Drift in Transaction-Level Asset Price Models

Wen Cao* Clifford Hurvich† Philippe Soulier‡

November 9, 2012

Abstract

We study the effect of drift in pure-jump transaction-level models for asset prices in continuous time, driven by point processes. The drift is assumed to arise from a nonzero mean in the efficient shock series. It follows that the drift is proportional to the driving point process itself, i.e. the cumulative number of transactions. This link reveals a mechanism by which properties of intertrade durations (such as heavy tails and long memory) can have a strong impact on properties of average returns, thereby potentially making it extremely difficult to determine growth rates. We focus on a basic univariate model for log price, coupled with general assumptions on durations that are satisfied by several existing flexible models, allowing for both long memory and heavy tails in durations. Under our pure-jump model, we obtain the limiting distribution for the suitably normalized log price. This limiting distribution need not be Gaussian, and may have either finite variance or infinite variance. We show that the drift can affect not only the limiting distribution for the normalized log price, but also the rate in the corresponding normalization. Therefore, the drift (or equivalently, the properties of durations) affects the rate of convergence of estimators of the growth rate, and can invalidate standard hypothesis tests for that growth rate. Our analysis also sheds some new light on two longstanding debates as to whether stock returns have long memory or infinite variance.

*Stern School of Business, New York University, Henry Kaufman Management Center, 44 West Fourth Street, New York, NY 10012, USA, email: `wcao@stern.nyu.edu`

†Stern School of Business, New York University, Henry Kaufman Management Center, 44 West Fourth Street, New York, NY 10012, USA, email: `churvich@stern.nyu.edu`

‡Université Paris X, 200 avenue de la République, 92001 Nanterre cedex, France, email: `philippe.soulier@u-paris10.fr`. The research of Philippe Soulier was partially supported by the grant ANR-08-BLAN-0314-02.

The authors thank participants of the 2012 NBER-NSF Time Series Conference for helpful comments.

1 Introduction

In recent years, transaction-level data on financial markets has become increasingly available, and is now often used to make trading decisions in real time. Such data typically consist of the times at which transactions occurred, together with the price at which the transaction was executed, and may include other concomitant variables (”marks”) such as the number of shares traded. Our focus here is on actual transactions rather than quotes, but regardless of which type of event is being considered it is important to recognize that a useful framework for modeling and analyzing such data is that of marked point processes rather than, say, time series in discrete time. Though time series are typically provided for further analysis, such as daily (or high frequency) stock returns, these inevitably involve aggregation and entail a loss of information that may be crucial for trading and perhaps even for risk management and portfolio selection.

The perspective of asset prices as (marked) point processes has a long history in the financial and econometric literature. For example, Scholes and Williams [1977] allowed for a compound Poisson process. However, such a model is at odds with the stylized fact that time series of financial returns exhibit persistence in volatility. Recent interest in the point process approach to modeling transaction-level data was spurred by the seminal paper of Engle and Russell [1998], who proposed a model for inter-trade durations. Other work on modeling transaction-level data as point processes and/or constructing duration models includes that of Bowsher [2007], Bauwens and Veredas [2004], Hautsch [2012], Bacry et al. [2011], Deo et al. [2009], Deo et al. [2010], Hurvich and Wang [2010], Aue et al. [2011], Shenai [2012], Chen et al. [2012].

Nevertheless, it must be recognized that time series of asset returns in discrete (say, equally-spaced) time are still in widespread use, and indeed may be the only recorded form of the data that encompasses many decades. Such long historical series are of importance for understanding long-term trends (a prime focus of this paper) and, arguably, for a realistic assessment of risk. So given the ubiquitous nature of the time series data but also keeping in mind the underlying price-generating process that occurred at the level of individual transactions, it is important to make sure that transaction-level models obey the stylized facts, not only for the intertrade durations but also for the lower-frequency time series.

It has been observed empirically that time series of financial returns are weakly autocorrelated (though perhaps not completely uncorrelated), while squared returns or other proxies for volatility show strong autocorrelations that decay very slowly with increasing lag, possibly suggesting long memory (see Andersen et al. [2001]). It is also generally accepted that such time series show asymmetries, such as a correlation between the current return and the next period’s squared return, and this effect (often referred to traditionally as the ”leverage effect”) is addressed for example by the EGARCH model of Nelson [1991]. The average return often differs significantly from zero based on a traditional t-test, possibly suggesting a linear trend in the series of log prices. Meanwhile, Deo et al. [2010] found that intertrade durations have long memory (this was also found by Chen et al. [2012]), and they investigated the possibility that the durations have heavy tails.

One more fact that we wish to stress is that in continuous time, realizations of series of transaction-based asset prices are step functions, since the price is constant unless a transaction occurs. Thus, we choose to focus on pure-jump models for the log price (viewed as a time series in continuous time), driven by a point process that counts the cumulative number of transactions. This is equivalent to a marked point process approach where the points correspond to the transaction times and the marks are the transaction-level return shocks, which can have both an efficient component and a microstructure component.

Within this context, we will in this paper investigate the effect of drift (modeled at the transaction level) on the behavior of very-long-horizon returns, or equivalently, on the asymptotic behavior of the log price as time increases. The drift is assumed to arise from a nonzero mean in the efficient shock series. It follows that the drift is proportional to the driving point process itself, i.e. the cumulative number of transactions. This link reveals a mechanism by which properties of intertrade durations (such as heavy tails and long memory) can have a strong impact on properties of average returns, thereby potentially making it extremely difficult to determine long-term growth rates or to reliably detect an equity premium.

