HAL
open science

Drift in Transaction-Level Asset Price Models

Wen Cao, Clifford Hurvich, Philippe Soulier

To cite this version:

Wen Cao, Clifford Hurvich, Philippe Soulier. Drift in Transaction-Level Asset Price Models. 2012. hal-00756372v2

HAL Id: hal-00756372

https://hal.science/hal-00756372v2

Preprint submitted on 5 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Drift in Transaction-Level Asset Price Models

Wen Cao* Clifford Hurvich† Philippe Soulier‡

January 6, 2015

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 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 the point process 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. As a remedy to these problems, we propose a new ratio statistic which behaves more

*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 Ouest, 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.

robustly, and employ subsampling methods to carry out inference for the growth rate. Our analysis also sheds some new light on two longstanding debates as to whether stock returns have long memory or infinite variance.

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 Prigent (2001), Bowsher (2007), Billingsley (1968), Hautsch (2012), Bacry et al. (2011), Deo et al. (2009), Deo et al. (2010), Hurvich and Wang (2010), Aue et al. (2014), 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 the point process 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 present Proposition 2.1 on the limiting behavior of the log price process, as determined by the properties of the point process. In Section 3 we begin a study of statistical inference for the trend, and obtain the behavior of the ordinary t-statistic under the null hypothesis. We also propose a ratio statistic which behaves robustly, in that (unlike the t-statistic) it converges in distribution under a broad range of conditions. This motivates our discussion in Section 4 of the use of subsampling based on the ratio statistic to conduct statistical inference on the trend. We then present, in Section 5, a series of examples of specific duration and point process models that have been proposed in the literature, including the Autoregressive conditional duration (ACD) model of Engle and Russell (1998) and the long memory stochastic duration (LMSD) model of Deo et al. (2010). These examples provide for great diversity of the asymptotic distributions of sums of durations and therefore (by Proposition 2.1 below) for the asymptotic distribution of the log price. Simulations are presented in Section 6. There, we consider the size and power for both the t-test and the test based on the new ratio statistic under three duration models: *iid* exponential, ACD and LMSD. The parameters of simulations are calibrated using empirical data. Section 7 provides a data analysis in which we employ the t-test and the test based on the new ratio statistic to gauge the strength of the equity premium. Section 8 provides a concluding discussion on how our results may help to reconcile some longstanding debates. Proofs of the mathematical results are provided in Section 9.

2 A Univariate Model for Log Price

We start with a basic univariate pure-jump model for a log price series y. Let N_1, N_2 be mutually independent point processes on the real line. Points of $N_1 + N_2$ correspond to transactions on the asset[1] and $(N_1 + N_2)(t)$ is the number of

[1]From a modeling perspective it may be desirable to instead have the points of $N_1 + N_2$ 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 $(N_1+N_2)(t)$ count actual transactions, but our theoretical results do not depend on this particular choice of

transactions in $(0, t]$. We set $y(0) = 0$ and define, for $t \geq 0$,

$$y(t) = \sum_{k=1}^{N_1(t)} (\tilde{e}_{1,k} + \eta_{1,k}) + \sum_{k=1}^{N_2(t)} (\tilde{e}_{2,k} + \eta_{2,k}) \tag{2.1}$$

where for $i = 1, 2$,

$$\tilde{e}_{i,k} = \mu_i + e_{i,k}$$

and the $\{e_{i,k}\}$, which are independent of (N_1, N_2), are i.i.d. with zero mean and finite variance σ_i^2. We assume that the microstructure noise sequences $\{\eta_{i,k}\}$ are independent of the efficient shocks $\{e_{1,k}\}$, $\{e_{2,k}\}$, but not necessarily of the counting process (N_1, N_2). This allows for a leverage effect. See Aue et al. (2014). Assumption 2.2 below implies that the microstructure noise becomes negligible after aggregation.

We assume that μ_1 and μ_2 are constants, not both zero. This model with $\mu_1 = \mu_2 = 0$, $\sigma_2^2 = 0$ and $\{\eta_{2,k}\} = 0$ was considered in Deo et al. (2009), who showed that it can produce long memory in the realized volatility. We generalize the model here to allow for two driving processes, N_1 and N_2. It follows from (2.1) that

$$y(t) = \mu_1 N_1(t) + \mu_2 N_2(t) + \sum_{k=1}^{N_1(t)} (e_{1,k} + \eta_{1,k}) + \sum_{k=1}^{N_2(t)} (e_{2,k} + \eta_{2,k}) \ . \tag{2.2}$$

The quantity $\mu_1 N_1(t) + \mu_2 N_2(t)$ can be viewed as a random drift term in the log price $y(t)$. If, for example, $\mu_1 > 0$ and $\mu_2 < 0$ then we can think of N_1 and N_2 as governing the activity of buyers and sellers, respectively. An active period for N_1 puts upward pressure on the price, while an active period for N_2 exerts downward pressure. One could envision including other point processes to differentiate, say, between information traders and noise traders, but we find that the model with just N_1 and N_2 is quite flexible and able to reproduce the stylized facts.

To simplify the discussion, we assume temporarily that N_1 and N_2 are stationary point processes with intensities λ_1 and λ_2. Then the expectation of the drift term is a linear trend, that is, $\mathbb{E}[\mu_1 N_1(t) + \mu_2 N_2(t)] = \mu^* t$, where

$$\mu^* = \lambda_1 \mu_1 + \lambda_2 \mu_2 \ .$$

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 the random drift term $\mu_1 N_1(t) + \mu_2 N_2(t)$ is replaced by the deterministic

the definition of an event.

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 $\mathbb{E}[y(t)]$ is a linear function of t, to account for such phenomena as equity premia and inflation. This is indeed the case for Model (2.2) if it is assumed in addition that the microstructure noise terms have zero mean under their Palm distribution (see below), which implies $\mathbb{E}[y(t)] = \mu^* t$. But in actual realizations of $y(t)$, the trend is only impounded when a transaction occurs, due to the nonzero means in $\{\tilde{e}_{1,k}\}$ and/or $\{\tilde{e}_{2,k}\}$.

We now obtain conditions under which the two summations on the righthand side of (2.2) have mean zero, so that $\mathbb{E}[y(t)] = \mu^* t$, still under the assumption that N_1 and N_2 are stationary. Since the efficient shocks are assumed independent of N_1 and N_2 and have mean zero, it follows that $\mathbb{E}\left[\sum_{k=1}^{N_i(t)} e_{i,k}\right] = 0$. However, due to potential leverage-type effects, even if one were to assume that $\mathbb{E}[\eta_{i,k}] = 0$, this alone would not ensure that $\mathbb{E}\left[\sum_{k=1}^{N_i(t)} \eta_{i,k}\right] = 0$. Therefore, if we want to interpret expectation of the drift term as a linear trend in the log price, we must make the additional assumption that $\mathbb{E}\left[\sum_{k=1}^{N_i(t)} \eta_{i,k}\right] = 0$. As we will show, this assumption is equivalent to assuming that the $E_i^0[\eta_{i,k}] = 0$, where P_i^0 is the Palm probability measure associated with the point process N_i, and E_i^0 is the expectation under the Palm measure.

We briefly recall here (see Deo et al. (2009) for more details) that in general there is no single measure under which both the durations and the marked point process are stationary. If the durations for a point process N are stationary under a measure P^0, then we refer to P^0 as the Palm probability measure. This, if such a measure exists, would be the appropriate measure for assessing properties of durations as a stationary sequence. In such situations, there exists a corresponding measure $\mathbb{P}$ under which the point process is stationary. This would be the appropriate distribution for assessing counts (the number of events occurring in equally spaced intervals of time), and also the differences of $y(t)$ in (2.2), as a stationary time series in models that admit stationarity.

If $\{t_k, z_k, k \in \mathbb{Z}\}$ are the (marked) points of a stationary (under $\mathbb{P}$) marked point process N 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 $\mathbb{E}\left[\sum_{k=1}^{N(t)} z_k\right] = 0$, even if,

under $\mathbb{P}$, it might happen that $\mathbb{E}[z_0] = \lambda E^0[t_1 z_1] \neq 0$.

It follows that if the microstructure noise sequences $\{\eta_{i,k}\}$ have zero mean under their respective Palm measures P_i^0 then in the model (2.2) under $\mathbb{P}$ we have $\mathbb{E}[y(t)] = \mu^* t$.

There are a variety of reasons why it may be considered unrealistic to assume that the point process (N_1, N_2) is stationary. Indeed, it is well-known that the counts of financial transactions can show intraday seasonality (see, e.g., Engle and Russell (1998), Deo et al. (2010)), and the transaction counts will be guaranteed to be zero when the market is not open for trading. Furthermore, regime changes are widely understood to be an important feature of economic processes. We will consider a time deformation mechanism below which allows for such effects. In the absence of stationarity we are no longer able to compute $\mathbb{E}[y(t)]$ in (2.2), but we can still interpret μ^* as a long term trend, i.e. $y(t)/t \xrightarrow{\mathbb{P}} \mu^*$, under the assumptions below. These assumptions are stronger than needed for this purpose but will be used later to derive limiting distributions for suitably rescaled versions of y.

In the sequel, we will consider a single probability measure $\mathbb{P}$, without any requirement that (N_1, N_2) be stationary under $\mathbb{P}$, unless specified otherwise. Convergence in $\mathbb{P}$-probability will be denoted by $\xrightarrow{\mathbb{P}}$, convergence in distribution under $\mathbb{P}$ of sequences of random variables will be denoted by $\to$ and convergence of finite dimensional distributions of a sequence of stochastic processes will be denoted by $\xrightarrow{fi.di.}$.

Assumption 2.1. *The point processes N_1 and N_2 are mutually independent and there exists $\gamma \geq 1/2$ such that*

$$\{n^{-\gamma}(N_i(nt) - \lambda_i nt), t \geq 0\} \xrightarrow{fi.di.} A_i \tag{2.3}$$

for $i = 1, 2$, as $n \to \infty$ where A_i are stochastic processes, at least one of which is nonzero.

We will give examples of point processes satisfying Assumption 2.1 in Section 5. As will be seen, the limiting processes A_i can be Gaussian or have infinite variance, can have independent or dependent increments, and γ can take any value in $[1/2, 1)$. However, whatever A and γ, a consequence of the convergence (2.3) is that $N_i(t)/t \xrightarrow{\mathbb{P}} \lambda_i$ and $t^{-1/2} \sum_{k=1}^{N_i(t)} e_{i,k}$ converges weakly to a Brownian motion.

Two distinct modeling approaches seem natural. One is to model the point process directly as in Bacry et al. (2011) who use Hawkes processes or Delattre et al. (2013) who use Cox processes. Another approach consists of modeling the durations. This was done by Engle and Russell (1998) who defined the ACD model

and Deo et al. (2010) who defined the LMSD model. These two approaches are equivalent with respect to Assumption 2.1, since (2.3) with limit process A holds for a point process N with intensity λ and duration sequence $\{\tau_k\}$ if and only if

$$n^{-\gamma}\sum_{k=1}^{[n\cdot]}(\tau_k - \lambda^{-1}) \xrightarrow{fi.di.} -\lambda^{-(1+\gamma)}A \ , \tag{2.4}$$

where convergence holds in the sense of finite dimensional distributions. See Whitt (2002, Theorem 7.3.1).

Assumption 2.2. *The sequence of processes $n^{-1/2}\sum_{k=1}^{[n\cdot]}\eta_{i,k}$ converges in probability uniformly on compact sets to 0.*

Assumptions 2.1 and 2.2 imply that $t^{-1/2}\sum_{k=1}^{N_i(t)}(e_{i,k}+\eta_{i,k})$ converges weakly to a Brownian motion and so the microstructure terms are asymptotically negligible.

The following proposition, which follows easily from Assumptions 2.1 and 2.2, gives the limiting distribution for suitably rescaled versions of the log price process y. Recall that Assumption 2.1 implies that $N_i(t)/t \xrightarrow{\mathbb{P}} \lambda_i$. Let $\mu^* = \lambda_1\mu_1 + \lambda_2\mu_2$.

Proposition 2.1. *Let Assumptions 2.1 and 2.2 hold. Then $y(t)/t \xrightarrow{\mathbb{P}} \mu^*$. Moreover, as $n \to \infty$,*

$$\{n^{-\gamma}(y(nt) - \mu^* nt), t \geq 0\} \xrightarrow{fi.di.} \mu_1 A_1 + \mu_2 A_2 \ , \tag{2.5}$$

if $\gamma > 1/2$ and

$$\{n^{-1/2}(y(nt) - \mu^* nt), t \geq 0\} \xrightarrow{fi.di.} \mu_1 A_1 + \mu_2 A_2 + \sigma B \tag{2.6}$$

if $\gamma = 1/2$ or if $\mu_1 = \mu_2 = 0$, where B is a standard Brownian motion independent of A_i and σ is a positive constant.

Next, we discuss the use of time deformation. This discussion is taken from Aue et al. (2014). Let f be a deterministic or random function such that f is nondecreasing and has càdlàg paths with probability one. Let $\tilde{N}$ be a point process and define

$$N(t) = \tilde{N}(f(t)) \ .$$

We do not require that $\tilde{N}$ be stationary, but one useful application of time deformation is to start with a stationary process $\tilde{N}$ and then deform it as above to obtain a nonstationary process N. If the function f is random, we assume moreover that it is independent of $\tilde{N}$. The use of the time-varying intensity function f

may render the counting process N nonstationary even if $\tilde{N}$ is a stationary point process. Since it is possible that f has (upward) jumps, the point process N may also not be simple even though $\tilde{N}$ is simple. We now show, however, that if $\tilde{N}$ satisfies Assumption 2.1, then so does the time-deformed N under some restrictions on f.

Lemma 2.1. *Assume that f is a nondecreasing (random) function such that $t^{-1}f(t) \xrightarrow{\mathbb{P}} \delta \in (0,\infty)$ and*

$$\sup_{t\geq 0} |f(t) - f(t^-)| \leq C$$

with probability one, where $C \in (0,\infty)$ is a deterministic constant. Let $\tilde{N}$ be a point process such that Assumption 2.1 holds for some $\tilde{\lambda} \in (0,\infty)$ and $\gamma > 0$. Let N be the counting process defined by $N(\cdot) = \tilde{N}(f(\cdot))$. Then Assumption 2.1 holds for N with $\lambda = \tilde{\lambda}\delta$.

