of Pennsylvania.
- Diebold, F.X. and Vega, C. (2002), "Filtering Microstructure Noise," Manuscript, University of Pennsylvania.
- Easley, D. and O'Hara, M. (1987), "Price, Trade Size, and Information in Securities Markets." *Journal of Financial Economics*, 19, 69-90.
- Easley, D. and O'Hara, M. (1992), "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 576-605.
- Elliott, G. (1998), "On the Robustness of Cointegration Methods when Regressors Almost Have Unit Roots," *Econometrica*, 66, 149-158.
- Glosten, L.R., and Milgrom, P.R. (1985), "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71-100.
- Hasbrouck, J. (1999a), "Security Bid/Ask Dynamics with Discreteness and Clustering: Simple Strategies for Modeling and Estimation," *Journal of Financial Markets*, 2, 1-28.
- Hasbrouck, J. (1999b), "The Dynamics of Discrete Bid and Ask Quotes," *Journal of Finance*, 6, 2109-

Hasbrouck, J. (2004), "Empirical Market Microstructure: Economic and Statistical Perspectives on the Dynamics of Trade in Securities Markets," Manuscript, Stern School of Business, New York University.

Kyle, A. (1985), "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315-1335.

Lo, A.W. and MacKinlay, A.C. (2001), *A Non-Random Walk Down Wall Street*. Princeton: Princeton University Press.

Owens, J.P. and Steigerwald, D.G. (2005), "Noise Reduced Realized Volatility: A Kalman Filter Approach," in T. Fomby and D. Terrell (eds.), *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series*, Volume 20, Part A, forthcoming.

Roll, R. (1984), "A Simple Implicit Measure of the Effective Bid-ask Spread in an Efficient Market," *Journal of Finance*, 39, 1127-1139.