We focus on a basic univariate model for log price, coupled with general assumptions on durations that are satisfied by several existing flexible models, allowing for both long memory and heavy tails in durations. Under our pure-jump model (which can capture all the stylized facts described above), we obtain the limiting distribution for the suitably normalized log price. This limiting distribution need not be Gaussian, and may have either finite variance or infinite variance. The diversity of limiting distributions here may be considered surprising, since our assumptions imply that the return shocks obey an ordinary central limit theorem under aggregation across transactions (i.e. in *transaction time* but generally not in *calendar time*). We show that the drift can affect not only the limiting distribution for the normalized log price, but also the rate in the corresponding normalization. Therefore, the drift (or equivalently, the properties of durations) affects the rate of convergence of estimators of the growth rate, and can invalidate standard hypothesis tests for that growth rate. Our analysis also sheds some new light on two longstanding debates as to whether stock returns have long memory or infinite variance.

The remainder of this paper is organized as follows. In Section 2, we provide a simple univariate model for the log price, discuss the trend term and state our first main theorem on the limiting behavior of the log price process, as determined by the properties of intertrade durations. In Section 3 we study statistical inference for the trend, and obtain the behavior of the ordinary t-statistic under the null hypothesis. We then examine a series of examples based on specific duration models that have been proposed in the literature, including the ACD model of Engle and Russell [1998] and a generalized form of the LMSD model originally proposed in Deo et al. [2010]. These examples provide for great diversity of the asymptotic distributions of sums of durations and therefore (by our Theorem 2.1

below) for the asymptotic distribution of the log price. Section 4 provides a concluding discussion on how our results may help to reconcile some longstanding debates. Proofs of the mathematical results are provided in Section 5.

2 A Simple Univariate Model for Log Price

We start with a basic univariate pure-jump model for a log price series $y(t)$. Let $N(t)$ be a point process on the real line. Points of N correspond to transactions on the asset [1] and $N(t)$ is the number of transactions in $(0, t]$. We set $y(0) = 0$ and define, for $t \geq 0$,

$$y(t) = \sum_{k=1}^{N(t)} \tilde{e}_k \tag{2.1}$$

where $\tilde{e}_k = \mu + e_k$ and the $\{e_k\}$, which are independent of $N(\cdot)$, are i.i.d. with zero mean and finite variance σ_e^2.

We assume that μ is a nonzero constant. This model with $\mu = 0$ was considered in Deo et al. [2009], who showed that it can produce long memory in the realized volatility. We therefore have from (2.1),

$$y(t) = \mu N(t) + \sum_{k=1}^{N(t)} e_k \ . \tag{2.2}$$

Since we are modeling the log prices $y(t)$ as a pure-jump process the log price is constant when no trading occurs. Unfortunately, the modified version of (2.2) in which $\mu N(t)$ is replaced by the deterministic time trend ct (where c is a nonzero constant) would not yield a pure-jump process. Nevertheless, it is quite reasonable from an economic viewpoint to imagine that $E[y(t)]$ is a linear function of t, to account for such phenomena as equity premia and inflation. This is indeed the case for Model (2.1) if it is assumed that the point process is stationary with intensity $\lambda > 0$, which implies $E[y(t)] = E[\mu N(t)] = \mu\lambda t$. But in actual realizations of $y(t)$, the trend is only impounded when a transaction occurs, due to the nonzero mean in $\{\tilde{e}_k\}$.

Denote the transaction times by t_k with $\cdots t_{-1} \leq t_0 \leq 0 < t_1 \leq t_2 \cdots$ and define the durations by $\tau_k = t_k - t_{k-1}$. As we will see the properties of the point process N can play an important role in determining the (asymptotic) properties of statistics of interest. Two distinct modeling approaches seem natural. One is to model the point process directly as in Bacry et al. [2011] who use Hawkes processes. Another approach

[1] From a modeling perspective it may be desirable to instead have the points of N correspond to other relevant trading events, such as "every fourth transaction", "a transaction that moves the price", etc. For simplicity and definiteness in the paper, we simply let $N(t)$ count actual transactions, but our theoretical results do not depend on this particular choice of the definition of an event.

pioneered by Engle and Russell [1998] consists of modeling the durations as a stationary process. They defined the ACD model and Deo et al. [2010] defined the LMSD model. We will consider these duration models as examples in this paper.

Our first theorem demonstrates how the asymptotic distribution of suitably normalized sums of durations determines that of the correspondingly normalized log price, under the model (2.2). The theorem is a consequence of the CLT equivalence of Whitt [2002, Theorem 7.3.1]. Proofs of all theorems and lemmas are provided in Section 5. The symbol $\to$ below denotes convergence in distribution, and $\xrightarrow{\mathbb{P}}$ denotes convergence in probability.

Theorem 2.1. *Assume that for $\gamma \geq 1/2$,*

$$n^{-\gamma} \sum_{k=1}^{n} (\tau_k - 1/\lambda) \to A \tag{2.3}$$

where A is some nonzero random variable. If $\gamma > 1/2$ then $n^{-\gamma}(y(n) - \lambda\mu n) \to -\mu\lambda^{1+\gamma} A$. If $\gamma = 1/2$, then $n^{-1/2}(y(n) - \lambda\mu n) \to -\mu\lambda^{3/2} A + \sqrt{\lambda}\sigma_e Z$ where Z is a standard Gaussian random variable, independent of A.