The function f is used to speed up or slow down the trading clock. To incorporate dynamic intraday seasonality in volatility, the same time deformation can be used in each trading period (of length, say, T), assuming that $f(t)$ has a periodic derivative (with period T and with probability one), for example, $f(t) = t + .5\sin(2\pi t/T)$. Fixed nontrading intervals, say, $t \in [T_1, T_2)$, could be accommodated by taking $f(t) = f(T_1)$ for $t \in [T_1, T_2)$ so that $f(t)$ remains constant for t in this interval, and then taking $f(T_2) > f(T_1)$ so that $f(t)$ jumps upward when trading resumes at time T_2. The jump allows for the possibility of one or more transactions at time T_2, potentially reflecting information from other markets or assets that did trade in the period $[T_1, T_2)$. Since it is possible that f has (upward) jumps, N may not be simple even though $\tilde{N}$ is simple.

3 Statistical inference for the trend

For integer k, (assuming a time-spacing of 1 without loss of generality) we define the calendar-time returns as $r_k = y(k) - y(k-1)$ and the average return over a time period that extends from time 0 to time n as $\bar{r}_n = n^{-1}y(n) = n^{-1}\sum_{k=1}^{n} r_k$. We can think of n as the length of the time interval spanned by the observations, or (since we have assumed a time-spacing of 1) as the sample size, *i.e.*, the number of observed returns. It follows from Proposition 2.1 that if $\gamma > 1/2$, $\bar{r}_n - \mu^*$ will not be $O_p(n^{-1/2})$, making it difficult to accurately estimate growth rates based on the data set $\{r_k\}_{k=1}^{n}$.

We recall that model (2.2) under stationarity of N_1 and N_2 implies that $\mathbb{E}[y(n)] = \mu^* n$, so that the growth rate per unit time is μ^*, but that under the more general

assumptions 2.1 and 2.2 μ^* may still be viewed as a long-term growth rate, since $y(n)/n \xrightarrow{\mathbb{P}} \mu^*$.

We consider the problem of statistical inference for μ^*. We focus on testing a null hypothesis of the form $H_0 : \mu^* = \mu_0^*$ based on $\bar{r}_n$.

3.1 The t statistic

The corresponding t-statistic for testing H_0 is

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

where

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

The sample variance s_n^2 consistently estimates a positive constant, under suitable moments assumptions.

Lemma 3.1. *Under Assumptions 2.1 and 2.2, if N_1 and N_2 are stationary and ergodic and*

$$\mathbb{E}\left[\left(\sum_{k=1}^{N_i(1)} (\mu_i + e_{i,k} + \eta_{i,k})\right)^2\right] < \infty \,, \tag{3.1}$$

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

Since the efficient shocks have zero mean and finite variance and are independent of the point processes, a sufficient condition for (3.1) to hold is $\mathbb{E}[N_i^2(1)] < \infty$ and

$$\mathbb{E}\left[\left(\sum_{k=1}^{N_i(1)} \eta_{i,k}\right)^2\right] < \infty \,.$$

If $\gamma > 1/2$ and μ_1 and μ_2 not both zero, it follows from the convergence (2.5) and Lemma 3.1 that if the null hypothesis is true, then $t_n = O_p(n^{\gamma - 1/2})$ and the t-statistic diverges under the null hypothesis. Examples where this scenario would occur include durations generated by an LMSD model with long memory and an exponential volatility function. This scenario therefore is consistent with the empirical properties of durations found in Deo et al. (2010).

If $\gamma = 1/2$ or if $\mu_1 = \mu_2 = 0$, then $t_n \to A(1)/\varsigma$ where A is the limiting process in (2.6), i.e. the t-statistic converges to a non degenerate distribution, but the t-test will not be asymptotically correctly sized, except under very specific circumstances. First the limits A_1, A_2 must have normal distributions. 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. Then it would also be necessary that $\lim_{t\to\infty} t^{-1/2}\text{var}[N_i(t)] = \text{var}[N_i(1)]$, which would hold for instance if N is a Poisson process but would fail if counts are autocorrelated as would typically be the case.

3.2 A ratio statistic

Instead of using the t-statistic which could be degenerate, unless the parameter γ is equal to 1/2, we propose to use another self-normalized statistic which will always have a non degenerate distribution and does not need additional assumptions. Define $\bar{r}(n) = \bar{r}_n$ and

$$T_n = \frac{\bar{r}(n) - \mu_0^*}{|\bar{r}(n) - \bar{r}(n/2)|} \ . \tag{3.2}$$

Theorem 3.1. *If Assumptions 2.1 and 2.2 hold, then under the null hypothesis $\mu^* = \mu_0^*$,*

$$T_n \to \frac{A(1)}{|A(1) - 2A(1/2)|} \ , \tag{3.3}$$

where A denotes the limiting process in (2.5) if $\gamma > 1/2$ and μ_1 and μ_2 not both zero, or (2.6) if $\gamma = 1/2$. If $\mu^* > \mu_0^*$, then $T_n \xrightarrow{\mathbb{P}} +\infty$ and if $\mu^* < \mu_0^*$, then $T_n \xrightarrow{\mathbb{P}} -\infty$.

4 Subsampling

Let $D_n(x)$ be the cumulative distribution of T_n, that is, $D_n(x) = \mathbb{P}(T_n \leq x)$. Let D denote the corresponding limiting distribution,

$$D = \frac{A(1)}{|A(1) - 2A(1/2)|}$$

To perform a hypothesis test on μ^* based on T_n, we need to approximate $D_n(x)$. Since the limiting distribution D will be unknown in practice, we will approximate $D_n(x)$ non-parametrically via subsampling.

In subsampling we split the sample into overlapping blocks of size b, where b depends on n ($b \to \infty, b/n \to 0$), given by $\{r_t, r_{t+1}, ..., r_{t+b-1}\}$, $t = 1, 2, ..., n-b+1$, and calculate the self-normalized statistic upon each block, treating each block as if it were a full sample. Moreover, the parameter μ_0^* is replaced by its full-sample estimate $\bar{r}(n)$. This leads to $n - b + 1$ subsampling statistics

$$T_{n,b,t} = \frac{\bar{r}(n,b,t) - \bar{r}(n)}{|\bar{r}(n,b,t) - \bar{r}(n,b/2,t)|},$$

$$t = 1, 2, ..., n - b + 1;\ b\ is\ even,$$

where

$$\bar{r}(n) = \frac{1}{n}\sum_{k=1}^{n} r_k,$$

$$\bar{r}(n,b,t) = \frac{1}{b}\sum_{k=t}^{t+b-1} r_k,$$

$$\bar{r}(n,b/2,t) = \frac{2}{b}\sum_{k=t}^{t+b/2-1} r_k.$$

The cumulative distribution $D_n(x)$ is then approximated by the empirical cdf of these statistics

$$\hat{D}_{n,b}(x) = \frac{1}{n-b+1}\sum_{t=1}^{n-b+1} 1_{\{T_{n,b,t} \le x\}}\ .$$

Let $c_{n,b}(\alpha)$ denote the α quantile of the subsampling distribution, $\hat{D}_{n,b}(x)$,

$$c_{n,b}(\alpha) = \inf\{x : \hat{D}_{n,b}(x) \ge \alpha\}.$$

Then the rejection region for the subsmapling test based on T_n in the two-sided case

$$H_0 : \mu^* = \mu_0^*$$
$$H_1 : \mu^* \neq \mu_0^*$$

is

$$\{T_n < c_{n,b}(\alpha/2)\ or\ T_n > c_{n,b}(1-\alpha/2)\}\ .$$

To show that the resulting subsampling test based on T_n has asymptotically the nominal size (α), we use Theorem 4 of Appendix B of Jach et al. (2012). In order to apply this theorem, we must establish the θ-weak dependence of the time

series $\{r_k\}$ (see Doukhan and Louhichi (1999); Bardet et al. (2008)). We will establish this θ-weak dependence for one of the processes we consider in Section 5 (see Proposition 5.1). We say that the subsampling estimator $\hat{D}_{n,b}(x)$ is consistent if $|\hat{D}_{n,b}(x) - D_n(x)| \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$ for all x. The following theorem can be proved based on Theorem 4, Appendix A of Jach et al. (2012) and its proof, as well as the discussion on page 941 of McElroy and Jach (2012).

Theorem 4.1. *Assume that the time series of calendar-time returns $\{r_n, n \geq 1\}$ is strictly stationary and θ-weak dependent with rate $\theta_h = O(h^{-a})$ up to slowly varying functions. Assume also that the cumulative distribution function of*

$$\frac{A(1)}{|A(1) - 2A(1/2)|}$$

in Theorem 3.1 is continuous. Then

(i) for $a \geq 1/2$, the subsampling estimator is consistent and the resulting test based on T_n is asymptotically correctly sized and has power tending to 1 under the alternative hypothesis for any choice of the block size b such that $b \to \infty$ and $b/n \to 0$;

(ii) for $a < 1/2$, the subsampling estimator is consistent and the resulting test based on T_n is asymptotically correctly sized and has power tending to 1 under the alternative hypothesis if $b \to \infty$ and $b = O(n^{\zeta})$ (up to slowly varying functions) for some $0 < \zeta < 2a$.

The theorem gives little guidance as to the choice of the block size b in practice. In the simulations of Section 6 , we will consider a variety of choices of the block size.

5 Examples

In this section we give several examples of point processes which satisfy Assumption 2.1 with $\gamma > 1/2$.

5.1 ACD durations

Assume that under the durations form an ACD(1,1) process, defined by

$$\tau_k = \psi_k \epsilon_k, \qquad \psi_k = \omega + \alpha \tau_{k-1} + \beta \psi_{k-1}, \qquad k \in \mathbb{Z}, \tag{5.1}$$

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 $\mathbb{E}[\epsilon_0] = 1$. 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 $\mathbb{E}[\tau_0] = \omega/(1-\alpha-\beta)$. The tail index κ of a ACD(1,1) process is the solution of the equation

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

Moreover the stationary distribution satisfies $\mathbb{P}(\tau_1 > x) \sim cx^{-\kappa}$ for some positive constant c. See e.g. Basrak et al. (2002).

- If $1 < \kappa < 2$, and if ϵ_0 has a positive density on $[0, \infty)$, then the finite dimensional distributions of $n^{-1/\kappa} \sum_{k=1}^{[nt]}(\tau_k - \mathbb{E}[\tau_0])$ converges to a totally skewed to the right κ-stable law. Cf. Bartkiewicz et al. (2011, Proposition 5).
- A necessary and sufficient condition for $\mathbb{E}[\tau_0^2] < \infty$ is $\mathbb{E}[(\alpha\epsilon_0+\beta)^2] = \alpha^2\mathbb{E}[\epsilon_0^2] + 2\alpha\beta + \beta^2 < 1$. Cf. Giraitis and Surgailis (2002, Example 3.3). Under this condition, it also holds that $\sum_{k=1}^{\infty} \mathrm{cov}(\tau_0, \tau_k) < \infty$. Since the ACD process is associated, the summability of the covariance function implies the functional central limit theorem for the partial sum process. See Newman and Wright (1981). Thus the convergence (2.3) holds with $\gamma = 1/2$ and the limit process is the Brownian motion.

This implies that the convergence (2.3) holds with $\gamma = 1/\kappa$ and the limit process is a Lévy-stable process if $\kappa < 2$ and with $\gamma = 1/2$ and the limit process is the Brownian motion if $\mathbb{E}[(\alpha\epsilon_0 + \beta)^2] < 1$.

5.2 LMSD durations

Assume that the durations form an 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. As in Deo et al. (2010), for simplicity we will assume that $\sigma(x) = \mathrm{e}^x$.

Assume that the covariance of the Gaussian process $\{Y_k\}$ is such that

$$\rho_n = \mathrm{cov}(Y_0, Y_n) \sim cn^{2H-2} \ ,$$

where $H \in (1/2, 1)$ and $c > 0$. Denote $\lambda^{-1} = \mathbb{E}[\epsilon_0]\mathbb{E}[\exp(Y_0)]$. Then we have the following possibilities.

- If $\mathbb{E}[\epsilon_k^2] < \infty$, then

$$n^{-H} \sum_{k=1}^{[n\cdot]} (\tau_k - \lambda^{-1}) \;\Rightarrow\; \varsigma B_H \;,$$

  where ς is a nonzero constant and B_H is the standard fractional Brownian motion.

- If $\mathbb{P}(\epsilon_1 > x) \sim c x^{-\alpha}$ as $x \to \infty$ with $\alpha \in (1,2)$, then $\mathbb{E}[\epsilon_k^2] = \infty$ and the following dichotomy is proved in Kulik and Soulier (2012).

  - If $H > 1 - 1/\alpha$, then

$$n^{-H} \sum_{k=1}^{[n\cdot]} (\tau_k - \lambda^{-1}) \;\Rightarrow\; \varsigma B_H \;.$$

  - If $H < 1 - 1/\alpha$, then

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

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

We thus see that (2.3) may hold with $\gamma = H$ in the first case and $\gamma = 1/\alpha$ with a stable non Gaussian limit in the latter case.

5.3 Superposition of independent point processes

This is an example showing that the limiting distribution of a centered and normalized counting process N may be stable (hence heavy tailed) even though its durations τ_k are light tailed. Let M_1 be a Poisson process with intensity λ and durations $\{\tau_k^{(1)}\}$ and M_2 be a renewal process, independent of M_1, with i.i.d. durations $\{\tau_k^{(2)}\}$ in the normal domain of attraction of a stable law, i.e.

$$n^{-1/\alpha} \sum_{k=1}^{[nt]} (\tau_k^{(2)} - \lambda_2^{-1}) \;\xrightarrow{fi.di.}\; \mathbf{S}(t) \;,$$

where $\alpha \in (1,2)$, $\lambda_2^{-1} = \mathbb{E}[\tau_1^{(2)}]$ and $\mathbf{S}$ is a totally skewed to the right α-stable Lévy process. By the CLT equivalence, this implies that

$$n^{-1/\alpha}(M_2(nt) - \lambda_2 nt) \;\xrightarrow{fi.di.}\; \mathbf{S}(t) \;.$$

Since the durations of the Poisson process are i.i.d. with finite variance, $n^{-1/2}(M_1(nt) - \lambda_1 nt)$ converges weakly to a Brownian motion. Thus defining the superposition N of these two point processes by

$$N = M_1 + M_2 \ ,$$

we obtain that

$$n^{-1/\alpha}(N(nt) - \lambda nt) \stackrel{fi.di.}{\longrightarrow} -\lambda_2^{1+1/\alpha}\mathbf{S}(t) \ ,$$

with $\lambda = \lambda_1 + \lambda_2$.