We will later make specific assumptions on the durations or on the point process which imply the assumption of the theorem and can yield a wide variety of limiting distributions A and rates of convergence γ. In particular, A may be normal, non-Gaussian with finite variance, or stable with infinite variance. The case where $\gamma > 1/2$ and the limiting distribution has finite variance may be indicative of long memory in returns. This possibility was discussed by Lo [1991]. But in our model, if μ is not very large, any long memory phenomena generated by the stochastic drift $\mu N(t)$ may be difficult to detect in a data analysis. We will further develop these remarks later.

3 Statistical inference for the trend

For integer j, (assuming a time-spacing of 1 without loss of generality) we define the calendar-time returns as $r_j = y(j) - y(j-1)$ and the average return over a time period of n as $\bar{r}_n = n^{-1} y(n) = n^{-1} \sum_{j=1}^{n} r_j$. Theorem 2.1 implies that, in general, $\bar{r}_n - E[\bar{r}_n]$ will not be $O_p(n^{-1/2})$, making it difficult to accurately estimate growth rates.

Since our model (under stationarity of N) implies that $E[y(t)] = \lambda\mu t$, the growth rate per unit time is $\mu^* = \lambda\mu$. We therefore consider the problem of statistical inference for μ^*. We focus on testing a null hypothesis of form $H_0 : \mu^* = \mu_0^*$ based on $\bar{r}_n$, which is unbiased for μ^*. The corresponding t-statistic for testing H_0 is

$$t_n = n^{1/2}(\bar{r}_n - \mu_0^*)/s_n,$$

where

$$s_n^2 = (n-1)^{-1} \sum_{j=1}^{n} (r_j - \bar{r}_n)^2.$$

We next establish that s_n^2 consistently estimates a positive constant, under suitable regularity assumptions.

Lemma 3.1. *Under the assumptions of Theorem 2.1 if N is stationary and ergodic and $\mathbb{E}[N^2(1)] < \infty$, then $s_n^2 \xrightarrow{\mathbb{P}} \mu^2 \mathrm{var}(N(1)) + \lambda\sigma_e^2$.*

If $\gamma > 1/2$ it follows from Theorem 2.1 and Lemma 3.1 that if the null hypothesis is true and $\mu \neq 0$ then, with $\sigma^2 = \mu^2 \mathrm{var}(N(1)) + \lambda\sigma_e^2$,

$$n^{1-\gamma} n^{-1/2} t_n \rightarrow -\mu\lambda^{1+\gamma} A/\sigma \ . \tag{3.1}$$

Thus, $t_n = O_p(n^{\gamma-1/2})$ and the t-statistic diverges under the null hypothesis if $\gamma > 1/2$. Examples where this scenario would occur include durations generated by an ACD model with infinite variance, or by an LMSD model with long memory and an exponential volatility function. (See Examples 3.1 and 3.2 below). This scenario therefore is consistent with the empirical properties of durations found in Deo et al. [2010].

If $\gamma = 1/2$, then it follows similarly that

$$t_n \rightarrow (-\mu\lambda^{3/2} A + \sqrt{\lambda}\sigma_e Z)/\sigma \ . \tag{3.2}$$

So in the case $\gamma = 1/2$ the t-test may be asymptotically correctly sized, but this is only possible if A has a normal distribution, i.e. if the durations satisfy an ordinary central limit theorem. This would happen, for example, if the durations are i.i.d. with finite variance (as would be the case for the Poisson process), or if the durations obey an ACD model with finite variance (see Example 3.1 below), but not if A is non-normal, as can happen in an example given in Surgailis [2004]. Even when $\gamma = 1/2$ and A is normal, t_n would only be asymptotically standard normal if $\lim_{t\to\infty} \mathrm{var}[(N(t) - \lambda t)/t^{1/2}] = \mathrm{var} N(1)$, which would hold if N is a Poisson process but would fail if counts are autocorrelated as would typically be the case.

Remark 3.1. We have assumed explicitly in this paper that the true value of μ is nonzero. Thus, we have excluded in the analysis above the asymptotic behavior of the t-statistic when the null hypothesis $H_0 : \mu^* = 0$ holds. In this case, the first term in (2.2) drops out and so does the dependence of the properties of the t-statistic on A. Indeed, it follows from the proof of Theorem 2.1 that if $\gamma \geq 1/2$ and $\mu^* = \mu_0^* = 0$ then $t_n \rightarrow Z$, so the t-test remains asymptotically correctly sized. Still, the problem of constructing a confidence interval for μ^* would be very difficult since once the possibility that $\mu^* \neq 0$ is entertained the distribution of the re-centered statistic $n^{1/2-\gamma}(\bar{r}_n - \mu^*)/s_n$ may depend on the parameter of interest μ^*, as well as on A, which may have any of a wide variety of distributions (unknown *a priori*). It is clear, then, that feasible statistical inference on μ^* based on $\bar{r}_n$ is difficult or impossible in the absence of knowledge of the generating mechanism for the durations $\{\tau_k\}$ or for the point process N.

Remark 3.2. An economically-motivated null hypothesis for which μ_0^* is not zero would arise if one wished to test whether the expected return for a particular stock exceeds the risk-free rate (assumed fixed and known). One might try to turn the problem into a hypothesis test for a zero mean by working with the excess returns (difference between the actual return and the fixed risk-free rate). Unfortunately, the t-test for the null hypothesis that the expected excess return is zero would in general fail to be asymptotically correctly sized, since the asymptotic distribution for the corresponding t-statistic remains exactly as in Equations (3.1) and (3.2), with the same value of μ, i.e. the expectation of the return shock $\tilde{e}_k$. In the absence of an equity premium, so that the expected return is equal to the risk-free rate, μ in our model would be equal to λ^{-1} times the risk-free rate, so we would have $\mu > 0$ even though the expected excess return is zero. The key point is that subtracting a linear time trend from both sides of Equation (2.2) does not change μ, and therefore does not prevent the t-statistic from having nonstandard asymptotic properties. A similar argument would hold if the risk-free rate is taken to be observable and stochastic, but then one would also need to make assumptions about the statistical behavior of the risk-free rate.

We next consider two examples of generating mechanisms for the durations: the ACD model and a generalized version of the LMSD model. In both cases, the durations are assumed to form a stationary process. Unfortunately, it is known (see, for example, Nieuwenhuis [1989]) that except for the Poisson process there is no single probability measure under which both the durations and the associated counting process are stationary. We refer to the measure under which durations are stationary as the Palm measure, denoted by P^0. The Palm theory of point processes (see also Baccelli and Brémaud [2003]) guarantees the existence, under suitable regularity conditions, of a corresponding measure, denoted by P, under which the point process N is stationary. An economic interpretation of the Palm duality was provided by Deo et al. [2009]. We will use E to denote expectation under the P measure, and E^0, cov^0 to denote expectation and covariance under the Palm measure, P^0.

Example 3.1 (ACD durations). Assume that under the Palm measure P^0 the durations form a stationary ACD(1,1) process, defined by

$$\tau_k = \psi_k \epsilon_k , \qquad \psi_k = \omega + \alpha \tau_{k-1} + \beta \psi_{k-1} , \qquad k \in \mathbb{Z}, \tag{3.3}$$

where $\omega > 0$ and $\alpha, \beta \geq 0$, $\{\epsilon_k\}_{k=-\infty}^{\infty}$ is an i.i.d. sequence with $\epsilon_k \geq 0$ and $E^0[\epsilon_0] = 1$. If $\alpha + \beta < 1$, there exists a strictly stationary solution determined by $\tau_k = \omega \epsilon_k \sum_{j=1}^{\infty} \prod_{i=1}^{j-1} (\alpha \epsilon_{k-i} + \beta)$, with finite mean $E^0[\tau_0] = \omega/(1-\alpha-\beta)$. Moreover, by [Carrasco and Chen, 2002, Proposition 17], if ϵ_0 has a positive density on $[0, \infty)$, then the sequence $\{\tau_k\}$ is geometrically β-mixing. The tail index κ of a ACD(1,1) process is the solution of the equation

$$E^0[(\alpha \epsilon_0 + \beta)^\kappa] = 1 .$$

Moreover the stationary distribution satisfies $P^0(\tau_1 > x) \sim c x^{-\kappa}$ for some positive constant c. See e.g. Basrak et al. [2002]. If $1 < \kappa < 2$, then $n^{-1/\kappa} \sum_{k=1}^{n} (\tau_k - E^0[\tau_0])$ converges to a totally skewed to the right κ-stable law. Cf. Bartkiewicz et al. [2011, Proposition 5].

A necessary and sufficient condition for $E^0[\tau_0^2] < \infty$ is $E^0[(\alpha\epsilon_0+\beta)^2] = \alpha^2 E^0[\epsilon_0^2]+2\alpha\beta+\beta^2 < 1$. Cf. Giraitis and Surgailis [2002, Example 3.3]. Under this condition, it also holds that $\sum_{k=1}^{\infty} \mathrm{cov}^0(\tau_0, \tau_k) < \infty$. Since the ACD process is associated (positively dependent), the summability of the covariance function implies the central limit theorem for the partial sums. See Newman and Wright [1981], Giraitis and Surgailis [2002, Theorem 6.2].

The previous convergences hold under P^0, hence also under the corresponding measure P for which the point process is stationary. See Nieuwenhuis [1989] and Deo et al. [2009].

In order to check the condition $E[N^2(1)] < \infty$ of Lemma 3.1, we must assume that $E^0[\tau_0^{q+1}] < \infty$ for some $q > 4$. See Aue et al. [2011, Lemma 4.11]. A sufficient condition is $E^0[\alpha\epsilon_0 + \beta)^{q+1}] < \infty$. See Carrasco and Chen [2002]. This rules out the convergence to a stable law.

Example 3.2 (LMSD durations). Assume that under the Palm measure P^0, the durations form a stationary LMSD process, defined by $\tau_k = \epsilon_k \sigma(Y_k)$, where $\{\epsilon_k, k \in \mathbb{Z}\}$ is an i.i.d. sequence of almost surely positive random variables with finite mean and $\{Y_k, k \in \mathbb{Z}\}$ is a stationary standard Gaussian process, independent of $\{\epsilon_k\}$ and σ is a positive function. Deo et al. [2010] made two assumptions not required here, namely that $E^0[\epsilon_k^2] < \infty$ and that $\sigma(Y_k) = \exp(Y_k)$.

Let τ be the Hermite rank of the function $\sigma - E^0[\sigma(Y_0)]$. Assume that the covariance of the Gaussian process $\{Y_k\}$ is regularly varying at infinity, i.e.

$$\rho_n = \mathrm{cov}^0(Y_0, Y_n) = \ell(n) n^{2H-2} ,$$

where $H \in (1/2, 1)$. Assume first that $E^0[\epsilon_k^2] < \infty$. Denote $\lambda^{-1} = E^0[\epsilon_0] E^0[\sigma(Y_0)]$. The following dichotomy is well known. See Embrechts and Maejima [2002, Chapter 3].