Let $\{\tau_k\}$ be the duration sequence of the point process N. We now show that the durations are light tailed under the Palm measure P^0 under which they form a stationary sequence. It follows from the formula (1.4.5) and example 1.4.1 in Baccelli and Brémaud (2003) that

$$P^0(\tau_1 > x) = \lambda^{-1}\mathrm{e}^{-\lambda_1 x}\left\{\lambda_1 \bar{F}_2(x) + \lambda_2 \bar{H}_2(x)\right\} \ ,$$

where F_2 is the distribution function of $\tau_1^{(2)}$ and H_2 is the corresponding delay distribution defined by $H_2(x) = \lambda \int_0^x \bar{F}_2(t)\,\mathrm{d}t$. This yields that $P^0(\tau_1 > x) \leq \mathrm{e}^{-\lambda_1 x}$, i.e. the durations of the superposition process are light tailed.

5.4 Cox processes

Consider now a Cox process N driven by a stationary random measure ξ, which means that conditionally on ξ, N is a Poisson point process with mean measure ξ. Then, denoting $\xi(t) = \xi([0,t])$, we have $\mathbb{E}[N(t)] = \mathbb{E}[\xi(t)] = \lambda t$ with $\lambda = \mathbb{E}[\xi(1)]$ and

$$\mathrm{var}(N(t)) = \mathbb{E}[\xi(t)] + \mathrm{var}(\xi(t)) = \lambda t + \mathrm{var}(\xi(t)) \ .$$

If the stationary random measure ξ has long memory with Hurst index $H > 1/2$ then so has N. We give examples.

(i) Consider the random measure ξ with density $\mathrm{e}^{\sigma Z_H(s)}$ with respect to Lebesgue's measure on $\mathbb{R}$, where $\sigma > 0$ and Z_H is a standard fractional Gaussian noise with Hurst index $H \in (1/2, 1)$, i.e. a Gaussian stationary process with covariance function $\rho_H(t) = \frac{1}{2}\{|t-1|^{2H} - 2|t|^{2H} + |t+1|^{2H}\}$. This means that for any Borel set A, $\xi(A) = \int_A \mathrm{e}^{Z_H(s)}\,\mathrm{d}s$, and in particular $\xi(t) = \int_0^t \mathrm{e}^{Z_H(s)}\,\mathrm{d}s$. Then, by stationarity,

$$\lambda = \mathbb{E}\left[\int_0^1 \mathrm{e}^{\sigma Z_H(s)}\,\mathrm{d}s\right] = \mathbb{E}\left[\mathrm{e}^{\sigma Z_H(0)}\right] = \mathrm{e}^{-\sigma^2/2} \ .$$

Since the function $x \to \mathrm{e}^x$ has Hermite rank 1, it holds that $\mathrm{cov}(\mathrm{e}^{Z_H(0)}, \mathrm{e}^{Z_H(t)}) \sim c\rho_H(t)$ and $\mathrm{var}(\xi(t)) \sim C^2 t^{2H}$ where c and C are positive constants, and $n^{-H}\{\int_0^{nt} \mathrm{e}^{Z_H(s)}\,\mathrm{d}s - \lambda nt\} \xrightarrow{fi.di.} CB_H(t)$. See Arcones (1994) and Dobrushin and Major (1979) for details about Hermite ranks, covariance inequality and convergence in distribution of integrals of functions of Gaussian processes. Applying Lemma 9.1, we obtain the convergence

$$n^{-H}\{N(nt) - \lambda nt\} \xrightarrow{fi.di.} CB_H(t)\ .$$

(ii) Consider the case where the stochastic intensity can be expressed as $\xi(t) = \int_0^t W(s)\,\mathrm{d}s$, where $W(t)$ is an alternating renewal or ON-OFF process (cf.Heath et al. (1998), Daley (2010)), defined by $W(t) = 1$ if $T_{2n} \le t < T_{2n+1}$ and $T_{2n+1} \le t < T_{2n+2}$, where $T_{2n} = \sum_{i=0}^n X_i + Y_i$ and $T_{2n} = \sum_{i=0}^n X_i + Y_i + X_{n+1}$, $\{X_i, i \ge 1\}$ and $\{Y_i, i \ge 1\}$ are two independent i.i.d. sequences of positive random variables with finite mean (the ON and OFF periods) and X_0 and Y_0 are mutually independent and independent of the other random variables, whose distributions are the delay distributions which make the renewal processes stationary. If the ON distribution is regularly varying at infinity with tail index $\alpha \in (1,2)$ i.e. $\mathbb{P}(X_1 > x) \sim cx^{-\alpha}$ as $x \to \infty$ (which implies infinite variance) and if $\mathbb{E}[Y_i^{\alpha+\epsilon}] < \infty$ for some $\epsilon > 0$, then $\mathrm{var}(\xi(t)) \sim Ct^{2H}$ with $H = (3-\alpha)/2$. Moreover, $n^{-1/\alpha}\{\xi(nt) - \lambda nt\} \xrightarrow{fi.di.} \Lambda(t)$, where Λ is a totally skewed to the right α-stable process. Applying Lemma 9.1, we obtain the convergence

$$n^{-1/\alpha}\{N(nt) - \lambda nt\} \xrightarrow{fi.di.} \Lambda(t)\ .$$

Proposition 5.1. *Consider the returns $\{r_n\}_{n\in\mathbb{N}^*}$ driven by a Cox process of case (i), then the conclusion of Theorem 4.1 holds with $a = 2 - 2H$.*

Proof of Proposition 5.1. First note that the causal Gaussian process $Z_H(s)$ is θ-weak dependent with rate $\theta_h = L(h)h^{2H-2}$ by Bardet et al. (2008), where $L(h)$ is a slowly varying function. Then by the Proposition 1 of Jach et al. (2012), the stationary process $w(s) = e^{\sigma Z_H(s)}$ is θ-weak dependent with the same rate. Using Lemma 9.2, we have that the returns $\{r_n\}_{n\in\mathbb{N}^*}$ are θ-weak dependent with the same rate $\theta_h = O(h^{2H-2})$. The cumulative distribution function of

$$\frac{\Lambda(1)}{|\Lambda(1) - 2\Lambda(1/2)|}$$

is continuous, where $\Lambda(1)$ and $\Lambda(1/2)$ have totally skewed to the right α-stable distributions. Finally, by applying Theorem 4.1 with $a = 2 - 2H$, the result is proved. □

6 Simulation

We study the performance of the ordinary t-test and the subsampling-based T test of

$$H_0 : \mu^* = 0 \tag{6.1}$$
$$H_1 : \mu^* > 0 \tag{6.2}$$

We consider exponential, ACD(1,1), and LMSD durations. In each case, we simulate two mutually independent durations process $\{\tau_k^{(i)}\}$ for $i = 1, 2$, and then obtain the corresponding counting processes $\{N_i(t)\}$. Next, we generate mutually independent disturbance series $\{\tilde{e}_k^{(1)}\}$ and $\{\tilde{e}_k^{(2)}\}$ that are i.i.d. Gaussian with means $\mu_1 > 0$, $\mu_2 < 0$, $\mu_1 + \mu_2 > 0$, and with finite variances σ_i^2. For simplicity, we assume all parameters other than μ_i for the two processes to be the same. For the exponential and EACD(1,1) models, we only include the efficient shocks in simulations, and for the LMSD, we also include microstructure shocks, $\{\eta_k^{(i)}\}$, that induce a leverage effect. We then construct the log-price series $\{y(k)\}_{k=1}^n$ and return series $\{r_k\}_{k=1}^n$ from (2.2) and $r_k = y(k) - y(k-1)$ with $y(0) = 0$. We calculate t_n and T_n based on return series and apply subsampling to obtain the empirical quantiles of T_n and thereby carry out the hypothesis test.

We do a preliminary study on the sensitivity of our results to the choice of the block size b. These results show insensitivity except for the largest values of b.

The simulation results show that in all three models, the T test is generally correctly sized. But in the exponential case where the usual t test is also correctly sized, the power of T is lower than that of the t test. To improve the power, we modify the test statistic defined in (3.2) as

$$T_{n,2} = \frac{\bar{r}(n,n,1) - \mu_0^*}{|\bar{r}(n,n/2,1) - \bar{r}(n,n/4,1)| + |\bar{r}(n,n/2,n/2+1) - \bar{r}(n,n/4,n/2+1)|}, \tag{6.3}$$

where $\bar{r}(n,b,t) = \frac{1}{b}\sum_{k=t}^{t+b-1} r_k$. In the remainder of this section, we will use T_1 to denote (3.2), T_2 to denote (6.3), and "T test" to denote a test based on either T_1 or T_2.

6.1 Exponential durations

We generate $\{\tau_k^{(1)}\}$, $\{\tau_k^{(2)}\}$ from exponential distributions $\tau_k^{(i)} \overset{i.i.d.}{\sim} Exp(\lambda_i)$, which result in $N_1(t)$ and $N_2(t)$ being Poisson processes. Since the returns are i.i.d, the

t-test is correctly sized. Our goal is to study the size of the T test and to compare the power of the T to that of the t test.

6.1.1 Parameter calibration

We consider a total of 30 configurations of parameters and sample size, 15 for evaluating size (Table 2), and 15 for evaluating power (Table 3). The model has six parameters (λ_1, λ_2, σ_1, σ_2, μ_1, μ_2). Note that we assume that ($\lambda_1 = \lambda_2 = \lambda$, $\sigma_1 = \sigma_2 = \sigma$). For the first configuration corresponding to power (Table 3), we calibrate the model to match the mean and standard deviation of the returns to those of the Fama/French factor, Rm-Rf (see Section 7), and the intensity observed in the Boeing series used in Deo et al. (2010).

We consider one unit of time to represent 5 minutes, whereas Deo et al. (2010) used 1 minute and Rm-Rf is a daily series. Our calibrated model (which accounts for the discrepancies in the time units) is given in the first row of Table 3. We compute the parameter values using the fact that for the Poisson process, there is an explicit functional relationship between the model parameters and the mean and variance of the calendar-time returns.

Figure 1, for one simulated realization from this model with $n = 15000$, shows the time series plots of returns and log price, as well as the ACF and PACF of the returns. Neither ACF nor PACF shows statistically significant lags.

For the remaining configurations in Table 3, we increase μ^* by varying λ and μ_i. We also consider the two sample sizes, $n = 5000$ (corresponding to roughly one quarter of 5-minute observations) and $n = 10000$. The power should increase as μ^* increases. Figure 2 shows the time series plots of the log price for the models in rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) in Table 3.

For the first configuration corresponding to size (Table 2), we only match the standard deviation of the returns to that of Rm-Rf and the intensity of Boeing series, while setting the mean to zero. For the remaining configurations in Table 2, we vary the λ and μ_i similarly as in Table 3 while keeping $\mu^* = 0$.

For each configuration in Tables 2 and 3, we generate 1000 realizations. For each realization, we calculate the t and T statistics, and use subsampling to obtain the rejection region of T for the hypothesis test (6.1) (6.2) at the 0.05 significance level,

$$\{T > c_{n,b}(0.95)\}$$

where $c_{n,b}(p)$ is the pth quantile of the subsampling distribution as defined in Section 4.

6.1.2 Block size

For the two calibrated models, we evaluate the size and power of T_1 with different block sizes b, sample size $n = 15000$ (corresponding to roughly three quarters of observations), and 1000 realizations. We also evaluate the size and power. Results are given in Table 1. When $b < 1600$, the size of T_1 is not significantly different from 0.05 using the binomial test at the 0.05 significance level. When b becomes larger, the size is significantly larger than 0.05. The size and power of the T_1 test are insensitive to the block size except at the largest values of b. In view of this insensitivity, and in view of the fact that in Theorem 4.1 b may approach ∞ arbitrarily slowly, we will fix the block size $b = 160$ in the following simulations. (Other values of b were tried there and gave similar results.)

	T_1 **test**								t **test**
Block size (b)	160	320	480	640	800	1120	1600	3200	
Size	0.053	0.052	0.053	0.053	0.053	0.054	0.065	0.077	0.053
Power	0.071	0.07	0.071	0.07	0.078	0.082	0.090	0.110	0.089

Table 1: Size and power of T_1 with different block sizes, $n = 15000$, exponential durations, using the calibrated models (from first row of Tables 2 and 3).

6.1.3 Size/Power of the t and T tests for exponential durations

Table 2 shows that both t and T are correctly sized. Indeed, none of the sizes reported in Table 2 are significantly different from 0.05 according to the binomial test at the 0.05 significance level. Table 3 shows that in general the power of T is lower than that of t, and power of all tests increases as μ^* increases. Furthermore, as n is increased holding all model parameters fixed, the power of all tests increases. Finally, the power of the T tests is lower than that of the t test, but T_2 has higher power than T_1.