- If $\tau(1-H) < 1/2$, then

$$n^{-1}\rho_n^{-\tau/2} \sum_{k=1}^{[n\cdot]} (\tau_k - \lambda^{-1}) \ \Rightarrow \ \varsigma R_{\tau,H} ,$$

  where ς is a nonzero constant, $R_{1,H}$ is the standard fractional Brownian motion, and for $q \geq 2$ such that $q(1-H) < 1/2$, $R_{q,H}$ is the so-called Hermite or Rosenblatt process of order q and self-similarity index $1 - q(1-H)$.

- If $\tau(1-H) > 1/2$, then

$$n^{-1/2} \sum_{k=1}^{n} (\tau_k - \lambda^{-1}) \to N(0, \Sigma) ,$$

  with Σ a positive constant.

We thus see that (2.3) may hold with $\gamma = 1-\tau(1-H) > 1/2$ in the first case and a possibly non Gaussian limit, and $\gamma = 1/2$ with a Gaussian limit in the latter case. Consider now the case where ϵ_k has infinite variance. Assume then that

$$P^0(\epsilon_k > x) = L(x)x^{-\alpha} ,$$

with $\alpha \in (1,2)$ and L a slowly varying function. It is then shown in Kulik and Soulier [2012] that a similar dichotomy exists.

- If $\tau(1-H) < 1/\alpha$, then

$$n^{-1}\rho_n^{-\tau/2} \sum_{k=1}^{[n\cdot]} (\tau_k - \lambda^{-1}) \ \Rightarrow \ \varsigma R_{\tau,H} \ .$$

- If $\tau(1-H) > 1/\alpha$, then

$$n^{-1/\alpha}\tilde{L}(n) \sum_{k=1}^{[n\cdot]} (\tau_k - \lambda^{-1}) \ \Rightarrow \ L_\alpha \ ,$$

  where L_α is a totally skewed to the right α-stable Lévy process.

We thus see that (2.3) may hold with $\gamma = 1-\tau(1-H) > 1/2$ in the first case and a possibly non Gaussian limit, and $\gamma = 1/\alpha$ with a stable non Gaussian limit in the latter case. These convergences hold under P^0, hence also under P.

In order to check the condition $E[N^2(1)] < \infty$ of Lemma 3.1, we must assume that $E^0[\epsilon_0^{q+1}] < \infty$ for some $q > 2/(1-H)$. See Aue et al. [2011, Lemma 4.11]. This rules out the convergence to a stable law.

4 Discussion: Long Memory and Heavy Tails of Stock Returns

The introduction of a nonzero mean in the efficient shocks in the model (2.1) provides a link by which properties of intertrade durations can affect those of certain quantities that are observed at a macroscopic level. We have focused so far on inference for the trend (based on studying the asymptotic distribution of the log price). To illustrate just one of the variety of possible additional quantities of interest, we now turn our attention to properties of returns.

Lo [1991] investigated whether stock returns have long memory, and Mandelbrot [1963] argued that returns have infinite variance. Both of these propositions have met with considerable controversy, but under the model (2.2) both could contain an important grain of truth. Generalizing the analysis presented so far leads to a more nuanced interpretation of what these propositions could mean.

From here on in this section, when we mention sequences of random variables, we allow for suitable renormalization (centering and scaling) without always specifically mentioning or writing the renormalization. So the discussion here is somewhat informal, but can be made mathematically rigorous. We focus here on the case $\gamma > 1/2$.

Theorem 2.1 implies that partial sums of returns (after suitable renormalization) converge in distribution to a random variable that need not be Gaussian. This theorem has allowed us to discuss issues related to inference for the slope parameter, which is the expectation of the average return. It is also of interest to ask if one can go further and say something about the joint distribution of the returns themselves, rather than their sum. By making the stronger assumption of functional convergence of partial sums of durations to a limiting stochastic process (which is the case for all examples considered in this paper) then it is indeed possible to discuss the joint limiting distribution of any fixed number of contiguous returns at long horizons.

Although we have so far taken the time spacing in defining the returns to be 1, there is no essential reason for this and here we replace it by an arbitrary $T > 0$, and we define the returns with respect to this time spacing as $r_{j,T} = y(jT) - y((j-1)T)$. Now consider the first M of these returns, where M is fixed. It follows from our assumptions here (by arguments similar to the proof of Theorem 2.1 and by Theorem 7.3.2 of Whitt [2002]) that the joint distribution of these M returns (after suitable renormalization) converges as $T \to \infty$ to the distribution of M contiguous increments of the limiting process.

In our LMSD example, assuming finite variance and an exponential volatility function, the limiting process is fractional Brownian motion. Thus in this case, the M returns converge in distribution to M contiguous observations of a fractional Gaussian noise. In this sense, it could be said that the returns (computed at a sufficiently high level of aggregation) have long memory. Simulations not shown here of the model (2.2) in this LMSD case show that it may be hard to detect this long memory due to the additive noise that arises from the second term on the righthand side of (2.2).