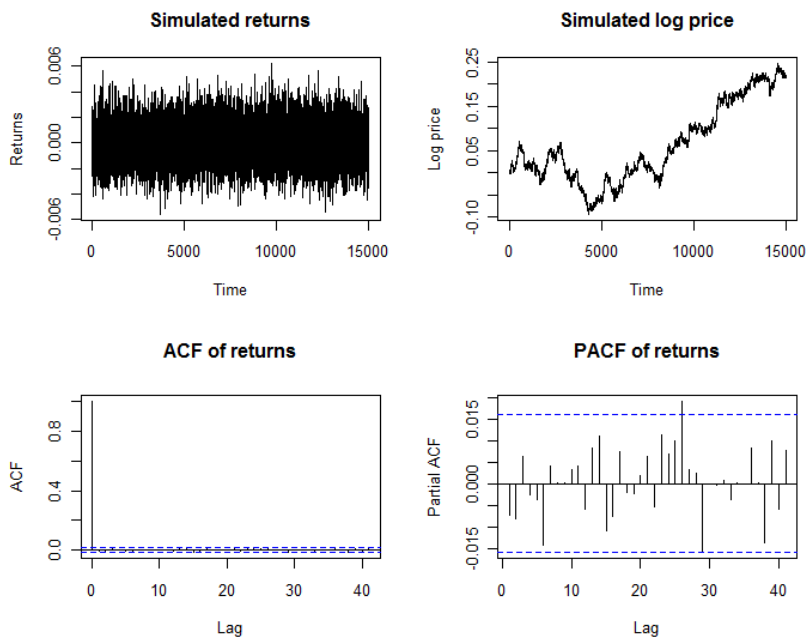

Figure 1: One simulated realization of the calibrated Poisson model (first row in Table 3). Time series plot of returns and log price, as well as ACF and PACF, $n = 15000$. The model is calibrated using Boeing's intensity and Fama/French factor Rm-Rf, such that $\lambda = 2.71$, $\sigma = 0.0003$, and the mean and standard deviation for the 5-minute returns are $\mu^* = 3.846 \times 10^{-6}$, $\tilde{\sigma} = 0.0012$.

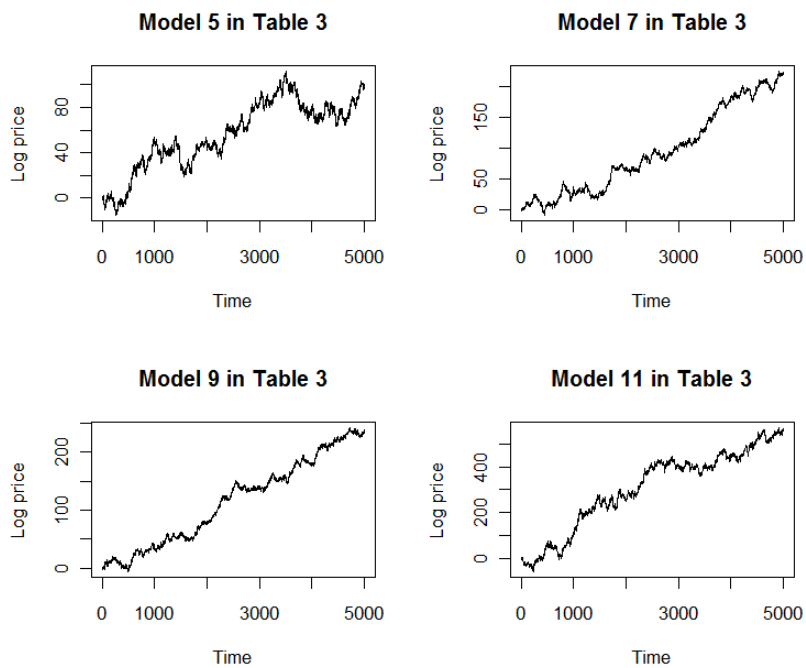

Figure 2: Simulated log price for models in rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) of Table 3, n=5000.

Parameter values					Size		
n	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2	t
5000	2.71	4.266×10^{-4}	-4.266×10^{-4}	0	0.055	0.063	0.045
5000	2.71	4.173×10^{-4}	-4.173×10^{-4}	0	0.057	0.050	0.048
5000	2.71	4.359×10^{-4}	-4.359×10^{-4}	0	0.048	0.056	0.048
5000	27.1	4.266×10^{-4}	-4.266×10^{-4}	0	0.055	0.054	0.052
5000	2.71	0.4266	-0.4266	0	0.043	0.054	0.049
5000	2.71	0.4173	-0.4173	0	0.047	0.055	0.049
5000	2.71	0.3273	-0.3273	0	0.058	0.056	0.053
5000	2.71	0.4359	-0.4359	0	0.061	0.057	0.063
5000	2.71	0.5259	-0.5259	0	0.060	0.058	0.053
5000	27.1	0.4266	-0.4266	0	0.058	0.057	0.046
5000	271	0.4266	-0.4266	0	0.053	0.051	0.058
10000	2.71	0.4173	-0.4173	0	0.050	0.064	0.056
10000	2.71	0.3273	-0.3273	0	0.052	0.051	0.050
10000	2.71	0.4359	-0.4359	0	0.058	0.056	0.057
10000	2.71	0.5259	-0.5259	0	0.060	0.058	0.056

Table 2: Size of the t and T tests for i.i.d. exponential durations.

Parameter values					Power		
n	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2	t
5000	2.71	4.273×10^{-4}	-4.259×10^{-4}	3.846×10^{-6}	0.066	0.077	0.068
5000	2.71	4.273×10^{-4}	-4.173×10^{-4}	3.846×10^{-5}	0.237	0.363	0.575
5000	2.71	4.359×10^{-4}	-4.259×10^{-4}	3.846×10^{-5}	0.242	0.242	0.582
5000	27.1	4.273×10^{-4}	-4.259×10^{-4}	3.846×10^{-5}	0.103	0.105	0.147
5000	2.71	0.4273	-0.4259	3.846×10^{-3}	0.068	0.089	0.090
5000	2.71	0.4273	-0.4173	3.846×10^{-2}	0.361	0.615	0.878
5000	2.71	0.4273	-0.3273	3.846×10^{-1}	1	1	1
5000	2.71	0.4359	-0.4259	3.846×10^{-2}	0.359	0.593	0.853
5000	2.71	0.5259	-0.4259	3.846×10^{-1}	1	1	1
5000	27.1	0.4273	-0.4259	3.846×10^{-2}	0.151	0.204	0.222
5000	271	0.4273	-0.4259	3.846×10^{-1}	0.357	0.596	0.880
10000	2.71	0.4273	-0.4173	3.846×10^{-2}	0.492	0.788	0.993
10000	2.71	0.4273	-0.3273	3.846×10^{-1}	1	1	1
10000	2.71	0.4359	-0.4259	3.846×10^{-2}	0.461	0.790	0.991
10000	2.71	0.5259	-0.4259	3.846×10^{-1}	1	1	1

Table 3: Power of the t and T tests for i.i.d. exponential durations.

6.2 ACD durations

We generate $\{\tau_k^{(i)}\}$ from the exponential ACD(1,1) model (EACD(1,1)) that satisfies (5.1) with $\{\epsilon_k^{(i)}\} \overset{i.i.d.}{\sim} exp(1)$, $i = 1, 2$. We generate models with finite and infinite variance for the durations. For the finite variance model, we try two block sizes $b = 160$ and 320. For the EACD(1,1), t is oversized and therefore we will not examine its power. We will also study the size and power of the T tests.

6.2.1 Parameter calibration

Similarly as in Section 6.1, for both finite and infinite variance EACD models, we consider a total of 30 configurations of parameters and sample size, 15 for evaluating size (Table 5 for the finite variance model, Table 7 for the infinite variance model), and 15 for evaluating power (Table 6 for the finite variance model, Table 8 for the infinite variance model). The model has six parameters $(\alpha, \beta, \omega, \sigma, \mu_1, \mu_2)$. Note that we assume that $\alpha, \beta, \omega, \sigma$ are the same for both processes $\{\tau_k^{(1)}\}$, $\{\tau_k^{(2)}\}$.

EACD(1,1) with finite variance

The necessary and sufficient condition for finite variance of the EACD(1,1) model is

$$2\alpha^2 + \beta^2 + 2\alpha\beta = \alpha^2 + (\alpha + \beta)^2 < 1 \ , \tag{6.4}$$

For the first configuration corresponding to power (Table 6), we still calibrate the model to match the mean and standard deviation of the returns to those of the Fama/French factor, Rm-Rf, and the (α, β, ω) observed in the Boeing series used in Deo et al. (2010), which satisfy (6.4). We use the same σ as for the exponential durations and set μ_1 and μ_2 to match the mean and standard deviation of the Rm-Rf. This calibrated model also accounts for the discrepancies in the time units, and is given in the first row of Table 6. For the EACD(1,1) model, there is no explicit functional relationship between the model parameters and the mean and variance of the calendar-time returns, since these now also depend on the variance of counts (the number of events occurring in a given 5-minute time interval). We estimate the variance of counts using simulation and set μ_1 and μ_2 accordingly.

Figure 3, for one simulated realization from this model with $n = 5000$, shows the time series plots of returns and log price, as well as the ACF and PACF of the returns. Both ACF and PACF show statistically significant autocorrelations at several lags.

For the remaining configurations in Table 6, we increase μ^* by varying ω (which changes λ) and μ_i. We also consider the two sample sizes, $n = 5000$ and $n = 10000$. The power should generally increase as μ^* increases. Figure 4 shows the time series plots of the log price for the models of rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) of Table 6.

For the first configuration corresponding to size (Table 5), similarly as in Section 6.1.1, we set the mean to zero. For the remaining configurations in these two tables, we vary the ω and μ_i similarly as in Table 6 while keeping $\mu^* = 0$.

For each configuration of size and power, we generate 1000 realizations. The simulation procedure is the same as for the exponential durations.

EACD(1,1) with infinite variance

The necessary and sufficient condition for the EACD(1,1) model to have infinite variance is

$$\alpha, \beta \geq 0, \quad \alpha + \beta < 1, \quad \alpha^2 + (\alpha + \beta)^2 \geq 1 \ .$$

We fix $\alpha + \beta = 0.99$ which is the same as for all of the finite variance models. Table 4 shows the (α, β) values that satisfy the above constraint while attaining the values of $\alpha^2+(\alpha+\beta)^2$ given in the top row. Since for the ACD model, estimates of β are typically close to 1, we choose the first pair $\alpha = 0.1729$, $\beta = 0.817$. We adjust ω to obtain the same values of λ as used in the finite variance EACD(1,1) models. The value of σ is the same as before. Then following the same procedure as described in the finite variance model, we obtain the size (Table 7) and power (Table 8).

Figure 5, for one simulated realization from this model with $n = 5000$, shows the time series plots of returns and log price, as well as the ACF and PACF of the returns. Figure 6 shows the time series plots of the log price for the models of rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) of Table 8.

We obtained similar results not shown here for other configurations of (α, β, ω) both with and without the constraint $\alpha + \beta = 0.99$.

$\alpha^2 + (\alpha + \beta)^2$	1.01	1.02	1.04	1.06	1.08	1.10
α	0.1729	0.1997	0.2447	0.28267	0.3161	0.3463
β	0.8171	0.7903	0.7453	0.7073	0.6739	0.6437

Table 4: (α, β) values such that $\alpha + \beta = 0.99$ and $\alpha^2 + (\alpha + \beta)^2$ is given in the top row.

6.2.2 Size/Power of the t and T tests for EACD(1,1) durations

The size for both the finite variance (Table 5) and infinite variance models (Table 7) show that for EACD(1,1) durations, the t test is over-sized and not reliable, especially in the infinite variance cases. Thus, we will not examine the power of the t test here.

For the finite variance cases, the T test is correctly sized. Table 6 indicates that the power of T has the same properties as in the Poisson model, i.e., the power increases as μ^* or n increases; T_2 generally has higher power than T_1.

For the infinite variance cases, Table 7 for size shows that both T_1 and T_2 are under-sized in some cases and are approximately correctly sized in the other cases. Table 8 shows that the power of both T_1 and T_2 increases as μ^* or n increases, but perhaps not as fast as for the finite variance models. T_2 has higher power than T_1, particularly when μ^* is large.

The results of Tables 5 and 6 show no apparent sensitivity to the block size.

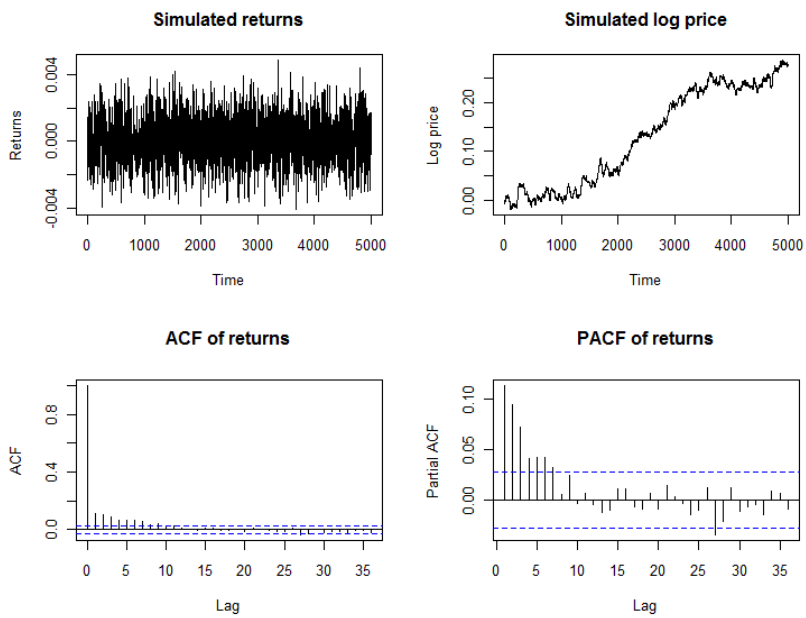

Figure 3: One simulated realization of the calibrated EACD(1,1) model with finite variance (first row in Table 6). Time series plot of returns and log price, as well as ACF and PACF, $n = 5000$. The model is calibrated using Boeing's and Fama/French factor Rm-Rf, such that $\alpha = 0.023161$ $\beta = 0.970158$, $\omega = 0.00247$, $\sigma = 0.0003$, and the mean and standard deviation for the 5-minute returns are $\mu^* = 3.846 \times 10^{-6}$, $\tilde{\sigma} = 0.0012$.

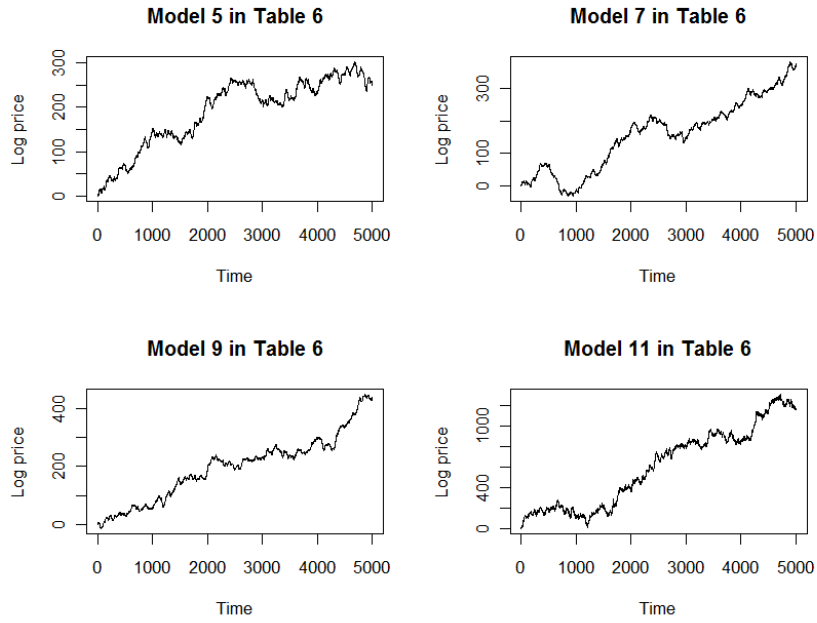

Figure 4: Simulated log price for models in rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) of Table 6, n=5000.

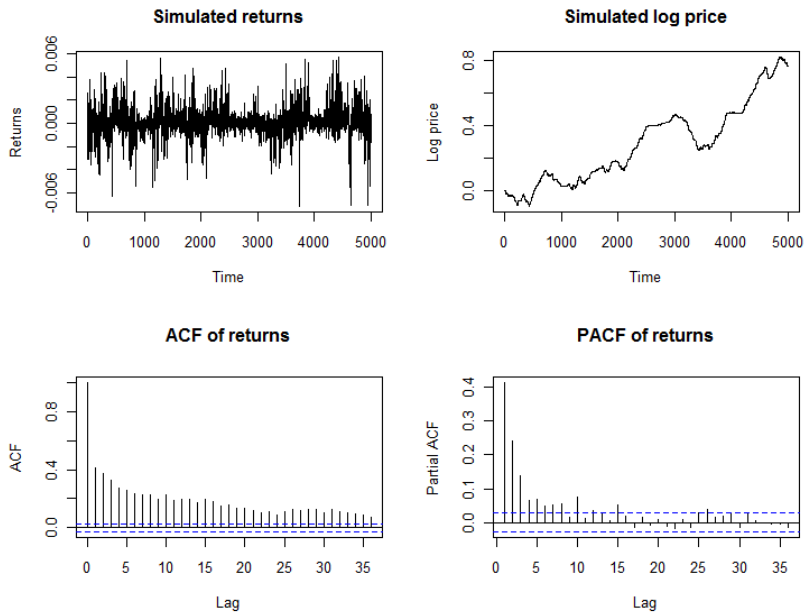

Figure 5: One simulated realization of the calibrated EACD(1,1) model with infinite variance (first row in Table 8). Time series plot of returns and log price, as well as ACF and PACF, $n = 5000$. The model is calibrated using Boeing's and Fama/French factor Rm-Rf, such that $\alpha = 0.1729$, $\beta = 0.8171$, $\omega = 0.00369$, $\sigma = 0.0003$, and the mean and standard deviation for the 5-minute returns are $\mu^* = 3.846 \times 10^{-6}$, $\tilde{\sigma} = 0.0012$.

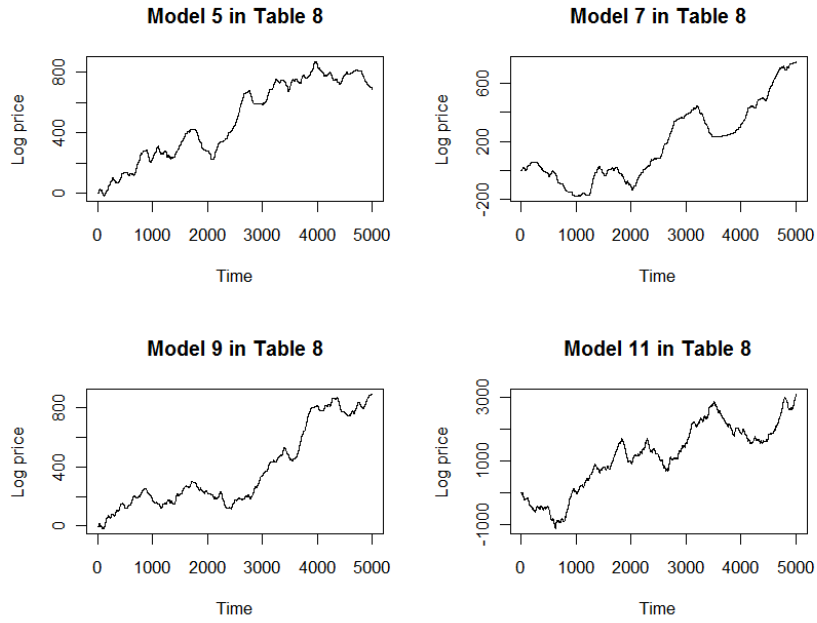

Figure 6: Simulated log price for models in rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) of Table 8, n=5000.

Parameter values						Size ($b = 160$)		Size ($b = 320$)		Size
n	ω	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2	T_1	T_2	t
5000	0.00247	2.71	3.380×10^{-4}	-3.380×10^{-4}	0	0.043	0.043	0.046	0.054	0.165
5000	0.00247	2.71	3.245×10^{-4}	-3.245×10^{-4}	0	0.058	0.045	0.057	0.047	0.185
5000	0.00247	2.71	3.515×10^{-4}	-3.515×10^{-4}	0	0.043	0.045	0.046	0.048	0.175
5000	0.000247	27.1	3.380×10^{-4}	-3.380×10^{-4}	0	0.058	0.057	0.056	0.056	0.171
5000	0.00247	2.71	0.3380	-0.3380	0	0.056	0.057	0.062	0.061	0.222
5000	0.00247	2.71	0.3245	-0.3245	0	0.055	0.045	0.057	0.057	0.232
5000	0.00247	2.71	0.1966	-0.2023	0	0.059	0.049	0.054	0.049	0.248
5000	0.00247	2.71	0.3515	-0.3515	0	0.057	0.052	0.060	0.064	0.250
5000	0.00247	2.71	0.4793	-0.4793	0	0.051	0.053	0.056	0.062	0.228
5000	0.000247	27.1	0.3380	-0.3380	0	0.054	0.050	0.059	0.046	0.233
5000	0.0000247	271	0.3380	-0.3380	0	0.050	0.054	0.050	0.065	0.226
10000	0.00247	2.71	0.3245	-0.3245	0	0.044	0.046	0.047	0.050	0.209
10000	0.00247	2.71	0.1966	-0.1966	0	0.048	0.050	0.043	0.050	0.213
10000	0.00247	2.71	0.3515	-0.3515	0	0.052	0.047	0.060	0.044	0.243
10000	0.00247	2.71	0.4793	-0.4793	0	0.052	0.048	0.055	0.045	0.235

Table 5: Size of the t and T tests for EACD(1,1) durations with finite variance.

Parameter values						Power ($b = 160$)		Power ($b = 320$)	
n	ω	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2	T_1	T_2
5000	0.00247	2.71	3.387×10^{-4}	-3.373×10^{-4}	3.846×10^{-6}	0.043	0.054	0.052	0.063
5000	0.00247	2.71	3.387×10^{-4}	-3.245×10^{-4}	3.846×10^{-5}	0.175	0.196	0.187	0.203
5000	0.00247	2.71	3.515×10^{-4}	-3.373×10^{-4}	3.846×10^{-5}	0.145	0.193	0.155	0.205
5000	0.000247	27.1	3.387×10^{-4}	-3.373×10^{-4}	3.846×10^{-5}	0.093	0.096	0.095	0.098
5000	0.00247	2.71	0.3387	-0.3373	3.846×10^{-3}	0.068	0.066	0.067	0.070
5000	0.00247	2.71	0.3387	-0.3245	3.846×10^{-2}	0.176	0.204	0.178	0.211
5000	0.00247	2.71	0.3387	-0.1966	3.846×10^{-1}	0.964	1	0.954	0.999
5000	0.00247	2.71	0.3515	-0.3373	3.846×10^{-2}	0.173	0.219	0.171	0.216
5000	0.00247	2.71	0.4793	-0.3373	3.846×10^{-1}	0.906	0.999	0.918	0.999
5000	0.000247	27.1	0.3387	-0.3373	3.846×10^{-2}	0.088	0.089	0.094	0.091
5000	0.0000247	271	0.3387	-0.3373	3.846×10^{-1}	0.198	0.282	0.190	0.276
10000	0.00247	2.71	0.3387	-0.3245	3.846×10^{-2}	0.250	0.311	0.253	0.328
10000	0.00247	2.71	0.3387	-0.1966	3.846×10^{-1}	0.997	1	0.992	1
10000	0.00247	2.71	0.3515	-0.3373	3.846×10^{-2}	0.230	0.302	0.243	0.335
10000	0.00247	2.71	0.4793	-0.3373	3.846×10^{-1}	0.972	1	0.962	1

Table 6: Power of the T tests for EACD(1,1) durations with finite variance.

Parameter values						Size		
n	ω	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2	t
5000	0.00369	2.71	1.848×10^{-4}	-1.848×10^{-4}	0	0.040	0.041	0.393
5000	0.00369	2.71	1.713×10^{-4}	-1.713×10^{-4}	0	0.043	0.044	0.409
5000	0.00369	2.71	1.983×10^{-4}	-1.983×10^{-4}	0	0.041	0.040	0.439
5000	0.000369	27.1	1.848×10^{-4}	-1.848×10^{-4}	0	0.042	0.041	0.372
5000	0.00369	2.71	0.1848	-0.1848	0	0.041	0.042	0.448
5000	0.00369	2.71	0.1713	-0.1713	0	0.043	0.040	0.417
5000	0.00369	2.71	0.0435	-0.0435	0	0.060	0.054	0.571
5000	0.00369	2.71	0.1983	-0.1983	0	0.046	0.043	0.448
5000	0.00369	2.71	0.3261	-0.3261	0	0.053	0.059	0.506
5000	0.000369	27.1	0.1848	-0.1848	0	0.043	0.041	0.385
5000	0.0000369	271	0.1848	-0.1848	0	0.045	0.045	0.379
10000	0.00369	2.71	0.1713	-0.1713	0	0.045	0.040	0.435
10000	0.00369	2.71	0.0435	-0.0435	0	0.055	0.047	0.536
10000	0.00369	2.71	0.1983	-0.1983	0	0.043	0.042	0.453
10000	0.00369	2.71	0.3261	-0.3261	0	0.046	0.051	0.517

Table 7: Size of the t and T tests for EACD(1,1) durations with infinite variance.

Parameter values						Power	
n	ω	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2
5000	0.00369	2.71	1.855×10^{-4}	-1.841×10^{-4}	3.846×10^{-6}	0.041	0.042
5000	0.00369	2.71	1.855×10^{-4}	-1.713×10^{-4}	3.846×10^{-5}	0.060	0.056
5000	0.00369	2.71	1.983×10^{-4}	-1.841×10^{-4}	3.846×10^{-5}	0.046	0.047
5000	0.000369	27.1	1.855×10^{-4}	-1.841×10^{-4}	3.846×10^{-5}	0.042	0.043
5000	0.00369	2.71	0.1855	-0.1841	3.846×10^{-3}	0.046	0.044
5000	0.00369	2.71	0.1855	-0.1713	3.846×10^{-2}	0.047	0.046
5000	0.00369	2.71	0.1855	-0.0435	3.846×10^{-1}	0.408	0.622
5000	0.00369	2.71	0.1983	-0.1841	3.846×10^{-2}	0.056	0.058
5000	0.00369	2.71	0.3261	-0.1841	3.846×10^{-1}	0.244	0.299
5000	0.000369	27.1	0.1855	-0.1841	3.846×10^{-2}	0.044	0.045
5000	0.0000369	271	0.1855	-0.1841	3.846×10^{-1}	0.057	0.059
10000	0.00369	2.71	0.1855	-0.1713	3.846×10^{-2}	0.043	0.046
10000	0.00369	2.71	0.1855	-0.0435	3.846×10^{-1}	0.453	0.682
10000	0.00369	2.71	0.1983	-0.1841	3.846×10^{-2}	0.050	0.052
10000	0.00369	2.71	0.3261	-0.1841	3.846×10^{-1}	0.258	0.358

Table 8: Power of the T tests for EACD(1,1) durations with infinite variance.

6.3 LMSD durations

We generate $\{\tau_k^{(i)}\}$ from the LMSD model

$$\tau_k = \epsilon_k e^{Y_k}$$

where $\{\epsilon_k\} \overset{i.i.d.}{\sim} Weibull(\delta, \gamma)$, and $(\delta,\ \gamma)$ are the scale and shape parameter; $\{Y_k\}$ is a Gaussian $ARFIMA(1, d, 0)$ process with innovations $w_k \overset{i.i.d.}{\sim} N(0, \sigma_w^2)$, given by

$$(1 - \alpha L)(1 - L)^d Y_k = w_k$$

with $|\alpha| < 1$, and d is the long memory parameter with $d = H - 1/2 \in (0, 1/2)$. We generate models without and with microstructure shocks $\{\eta_k\}$. For all of these models, we find that t is oversized and therefore we will not examine its power. We will also study the size and power of the T tests and the leverage effect induced by the microstructure shocks. In all simulations, we use 1000 realizations and block size $b = 160$.