In the ACD example (and certain cases of the LMSD example as well), it turns out that the limiting process can be a stable process, for which the increments are independent and have infinite-variance stable distributions. So here, our M long-horizon returns converge in distribution (as $T \to \infty$, and after suitable renormalization) to a sequence of M i.i.d. stable random variables. This would seem to correspond to the proposition that returns have infinite variance. But actually, the truth here may be more subtle. It can happen that, for each fixed T the variance of the returns is finite. See, for example, Daley et al. [2000] for the underlying point process theory under heavy-tailed durations. It is even possible to construct an example where durations have finite variance and still the limit of partial sums of durations is a stable process, so the returns would once again have finite variance but converge in distribution to i.i.d. stable random variables with infinite variance. Such an example may come from durations that obey a positive version of the renewal-reward process discussed in Levy and Taqqu [1991] (see also Hsieh et al. [2007]). In such a model, durations would have finite variance but their sums would converge to a process with infinite variance.

In the case where the limiting process is a Lévy-stable process, it is of interest to note that such continuous-time processes have discontinuities with probability 1. These may correspond to what practitioners refer to as "jumps" in the log price process, even though under our model the log price process is a pure-jump process so that all activity consists of jumps.

The main message of this paper is that even in the simple transaction-level model (2.1) there is a wide variety of possible behaviors of macroscopic quantities of interest. Some additional quantities we hope to study in future work based on this and similar models include: regression coefficients as used in the market model, estimated cointegrating parameters (which were considered without a trend term in Aue et al. [2011]), and sample autocorrelations.

5 Proofs of Mathematical Results

In this section, we prove the results of the previous sections in a more general framework. Specifically, we introduce a microstructure noise term which may be dependent on the counting process N, thereby allowing for leverage effects. In the mathematical theory presented in this paper, all random variables and stochastic processes are defined on a single probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Expectation with respect to $\mathbb{P}$ will be denoted by $\mathbb{E}$ and var and cov will denote the variance and covariance with respect to $\mathbb{P}$. Convergence in $\mathbb{P}$-probability will be denoted by $\stackrel{\mathbb{P}}{\to}$, convergence in distribution under $\mathbb{P}$ of sequences of random variables will be denoted by $\to$. We use $\Rightarrow$ to denote weak convergence under $\mathbb{P}$ in the space $\mathcal{D}([0,\infty))$ of left-limited, right-continuous (càdlàg) functions, endowed with Skorohod's J_1 topology. See Billingsley [1968] or Whitt [2002] for details about weak convergence in $\mathcal{D}([0,\infty))$. Whenever the limiting process is continuous, this topology can be replaced by the topology of uniform convergence on compact sets. The model is now described as follows.

$$y(t) = \mu N(t) + \sum_{k=1}^{N(t)} \{e_k + \eta_k\} \, , \tag{5.1}$$

where the sequence $\{\eta_k\}$ satisfies the assumption

$$n^{-1/2} \sum_{k=1}^{[n\cdot]} \eta_k \ \Rightarrow \ 0 \, . \tag{5.2}$$

Theorem 5.1. *Under (2.3) and (5.2), if $\gamma > 1/2$, then $n^{-\gamma}(y(n) - \lambda\mu n) \to -\mu\lambda^{1+\gamma} A$. If $\gamma = 1/2$, then $n^{-1/2}(y(n) - \lambda\mu n) \to -\mu\lambda^{3/2} A + \sqrt{\lambda}\sigma_e Z$, where Z is a standard Gaussian random variable, independent of A.*

Theorem 2.1 follows from Theorem 5.1 by taking $\eta_k \equiv 0$.

Proof of Theorem 5.1. Write

$$y(n) - \lambda\mu n = \mu\{N(n) - \lambda n\} + \sum_{k=1}^{N(n)} e_k + \sum_{k=1}^{N(n)} \eta_k$$

The assumption on the durations implies that $n^{-1}\sum_{k=1}^{n} \tau_k \xrightarrow{\mathbb{P}} \lambda^{-1}$ and by the CLT equivalence, they imply that $n^{-\gamma}\mu(N(n)-n\lambda) \to -\mu\lambda^{1+\gamma}A$. Thus it also holds that $N(n)/n \xrightarrow{\mathbb{P}} \lambda$. Denote $x(n) = \sum_{k=1}^{N(n)} e_k + \sum_{k=1}^{N(n)} \eta_k$. It is proved in Aue et al. [2011] that $n^{-1/2}x([n\cdot]) \Rightarrow \sqrt{\lambda}\sigma_e B(1)$, where B is a standard Brownian motion, hence $n^{-1/2}x(n) \to N(0, \lambda\sigma_e^2)$. Note moreover that $n^{-1/2}\sum_{k=1}^{N(n)} \eta_k \to 0$, and since $\{e_k\}$ is independent of N, $n^{-1/2}x(n)$ and $n^{-\gamma}(N(n) - n\lambda)$ converge jointly. □

Lemma 5.1. *Assume that the marked point process with marked points (t_k, e_k, η_k) is stationary and ergodic. Assume moreover that (2.3) and (5.2) hold, and*

$$\mathbb{E}\left[\left(\sum_{k=1}^{N(1)} (\mu + e_k + \eta_k)\right)^2\right] < \infty \, . \tag{5.3}$$

Then, there exists $\sigma > 0$ such that $s_n^2 \xrightarrow{\mathbb{P}} \sigma^2$.