6.3.1 Parameter calibration

For both LMSD models without and with $\{\eta_k\}$, we consider a total of 30 configurations of parameters and sample size, 15 for evaluating size (Table 10), and 15 for evaluating power (Table 11). Parameters for the LMSD model are (α, δ, γ, σ_w^2, d, σ, μ_1, μ_2). Note that we assume that (α, δ, γ, σ_w^2, d, σ) are the same for both processes $\{\tau_k^{(1)}\}$, $\{\tau_k^{(2)}\}$.

LMSD without microstructure shocks

For the first configuration corresponding to power (Table 11), we calibrate the model to match the $(\alpha, \delta, \gamma, \sigma_w^2, d)$ observed in the Boeing series used in Deo et al. (2010). We use the same σ as for the exponential durations and set μ_1 and μ_2 to match the mean and standard deviation of the Rm-Rf. This calibrated model also accounts for the discrepancies in the time units, and is given in the first row of Table 11. Similarly as for the EACD(1,1) model, for the LMSD model, the functional relationship between the model parameters and the mean and variance of the calendar-time returns also depend on the variance of counts. We estimate the variance of counts using simulation and set μ_1 and μ_2 accordingly. Figure 7, for one simulated realization from this model with $n = 5000$, shows the time series plots of returns and log price, as well as the ACF and PACF of the returns.

For the remaining configurations in Table 11, we increase μ^* by varying δ (which changes λ) and μ_i. We consider two sample sizes, $n = 5000$ and $n = 10000$. The power should generally increase as μ^* increases.

For the first configuration corresponding to size (Table 10), we set the mean to zero. For the remaining configurations in these two tables, we vary the δ and μ_i similarly as in Table 11 while keeping $\mu^* = 0$.

LMSD with microstructure shocks

For $i = 1, 2$, define the microstructure shock η_k as

$$\eta_k = Y_{k-1} - Y_k \ .$$

With $\{\eta_k\}$, the functional relationship between the model parameters and the variance of the calendar-time returns now depend on not only the variance of counts, but also the variance of $\{\eta_k\}$ and the covariance of counts and $\{\eta_k\}$. Since it is difficult to calibrate this variance and covariance, for simplicity, we will use the same values of all parameters (α, δ, γ, σ_w^2, d, σ, μ_1, μ_2) as those used in the LMSD models without $\{\eta_k\}$. The size and power results are given in Tables 10 (size) and 11 (power).

Figure 8, for one simulated realization from first configuration of Table 11 with $\{\eta_k\}$, shows the time series plots of returns and log price, as well as the ACF and PACF of the returns. Figure 9 shows the time series plots of the log price for the models with $\{\eta_k\}$ of rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) in Table 11.

6.3.2 Leverage effect

In model (2.1), we assume that the microstructure shocks $\{\eta_{i,k}\}$ are independent of the efficient shocks $\{e_{1,k}\}$, $\{e_{2,k}\}$, but not necessarily of the counting process (N_1, N_2). This allows for a leverage effect. See Aue et al. (2014). The leverage is then defined as the correlation between current calendar-time return r_k and absolute value of the next calendar-time return $|r_{k+1}|$

$$leverage = corr(r_k, |r_{k+1}|) \ .$$

We test the model with different memory parameter $d = .23$ and $d = .4$, keeping other parameters the same as in the first configuration of Table 11. We generate 100 realizations, with $n = 5000$, calculate sample leverage for each realization, and use the t-test for the mean,

$$\begin{aligned} H_0 &: \ leverage = 0 \\ H_1 &: \ leverage < 0 \ . \end{aligned}$$

Table 9 shows the mean of sample leverage and the p-value for models with different memory parameters d. Without the $\{\eta_k\}$, there is no leverage in returns, while including the $\{\eta_k\}$ can induce a leverage effect, i.e. negative correlation between current return and the next period's absolute return.

Model	$d = 0.23$		$d = 0.4$	
	Mean	**p-value**	**Mean**	**p-value**
Returns without η_k	0.00062	0.6265	-0.00374	0.3727
Returns with η_k	-0.02299	0.0000	-0.01883	0.0000

Table 9: Means and p-values of sample leverage for models with $d = .23$ and $d = .4$.

6.3.3 Size/Power of the t and T tests for LMSD durations

Tables 10 and 11 show that there is no significant difference in the size and power between models without and with microstructure shocks $\{\eta_k\}$. For both models, Table 10 shows that the t test is oversized. T_1 is correctly sized in all cases, while T_2 is a bit oversized. As we increase λ or n, the size of T_2 becomes smaller. Table 11 shows that the power of the T test performs similarly as in the Poisson model: the power increases as μ^* increases; as n is increased holding all model parameters fixed, the power increases; T_2 has higher power than T_1.

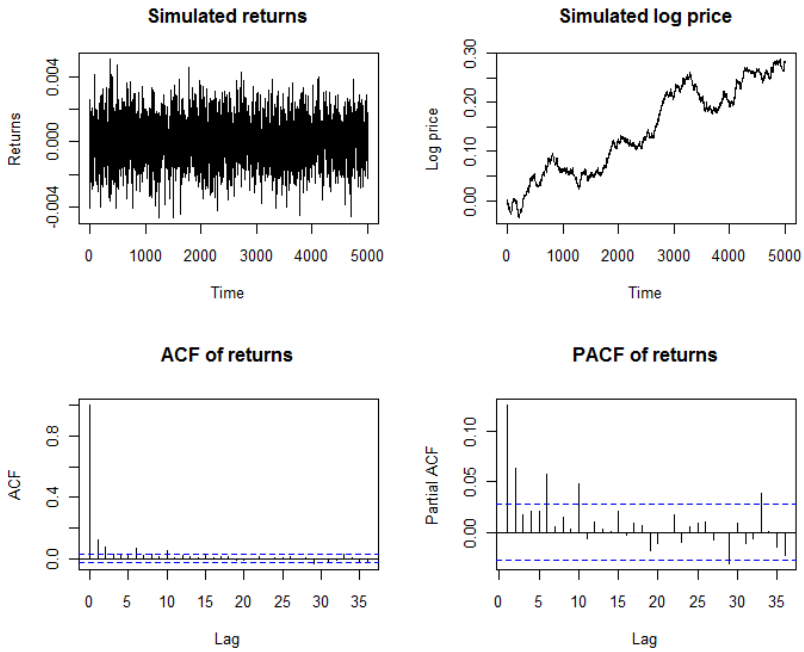

Figure 7: One simulated realization of the calibrated LMSD model without $\{\eta_k\}$ (first row in Table 11). Time series plot of returns and log price, as well as ACF and PACF, n=5000. The model is calibrated using Boeing's and Fama/French factor Rm-Rf, such that $\alpha = -.4212$, $\delta = 0.2$, $\gamma = 1.3376$, $\sigma_w^2 = 0.2368$, $d = 0.3545$, $\sigma = 0.0003$, and the mean and standard deviation for the 5-minute returns are $\mu^* = 3.846 \times 10^{-6}$, $\tilde{\sigma} = 0.0012$.

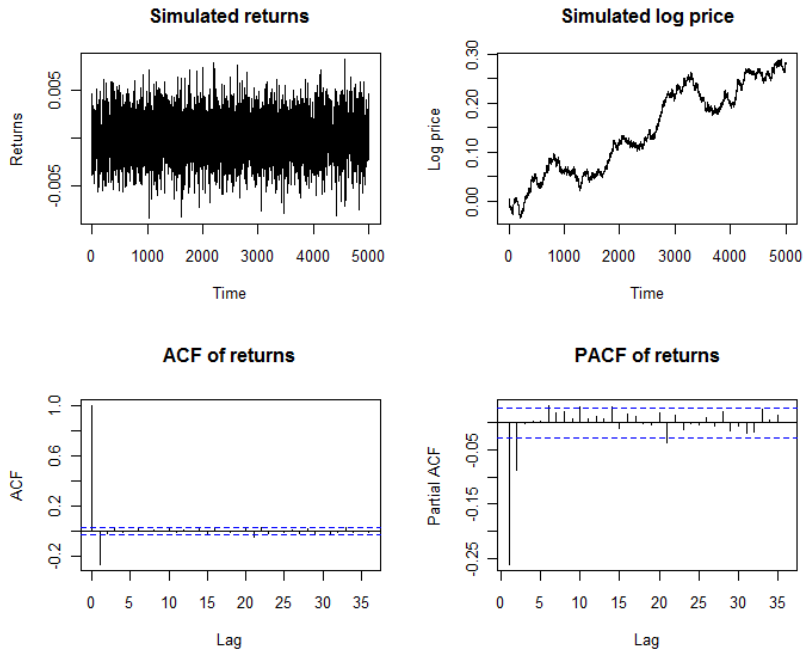

Figure 8: One simulated realization of the calibrated LMSD model with $\{\eta_k\}$ (first row in Table 11). Time series plot of returns and log price, as well as ACF and PACF, n=5000. The model is calibrated using Boeing's and Fama/French factor Rm-Rf, such that $\alpha = -.4212$, $\delta = 0.2$, $\gamma = 1.3376$, $\sigma_w^2 = 0.2368$, $d = 0.3545$, $\sigma = 0.0003$, and the mean and standard deviation for the 5-minute returns are $\mu^* = 3.846 \times 10^{-6}$, $\tilde{\sigma} = 0.0012$.

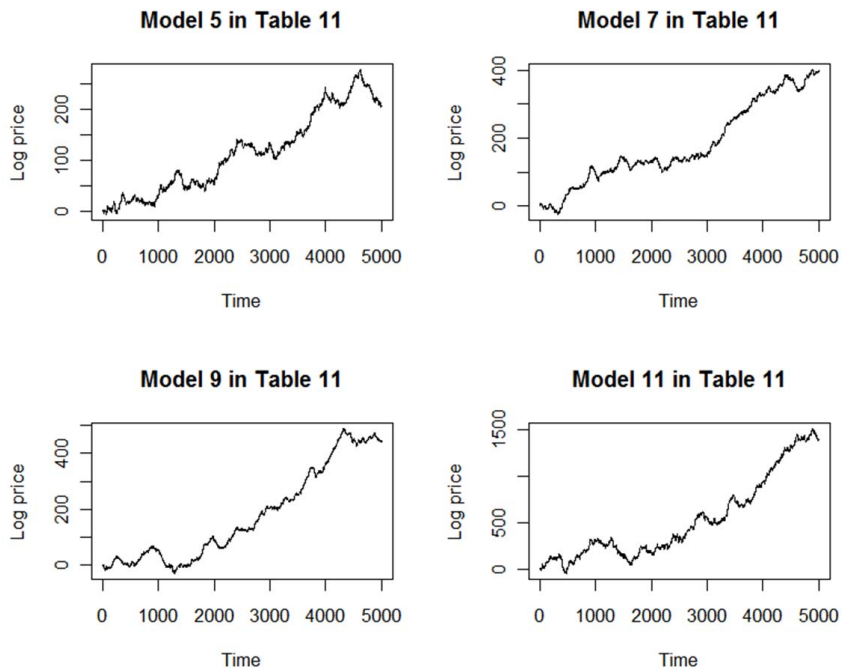

Figure 9: Simulated log price for models with $\{\eta_k\}$ in rows 5 ($\mu^* = 3.846 \times 10^{-3}$), 7, 9, and 11 (all with $\mu^* = 3.846 \times 10^{-1}$) of Table 11.

Parameter values						Size (without η_k)			Size (with η_k)		
n	δ	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2	t	T_1	T_2	t
5000	0.2	4.58	2.293×10^{-4}	-2.293×10^{-4}	0	0.057	0.071	0.307	0.051	0.080	0.307
5000	0.2	4.58	2.213×10^{-4}	-2.213×10^{-4}	0	0.059	0.080	0.312	0.057	0.084	0.311
5000	0.2	4.58	2.373×10^{-4}	-2.373×10^{-4}	0	0.052	0.064	0.291	0.061	0.079	0.320
5000	0.02	45.8	2.293×10^{-4}	-2.293×10^{-4}	0	0.052	0.083	0.337	0.059	0.072	0.314
5000	0.2	4.58	0.2293	-0.2293	0	0.058	0.073	0.341	0.061	0.063	0.345
5000	0.2	4.58	0.2213	-0.2213	0	0.063	0.088	0.352	0.057	0.089	0.351
5000	0.2	4.58	0.1458	-0.1458	0	0.046	0.073	0.376	0.044	0.081	0.341
5000	0.2	4.58	0.2373	-0.2373	0	0.057	0.078	0.366	0.053	0.076	0.339
5000	0.2	4.58	0.3129	-0.3129	0	0.056	0.077	0.343	0.062	0.077	0.372
5000	0.02	45.8	0.2293	-0.2293	0	0.052	0.070	0.357	0.054	0.071	0.364
5000	0.002	458	0.2293	-0.2293	0	0.054	0.057	0.371	0.053	0.052	0.389
10000	0.2	4.58	0.2213	-0.2213	0	0.057	0.068	0.366	0.062	0.070	0.368
10000	0.2	4.58	0.1458	-0.1458	0	0.053	0.070	0.368	0.056	0.071	0.370
10000	0.2	4.58	0.2373	-0.2373	0	0.048	0.071	0.368	0.051	0.070	0.376
10000	0.2	4.58	0.3129	-0.3129	0	0.052	0.071	0.373	0.051	0.071	0.374

Table 10: Size of the t and T tests for LMSD durations.