Proof of Lemma 5.1.

$$s_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(r_j - \lambda\mu)^2 + \frac{n}{n-1}(\bar{r}_n - \lambda\mu)^2$$

By Theorem 5.1, the second term is $o_P(1)$. Since

$$r_j - \lambda\mu = \sum_{k=N(j-1)+1}^{N(j)} (\mu + e_k + \eta_k) - \mu\lambda \, ,$$

by ergodicity of the marked point process, we have

$$\frac{1}{n}\sum_{j=1}^{n}(r_j - \lambda\mu)^2 \xrightarrow{\mathbb{P}} \mathbb{E}\left[\left(\sum_{k=1}^{N(1)}(\mu + e_k + \eta_k) - \lambda\mu\right)^2\right] \, .$$

□

Remark 5.1. If $\{t_k, z_k, k \in \mathbb{Z}\}$ are the (marked) points of a stationary (under $\mathbb{P}$) marked point process with finite intensity λ, then, by Baccelli and Brémaud [2003, Formula 1.2.9], for all $t > 0$

$$\mathbb{E}\left[\sum_{k=1}^{N(t)} z_k\right] = \lambda t E^0[z_0] \;,$$

where E^0 is the expectation with respect to the Palm probability P^0. If the marks $\{z_k\}$ have zero mean under the Palm measure P^0, then $\mathbb{E}\left[\sum_{k=1}^{N(t)} z_k\right] = 0$, and thus the trend is $\mathbb{E}[y(t)] = \mu\mathbb{E}[N(t)] = \lambda\mu t$, even if, under $\mathbb{P}$, it might happen that $\mathbb{E}[z_0] = \lambda E^0[t_1 z_1] \neq 0$.

Remark 5.2. Since the sequence $\{e_k\}$ has zero mean, finite variance and is independent of the point process N, if $\mathbb{E}[N^2(1)] < \infty$, then $\mathbb{E}[\{\sum_{k=1}^{N(t)}(\mu + e_k)\}^2] < \infty$. Next, for r, s such that $1/r + 1/s = 1$,

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{k=1}^{N(1)} \eta_k\right)^2\right] &= \sum_{k=1}^{\infty} \mathbb{E}\left[\left(\sum_{j=1}^{k} \eta_j\right)^2 \mathbf{1}_{\{N(1)=k\}}\right] \\ &\leq \sum_{k=1}^{\infty} \mathbb{E}^{1/s}\left[\left|\sum_{j=1}^{k} \eta_j\right|^{2s}\right] \mathbb{P}^{1/r}(N(1) = k) \;.\end{aligned}$$

Assume now that there exists some constant C_s such that

$$\mathbb{E}\left[\left|\sum_{j=1}^{k} \eta_j\right|^{2s}\right] \leq C_s k^s \;, \tag{5.4}$$

then, for $q > 0$, by Hölder's inequality,

$$\mathbb{E}\left[\left(\sum_{k=1}^{N(1)} \eta_k\right)^2\right] \leq C \sum_{k=1}^{\infty} k \mathbb{P}^{1/r}(N(1) = k) \leq C \left(\sum_{k=1}^{\infty} k^{-qs}\right)^{1/s} \mathbb{E}^{1/r}[N^{(q+1)r}(1)] \;.$$

If we can choose q such that $qs > 1$ and $\mathbb{E}[N^{(q+1)r}(1)] < \infty$, then the right hand side is finite and thus (5.3) holds.

Example 5.1 (Leverage). We now provide a specific example of a microstructure noise series $\{\eta_k\}$ that is dependent on N. Assume that under the Palm measure P^0, the durations form an LMSD sequence $\tau_k = \epsilon_k \mathrm{e}^{Y_k}$ as in Example 3.2, with memory parameter $d_\tau = H_\tau - 1/2 \in (0, 1/2)$. Assume moreover that the spectral density f_Y of the Gaussian process Y satisfies $f_Y(x) = |1 - \mathrm{e}^{\mathrm{i}x}|^{-2d_\tau} h(x)$, where the function h is slowly varying at 0. Let now η_k be defined as follows:

$$\eta_k = [(I - B)^\delta Y]_k \;,$$

and δ is such that $\delta \in (d_\tau, \infty)$ where B is the backshift operator. Note that $E^0[\eta_0] = 0$. The spectral density of the weakly stationary sequence $\{\eta_k\}$ is given by

$$f_\eta(x) = |1 - \mathrm{e}^{\mathrm{i}x}|^{2\delta} f_Y(x) = |1 - \mathrm{e}^{\mathrm{i}x}|^{2\delta - 2d_\tau} h(x) .$$

Thus the sequence $\{\eta_k\}$ has negative memory $d_\eta = d_\tau - \delta$ and Assumption (5.2) holds. There remain to check Condition (5.3). Since (5.4) holds for all $s \geq 1$, q and r can be chosen arbitrarily close to 0 and 1, respectively, and thus (5.3) holds if $\mathbb{E}[N^2(1)] < \infty$. If δ is chosen to be in $(d_\tau, d_\tau + 1/2)$, then $d_\eta \in (-1/2, 0)$. And if $\delta = 1$, then $\eta_k = Y_k - Y_{k-1}$.

References

Torben G. Andersen, Tim Bollerslev, Francis X. Diebold, and Paul Labys. The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc.*, 96 (453):42–55, 2001. ISSN 0162-1459. doi: 10.1198/016214501750332965. URL http://dx.doi.org/10.1198/016214501750332965.

Alexander Aue, Clifford M. Hurvich, Philippe Soulier, and Lajos Horváth. Limit laws in transaction-level asset price models. arxiv:1103:xxxx, 2011.