Parameter values						**Power** (without η_k)		**Power** (with η_k)	
n	δ	$\lambda_1 = \lambda_2$	μ_1	μ_2	μ^*	T_1	T_2	T_1	T_2
5000	0.2	4.58	2.297×10^{-4}	-2.289×10^{-4}	3.846×10^{-6}	0.072	0.080	0.071	0.090
5000	0.2	4.58	2.297×10^{-4}	-2.213×10^{-4}	3.846×10^{-5}	0.134	0.202	0.128	0.197
5000	0.2	4.58	2.373×10^{-4}	-2.289×10^{-4}	3.846×10^{-5}	0.116	0.161	0.130	0.204
5000	0.02	45.8	2.297×10^{-4}	-2.289×10^{-4}	3.846×10^{-5}	0.076	0.092	0.075	0.087
5000	0.2	4.58	0.2297	-0.2289	3.846×10^{-3}	0.074	0.08	0.067	0.069
5000	0.2	4.58	0.2297	-0.2213	3.846×10^{-2}	0.139	0.193	0.120	0.192
5000	0.2	4.58	0.2297	-0.1458	3.846×10^{-1}	0.881	0.999	0.905	0.998
5000	0.2	4.58	0.2373	-0.2289	3.846×10^{-2}	0.141	0.193	0.107	0.166
5000	0.2	4.58	0.3129	-0.2289	3.846×10^{-1}	0.697	0.965	0.688	0.974
5000	0.02	45.8	0.2297	-0.2289	3.846×10^{-2}	0.079	0.089	0.062	0.080
5000	0.002	458	0.2297	-0.2289	3.846×10^{-1}	0.091	0.105	0.083	0.098
10000	0.2	4.58	0.2297	-0.2213	3.846×10^{-2}	0.154	0.209	0.144	0.204
10000	0.2	4.58	0.2297	-0.1458	3.846×10^{-1}	0.940	1	0.949	1
10000	0.2	4.58	0.2373	-0.2289	3.846×10^{-2}	0.145	0.205	0.143	0.195
10000	0.2	4.58	0.3129	-0.2289	3.846×10^{-1}	0.793	0.997	0.780	0.990

Table 11: Power of the T tests for LMSD durations.

7 Data analysis

We study the daily returns of the Fama/French factor, Rm-Rf, from Kenneth French's data library. According to the website, "Rm-Rf is the excess return on the market, value-weight return of all CRSP (Center for Research in Security Prices) firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t, good shares and price data at the beginning of t, and good return data for t minus the one-month Treasury bill rate (from Ibbotson Associates)"(http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html). The data ranges from July 1, 1926 to December 31, 2013, a total 23133 daily excess returns with sample mean 0.000285.

Figure 10 shows the time series plot, log price, ACF and PACF of daily excess returns. Both ACF and PACF show statistically significant lags. We perform a t-test based on the usual standard error and the Newey-West standard error, which allows for serial correlation, and also our T_1 test for the following hypothesis:

$$H_0 : \mu^* = 0$$
$$H_1 : \mu^* > 0$$

where μ^* is the expected excess return (equity premium). To conduct the T_1 test, we calculate the T_1 statistic for the entire data set, and use the same subsampling procedure as defined in Section 6 to obtain the empirical quantiles of the limiting distribution of T. As suggested by Jach et al. (2012), with the sample size of $n = 23133$, we consider the block size $b \in \{732, 976, 1302, 1736, 2314, 3086, 4116, 5488\}$. This covers subsample sizes ranging from $3 - 24\%$ of the sample size. The results are given in Table 12.

At the 0.05 significant level, both the ordinary t-test and the Newey-West t-test reject the null hypothesis with very small p-values, 2.59e-05 and 7.40e-05 respectively. This seems to provide extremely strong evidence for an equity premium. However, the much larger p-values for all of the T_1 tests considered indicate far weaker evidence against the null hypothesis. This finding is consistent with the simulation results in Section 6 for the EACD and LMSD durations. The p-values corresponding to the ordinary t-tests may be spuriously low, and the ones corresponding to the T_1 tests may be more reliable.

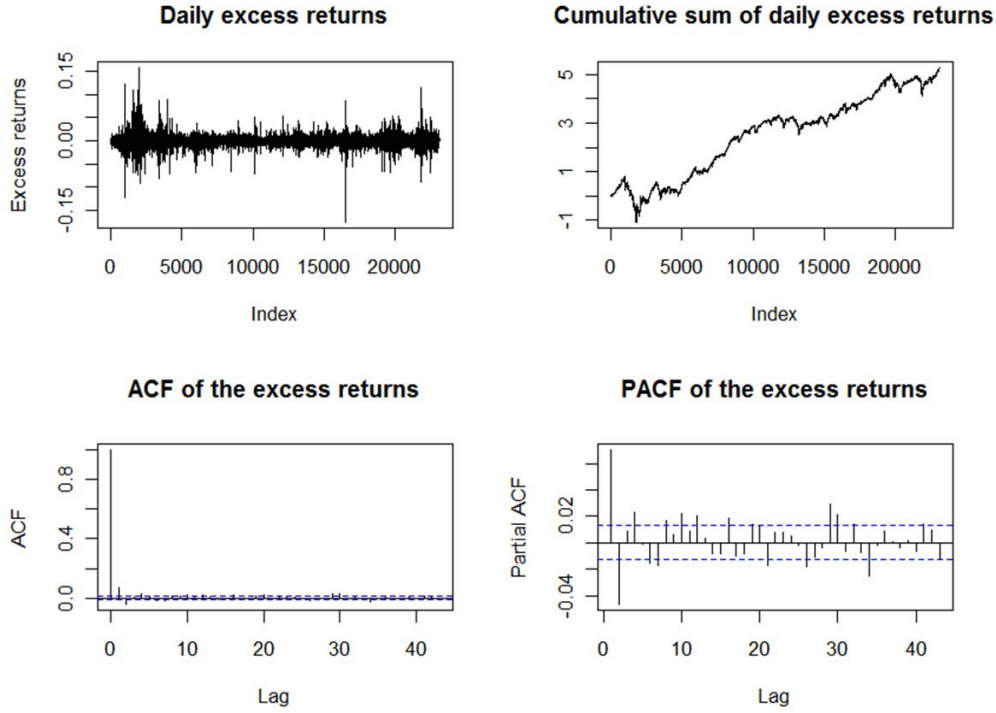

Figure 10: Time series plot, log price, ACF, and PACF of daily excess returns from July 1, 1926 to December 31, 2013.

T test				
Block size (b)	732	977	1302	1736
p**-value**	0.054	0.057	0.056	0.058
Block size (b)	2351	3087	4117	5489
p**-value**	0.064	0.062	0.065	0.057

Table 12: T_1 test results for daily excess returns Rm-Rf.

8 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 $\gamma > 1/2$.

Proposition 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. Therefore, we will now 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 Proposition 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, Whitt (2002) 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 Taqqu and Levy (1986) (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. (2014)), and sample autocorrelations.

9 Proofs

Proof of Proposition 2.1. We start by proving the convergences (2.5) and (2.6). Defining $\zeta_{i,k} = e_{i,k} + \eta_{i,k}$, we can write

$$y(n) - \mu^* n = \mu_1\{N_1(n) - \lambda_1 n\} + \mu_2\{N_2(n) - \lambda_2 n\} + \sum_{k=1}^{N_1(n)} \zeta_{1,k} + \sum_{k=1}^{N_2(n)} \zeta_{2,k}$$

Assumption 2.1 implies that $N_i(n)/n \xrightarrow{\mathbb{P}} \lambda_i$. For $t \geq 0$, define $x_{i,n}(t) = n^{-1/2} \sum_{k=1}^{N_i(nt)} \zeta_{i,k}$. The present assumptions imply assumptions 2.1, 2.2, 2.3 and 2.4 of Aue et al. (2014), thus Theorem 3.1 therein implies that 0 $x_{i,n} \xrightarrow{fi.di.} \sqrt{\lambda}\sigma_i B_i$, where B_i are mutually independent standard Brownian motions. Since the sequences $\{e_{i,k}\}$ are independent of the point processes N_i, the sequences of processes $x_{i,n}$ and $\{n^{-\gamma}(N_i(nt) - n\lambda_i t), t \geq 0\}$, $i = 1, 2$ converge jointly. Thus, if $\gamma > 1/2$, (2.5) holds and if $\gamma = 1/2$, (2.6) holds with $\sigma^2 = \lambda_1\sigma_1^2 + \lambda_2\sigma_2^2$. □

Proof of Theorem 3.1. The convergence (3.3) is a straightforward consequence of Proposition 2.1 and the continuous mapping Theorem. If $\mu^* \neq \mu_0^*$, write

$$T_n = \frac{\bar{r}(n) - \mu^*}{|\bar{r}(n) - \bar{r}(n/2)|} + \frac{\mu^* - \mu_0^*}{|\bar{r}(n) - \bar{r}(n/2)|} \ .$$

The first term converges weakly to $A(1)/|A(1) - 2A(1/2)|$ and the second term converges in probability to $+\infty$ if $\mu^* > \mu_0^*$ and to $-\infty$ if $\mu^* < \mu_0^*$. □

Proof of Lemma 3.1.

$$s_n^2 = \frac{1}{n-1}\sum_{j=1}^{n}(r_j - \mu^*)^2 + \frac{n}{n-1}(\bar{r}_n - \mu^*)^2 + 2(\mu^* - \bar{r})\sum_{j=1}^{n}(r_j - \mu^*) \tag{9.1}$$

Under Assumption 2.1, the convergences (2.5) and (2.6) imply that the second term in (9.1) is $o_P(1)$. Since

$$r_j - \mu^* = \sum_{i=1}^{2}\sum_{k=N_i(j-1)+1}^{N_i(j)}(\mu_i + \zeta_{i,k}) - \lambda_i\mu_i \ ,$$

by Assumption (3.1), ergodicity of the marked point processes and mutual independence, we have

$$\frac{1}{n}\sum_{j=1}^{n}(r_j - \lambda\mu)^2 \xrightarrow{\mathbb{P}} \sum_{i=1}^{2}\mathbb{E}\left[\left(\sum_{k=1}^{N_i(1)}(\mu_i + \zeta_{i,k}) - \lambda_i\mu_i\right)^2\right] \ .$$

Since the last term in (9.1) is a cross product, it is therefore also $o_P(1)$. □

Lemma 9.1. *Let ξ be a stationary random measure such that $n^{-\gamma}\{\xi(nt) - \lambda nt\} \xrightarrow{fi.di.} A(t)$. Let N be a Cox process with stochastic intensity ξ. Then $n^{-\gamma}\{N(nt) - \lambda nt\} \xrightarrow{fi.di.} A(t)$.*

Proof. We only prove the convergence in distribution of $n^{-\gamma}\{\xi(n) - \lambda n\}$ to $A(1)$. The convergence of the finite dimensional distribution is proved similarly. By conditioning on ξ, we have, for all $z \in \mathbb{R}$,

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{\mathrm{i}zt^{-\gamma}\{N(t)-\lambda t\}}\right] &= \mathbb{E}\left[\mathrm{e}^{\xi(t)\{\mathrm{e}^{\mathrm{i}zt^{-\gamma}}-1\}-\mathrm{i}zt^{1-\gamma}}\right] \\
&= \mathbb{E}\left[\mathrm{e}^{\mathrm{i}zt^{-\gamma}\{\xi(t)-\lambda t\}}\right] + \mathbb{E}\left[\mathrm{e}^{\mathrm{i}zt^{-\gamma}\{\xi(t)-\lambda t\}}\left\{\mathrm{e}^{\xi(t)\{\mathrm{e}^{\mathrm{i}zt^{-\gamma}}-1-\mathrm{i}zt^{-\gamma}\}} - 1\right\}\right] \\
&= \phi_t(z) + \mathrm{rest}_t(z) \ .
\end{aligned}$$

By assumption, $\lim_{t\to\infty}\phi_t(z) = \mathbb{E}[\mathrm{e}^{\mathrm{i}zA(1)}]$. Then, denoting $h(t,z) = \mathrm{e}^{\mathrm{i}zt^{-\gamma}} - 1 - \mathrm{i}zt^{-\gamma}$, the remainder term is bounded by

$$|\mathrm{rest}_t(z)| \leq \mathbb{E}\left[\left|\mathrm{e}^{h(t,z)\xi(t)} - 1\right|\right] .$$

For each z, $h(t,z) = O(t^{-2\gamma})$, so the assumption on ξ implies that $\mathrm{e}^{h(t,z)\xi(t)} - 1 \xrightarrow{\mathbb{P}} 0$. Moreover, for all $t > 0$ and $z \in \mathbb{R}$, $|\mathrm{e}^{h(t,z)\xi(t)} - 1| \leq 1 + \mathrm{e}^{(cos(z)-1)\xi(t)} \leq 2$. Thus, the bounded convergence theorem yields that $\lim_{t\to\infty}\mathbb{E}[|\mathrm{e}^{h(t,z)\xi(t)} - 1|] = 0$. □

We recall the definition of θ-weak dependence from Bardet et al. (2008). For each positive integer v, let $\mathbb{R}^v$ be equipped with the l^1-norm

$$\|\mathbf{x} - \mathbf{y}\|_1 = \sum_{i=1}^{v} |x_i - y_i|, \ \mathbf{x}, \mathbf{y} \in \mathbb{R}^v .$$

Let $\mathcal{G}_v$ be the class of bounded functions $g : \mathbb{R}^v \to \mathbb{R}$ such that

$$\mathrm{Lip}(g) = \sup_{\mathbf{x}\neq\mathbf{y}} \frac{|g(\mathbf{x}) - g(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_1} < \infty .$$

The quantity in the left hand side is denoted $\mathrm{Lip}(g)$ and called the Lipschitz modulus of g. A stationary sequence $\{X_k, k \in \mathbb{Z}\}$ is said to be θ weakly dependent with rate $\{\theta_k, k \geq 0\}$ if for any $u, v \in \mathbb{N}^*$, $(i_1, \ldots, i_u) \in \mathbb{Z}$, $(j_1, \ldots, j_v) \in \mathbb{Z}$ with $i_1 < \cdots < i_u \leq i_u + k \leq j_1 < \cdots < j_v$, and $f \in \mathcal{G}_u$, $g \in \mathcal{G}_v$,

$$\mathrm{cov}(f(r_{i_1}, \ldots, r_{i_u}), g(r_{j_1}, \ldots, r_{j_v})) \leq v\, \|f\|_\infty \,\mathrm{Lip}(g)\, \theta_k .$$

Lemma 9.2. *Let N be a Cox process with stochastic intensity ξ, where*

$$\xi(A) = \int_A w(s)\mathrm{d}s ,$$

and w is a positive stationary stochastic process with almost surely continuous paths. Assume that w is θ-weak dependent. Then

(i) the increments $\{\Delta N(k), k \in \mathbb{Z}\}$ is θ-weak dependent with the same rate function as the process w;

(ii) If the efficient shocks e_i are i.i.d. and independent of the point process N, the sequence of calendar-time returns $\{r_k, k \in \mathbb{N}^*\}$, defined by

$$r_k = \sum_{n=N(k-1)+1}^{N(k)} e_n ,$$

is θ-weak dependent, with the same rate as w.