François Baccelli and Pierre Brémaud. *Elements of queueing theory*, volume 26 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2003.

Emmanuel Bacry, Sylvain Delattre, Marc Hoffmann, and Jean-François Muzy. Modelling microstructure noise with mutually exciting point processes. To be published in Quantitative Finance, 2011.

Katarzyna Bartkiewicz, Adam Jakubowski, Thomas Mikosch, and Olivier Wintenberger. Stable limits for sums of dependent infinite variance random variables. *Probability Theory and Related Fields*, 150(3-4):337–372, 2011.

Bojan Basrak, Richard A. Davis, and Thomas Mikosch. A characterization of multivariate regular variation. *The Annals of Applied Probability*, 12(3):908–920, 2002.

Luc Bauwens and David Veredas. The stochastic conditional duration model: a latent variable model for the analysis of financial durations. *Journal of Econometrics*, 119(2): 381–412, 2004. Dynamic factor models.

Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

Clive G. Bowsher. Modelling security market events in continuous time: intensity based, multivariate point process models. *J. Econometrics*, 141(2): 876–912, 2007. ISSN 0304-4076. doi: 10.1016/j.jeconom.2006.11.007. URL http://dx.doi.org/10.1016/j.jeconom.2006.11.007.

Marine Carrasco and Xiaohong Chen. Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory*, 18(1):17–39, 2002.

Fei Chen, Francis X. Diebold, and Frank Schorfheide. A markov-switching multi-fractal inter-trade duration model, with application to u.s. equities. Preprint, 2012.

D. J. Daley, T. Rolski, and Rein Vesilo. Long-range dependent point processes and their Palm-Khinchin distributions. *Adv. in Appl. Probab.*, 32(4): 1051–1063, 2000. ISSN 0001-8678. doi: 10.1239/aap/1013540347. URL http://dx.doi.org/10.1239/aap/1013540347.

Rohit Deo, Clifford M. Hurvich, Philippe Soulier, and Yi Wang. Conditions for the propagation of memory parameter from durations to counts and realized volatility. *Econometric Theory*, 25(3):764–792, 2009.

Rohit Deo, Mengchen Hsieh, and Clifford M. Hurvich. Long memory in intertrade durations, counts and realized volatility of nyse stocks. *Journal of Statistical Planning and Inference*, 140(12):3715 – 3733, 2010. doi: DOI:10.1016/j.jspi.2010.04.037. Special Issue in Honor of Emanuel Parzen on the Occasion of his 80th Birthday and Retirement from the Department of Statistics, Texas A&M University - Emmanuel Parzen.

Paul Embrechts and Makoto Maejima. *Selfsimilar processes*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2002.

Robert F. Engle and Jeffrey R. Russell. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, 66(5):1127–1162, 1998.

Liudas Giraitis and Donatas Surgailis. ARCH-type bilinear models with double long memory. *Stochastic Processes and their Applications*, 100:275–300, 2002.

Nikolaus Hautsch. *Econometrics of financial high-frequency data*. Springer, Heidelberg, 2012. ISBN 978-3-642-21924-5. doi: 10.1007/978-3-642-21925-2. URL http://dx.doi.org/10.1007/978-3-642-21925-2.

Meng-Chen Hsieh, Clifford M. Hurvich, and Philippe Soulier. Asymptotics for duration-driven long range dependent processes. *Journal of Econometrics*, 141(2):913–949, 2007.

Clifford M. Hurvich and Yi Wang. A pure-jump transaction-level price model yielding cointegration. *Journal of Business & Economic Statistics*, 28(4):539–558, 2010.

Rafał Kulik and Philippe Soulier. Limit theorems for long memory stochastic volatility models with infinite variance: Partial sums and sample covariances. *Advances in Applied Probability*, 44(4), 2012.

Joshua B. Levy and Murad S. Taqqu. A characterization of the asymptotic behavior of stationary stable processes. In *Stable processes and related topics (Ithaca, NY, 1990)*, volume 25 of *Progr. Probab.*, pages 181–198. Birkhäuser Boston, Boston, MA, 1991.

Andrew W. Lo. Long-term memory in stock market prices. *Probability Theory and Related Fields*, 59:1279–1313, 1991.

Benoit Mandelbrot. The variation of certain speculative prices. *Journal of Business*, 36 (4), 1963.

Daniel B. Nelson. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59(2):347–370, 1991. ISSN 0012-9682. doi: 10.2307/2938260. URL http://dx.doi.org/10.2307/2938260.

Charles M. Newman and A. Larry Wright. An invariance principle for certain dependent sequences. *The Annals of Probability*, 9(4):671–675, 1981.

Gert Nieuwenhuis. Equivalence of functional limit theorems for stationary point processes and their Palm distributions. *Probability Theory and Related Fields*, 81(4):593–608, 1989.

Myron Scholes and Joseph Williams. Estimating betas from nonsynchronous data. *Journal of Financial Economics*, 5:309–327, 1977.

Nikhil Shenai. Dynamic modelling of irregular times, prices and volumes at high frequencies. PhD Thesis, Imperial College Business School, 2012.

Donatas Surgailis. Stable limits of sums of bounded functions of long-memory moving averages with finite variance. *Bernoulli*, 10(2):327–355, 2004. ISSN 1350-7265. doi: 10.3150/bj/1082380222. URL http://dx.doi.org/10.3150/bj/1082380222.

Ward Whitt. *Stochastic-process limits*. Springer Series in Operations Research. Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.