Proof. Since (i) is a particular case of (ii) with the shocks taken to be uniformly equal to 1, we only need to prove (ii). Let $f \in \mathcal{G}_u$ and $g \in \mathcal{G}_v$. Consider now the sequence of returns $\{r_k, k \in \mathbb{N}^*\}$. By conditioning on the stochastic intensity, we have

$$\begin{aligned}\operatorname{cov}(f(r_{i_1},\dots,r_{i_u}), g(r_{j_1},\dots,r_{j_v})) &= \mathbb{E}\left[\operatorname{cov}(f(r_{i_1},\dots,r_{i_u}), g(r_{j_1},\dots,r_{j_v})) \mid \xi\right] \\ &\quad + \operatorname{cov}\left(\mathbb{E}(f(r_{i_1},\dots,r_{i_u})|\xi), \mathbb{E}(g(r_{j_1},\dots,r_{j_v})|\xi)\right)\end{aligned}$$

Since the intervals are non intersecting, and $\{e_k\}$ is independent of N which is a Poisson point process conditionally on ξ, we have

$$\mathbb{E}\left(\operatorname{cov}(f(r_{i_1},\dots,r_{i_u}), g(r_{j_i},\dots,r_{j_v})|\xi)\right) = 0 \ .$$

To compute the second term, denote $h_n(x) = \frac{e^{-x}x^n}{n!}$ and let $\{e_k^{(i)}\}$, $i \in \mathbb{N}$ be i.i.d. copies of the sequence $\{e_k\}$ and define $S_i(0) = 0$ and $S_i(n) = \sum_{k=1}^n e_k^{(i)}$. Denote also $\xi_i = \xi(i-1, i]$. Then, by the independent increments property, we have

$$\begin{aligned}\mathbb{E}\left[f(r_{i_1},\dots,r_{i_u})|\xi\right] &= \mathbb{E}\left[f\left(\sum_{k=N(i_1-1)+1}^{N(i_1)} e_k, \dots, \sum_{k=N(i_u-1)+1}^{N(i_u)} e_k\right)\middle|\xi\right] \\ &= \sum_{n_1,\dots,n_u=0}^{\infty} \mathbb{E}[f(S_1(n_1),\dots,S_u(n_u))] \prod_{i=1}^{n} h_{n_i}(\xi_i) \ .\end{aligned}$$

Similarly, we have

$$\mathbb{E}\left(g(r_{j_1},\dots,r_{j_v})|\xi\right) = \sum_{m_1,\dots,m_v=0}^{\infty} \mathbb{E}\left[g(S_1(m_1),\dots,S_v(m_v))\right] \prod_{q=1}^{v} h_{m_q}(\xi_{j_q}) \ .$$

Define the functions F and G on $\mathbb{R}^u$ and $\mathbb{R}^v$ respectively by

$$\begin{aligned}F(x_1,\dots,x_u) &= \sum_{n_1,\dots,n_u=0}^{\infty} \mathbb{E}\left[f(S_1(n_1),\dots,S_u(n_u))\right] \prod_{i=1}^{u} h_{n_i}(x_i) \ , \\ G(x_1,\dots,x_v) &= \sum_{m_1,\dots,m_v=0}^{\infty} \mathbb{E}\left[g(S_1(m_1),\dots,S_v(m_v))\right] \prod_{i=1}^{v} h_{m_i}(x_i)\end{aligned}$$

Since f is bounded, we obtain, for all $\boldsymbol{x} \in \mathbb{R}^u$,

$$F(\boldsymbol{x}) \leq \|f\|_\infty \sum_{n_1,\dots,n_u=0}^{\infty} \prod_{p=1}^{u} h_{n_p}(x_p) = ||f||_\infty \ .$$

Thus $\|F\|_\infty \leq \|f\|_\infty$. We now prove that F is a Lipschitz function. For $\boldsymbol{x} = (x_1, \ldots, x_u) \in \mathbb{R}^u$, we have

$$\begin{aligned}
\frac{\partial F(\boldsymbol{x})}{\partial x_1} &= \sum_{n_1=1, n_2, \ldots, n_u=0}^{\infty} \mathbb{E}\left[f(S_1(n_1), \ldots, S_u(n_u))\right] \frac{\mathrm{e}^{-x_1} n_1 x_1^{n_1-1}}{n_1!} \prod_{p=2}^{u} h_{n_p}(x_p) \\
&\quad - \sum_{n_1, \ldots, n_u=0}^{\infty} \mathbb{E}\left[f(S_1(n_1), S_2(n_2), \ldots, S_u(n_u))\right] \prod_{p=1}^{u} h_{n_p}(x_p) \\
&= \sum_{n_1, \ldots, n_u=0}^{\infty} \mathbb{E}\left[f(S_1(n_1+1), \ldots, S_{i_u}(n_u))\right] \prod_{p=1}^{u} h_{n_p}(x_p) \\
&\quad - \sum_{n_1, n_2, \ldots, n_u=0}^{\infty} \mathbb{E}\left[f(S_1(n_1), \ldots, S_u(n_u))\right] \prod_{p=1}^{u} h_{n_p}(x_p) \\
&= \sum_{n_1, \ldots, n_u=0}^{\infty} \mathbb{E}\left[f(S_1(n_1+1), S_2(n_2), \ldots, S_u(n_u)) - f(S_1(n_1), \ldots, S_u(n_u))\right] \prod_{p=1}^{u} h_{n_p}(x_p) \ .
\end{aligned}$$

Applying the Lipschitz property of f, we have

$$\begin{aligned}
\left|\frac{\partial F(\boldsymbol{x})}{\partial x_1}\right| &\leq \sum_{n_1, \ldots, n_u=0}^{\infty} \mathbb{E}\left[|f(S_1(n_1+1), S_{i_2}(n_2), \ldots, S_{i_u}(n_u)) - f(S_{i_1}(n_1), \ldots, S_{i_u}(n_u))|\right] \prod_{p=1}^{u} h_{n_p}(x_p) \\
&\leq \mathrm{Lip}(f) \sum_{n_1, \ldots, n_u=0}^{\infty} \mathbb{E}\left[|S_{i_1}(n_1+1) - S_{i_1}(n_1)|\right] \prod_{p=1}^{u} h_{n_p}(x_p) \\
&= \mathrm{Lip}(f) \mathbb{E}[|e_1|] \sum_{n_1, \ldots, n_u=0}^{\infty} \prod_{p=1}^{u} h_{n_p}(x_p) = \mathrm{Lip}(f) \mathbb{E}[|e_1|] \ .
\end{aligned}$$

Similarly, for each $p = 2, \ldots, u$,

$$\left\|\frac{\partial F}{\partial x_p}\right\|_\infty \leq \mathrm{Lip}(f)\, \mathbb{E}[|e_1|] \ .$$

Thus F is also a Lipschitz function and $\mathrm{Lip}(F) \leq \mathbb{E}[|e_1|]\mathrm{Lip}(f)$. We can similarly prove that G is a bounded Lipschitz function. Let us assume for the moment that $\{\xi_i, i \in \mathbb{Z}\}$ is θ-weak dependent. We then conclude that

$$\begin{aligned}
\mathrm{cov}(f(r_{i_1}, \ldots, r_{i_u}), g(r_{j_i}, \ldots, r_{j_v})) &= \mathrm{cov}\left(\mathbb{E}[f(r_{i_1}, \ldots, r_{i_u})|\xi], \mathbb{E}[g(r_{j_i}, \ldots, r_{j_v}) \mid \xi]\right) \\
&= \mathrm{cov}(F(\xi_{i_1}, \ldots, \xi_{i_u}), G(\xi_{j_1}, \ldots, \xi_{j_v})) \\
&\leq v\|F\|_\infty \mathrm{Lip}(G)\theta_r \leq v\, \mathbb{E}[|e_1|]\, \|f\|_\infty\, \mathrm{Lip}(g)\, \theta_r \ .
\end{aligned}$$

This proves that $\{r_k, k \in \mathbb{N}^*\}$ is θ-weak dependent with the same rate as $\{\xi_i, i \in \mathbb{Z}\}$.

Let us now prove the θ-weak dependence of the sequence $\{\xi_i, i \in \mathbb{Z}\}$. Let $f \in \mathcal{G}_u$ and $g \in \mathcal{G}_v$. We must prove that

$$|\mathrm{cov}(f(\xi_{i_1}, \ldots, \xi_{i_u}), g(\xi_{j_i}, \ldots, \xi_{j_v}))| \leq C\, v\, \|f\|_\infty \,\mathrm{Lip}(g)\theta_r \ ,$$

for some constant C and all $(u+v)$-tuples of integers $i_1 < \cdots < i_u < j_1 < \cdots < j_v$ such that $j_1 > i_u + r$. Since the density w of the stochastic intensity has almost surely continuous path, it holds that ξ_i is the almost sure limit of the Riemann sums $\xi_{n,i} = n^{-1} \sum_{q=1}^{n} w(i-1+q/n)$. Since f and g are continuous and bounded, we obtain by bounded convergence

$$\mathrm{cov}(f(\xi_{i_1} \ldots, \xi_{i_u}), g(\xi_{j_1}, \ldots, \xi_{j_v})) = \lim_{n\to\infty} \mathrm{cov}(f(\xi_{i_1,n} \ldots, \xi_{i_u,n}), g(\xi_{j_1,n}, \ldots, \xi_{j_v,n})) \ . \tag{9.2}$$

Given a function $h : \mathbb{R}^p \to \mathbb{R}$, define the function h_n on $\mathbb{R}^{np}$ by

$$h_n(x_{i,q}, 1 \leq i \leq p, 1 \leq q \leq n) = h\left(\frac{x_{1,1} + \cdots + x_{1,n}}{n}, \ldots, \frac{x_{p,1} + \cdots + x_{p,n}}{n}\right)$$

If h is bounded and Lipschitz, then $\|h_n\|_\infty \leq \|h\|_\infty$ and $\mathrm{Lip}(h_n) \leq n^{-1}\mathrm{Lip}(h)$. Since the process w is θ weakly dependent, we obtain

$$\begin{aligned}
&|\mathrm{cov}(f(\xi_{i_1,n} \ldots, \xi_{i_u,n}), g(\xi_{j_1,n}, \ldots, \xi_{j_v,n}))| \\
&\quad = |\mathrm{cov}(f_n(w(i_s - 1 + \frac{q}{n}), s = 1, \ldots, u, q = 1, \ldots, n), g_n(w(j_t - 1 + \frac{q}{n}), t = 1, \ldots, v, q = 1, \ldots, n) \\
&\quad \leq vn \times \|f\|_\infty \times n^{-1}\mathrm{Lip}(g) \times \theta_{k-1} \\
&\quad = v\|f\|_\infty \,\mathrm{Lip}(g)\theta_{k-1} \ .
\end{aligned} \tag{9.3}$$

The bounds (9.2) and (9.3) proves that the sequence $\{\xi_i, i \in \mathbb{Z}\}$ is θ weak dependent with rate θ_{k-1}. □

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc.*, 96(453):42–55.

Arcones, M. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *The Annals of Probability*, 22(4):2242–2274.

Aue, A., Horváth, L., Hurvich, C., and Soulier, P. (2014). Limit laws in transaction-level asset price models. *Econometric Theory*, 30(3):536–579.

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

Bacry, E., Delattre, S., Hoffmann, M., and Muzy, J.-F. (2011). Modelling microstructure noise with mutually exciting point processes. To be published in Quantitative Finance.

Bardet, J.-M., Doukhan, P., Lang, G., and Ragache, N. (2008). Dependent Lindeberg central limit theorem and some applications. *ESAIM. Probability and Statistics*, 12:154–172.

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

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

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

Bowsher, C. G. (2007). Modelling security market events in continuous time: intensity based, multivariate point process models. *J. Econometrics*, 141(2):876–912.

Chen, F., Diebold, F. X., and Schorfheide, F. (2012). A markov-switching multifractal inter-trade duration model, with application to u.s. equities. Preprint.

Daley, D. J. (2010). Long-range dependence in a Cox process directed by an alternating renewal process. In *Probability and mathematical genetics*, volume 378 of *London Math. Soc. Lecture Note Ser.*, pages 169–184. Cambridge Univ. Press, Cambridge.

Delattre, S., Robert, C. Y., and Rosenbaum, M. (2013). Estimating the efficient price from the order flow : a brownian cox process approach. arXiv:1301.3114.

Deo, R., Hsieh, M., and Hurvich, C. M. (2010). Long memory in intertrade durations, counts and realized volatility of NYSE stocks. *Journal of Statistical Planning and Inference*, 140(12):3715–3733.

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

Dobrushin, R. L. and Major, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete*, 50(1):27–52.

Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Processes and their Applications*, 84(2):313–342.

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

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

Hautsch, N. (2012). *Econometrics of financial high-frequency data*. Springer, Heidelberg.

Heath, D., Resnick, S., and Samorodnitsky, G. (1998). Heavy tails and long range dependence in ON/OFF processes and associated fluid models. *Mathematics of Operations Research*, 23(1):145–165.

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

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

Jach, A., McElroy, T., and Politis, D. N. (2012). Subsampling inference for the mean of heavy-tailed long-memory time series. *Journal of Time Series Analysis*, 33(1):96–111.

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

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

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36(4):394–419.

McElroy, T. and Jach, A. (2012). Subsampling inference for the autocovariances and autocorrelations of long-memory heavy-tailed linear time series. *Journal of Time Series Analysis*, 33(6):935–953.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59(2):347–370.

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

Prigent, J.-L. (2001). Option pricing with a general marked point process. *Mathematics of Operations Research*, 26(1):50–66.

Scholes, M. and Williams, J. (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics*, 5:309–327.

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

Taqqu, M. S. and Levy, J. B. (1986). Using renewal processes to generate long-range dependence and high variability. In *Dependence in probability and statistics (Oberwolfach, 1985)*, volume 11 of *Progr. Probab. Statist.*, pages 73–89. Birkhauser.

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