Econometric Theory, **25**, 2009, 764–792. Printed in the United States of America.
doi:10.1017/S0266466608090294

CONDITIONS FOR THE PROPAGATION OF MEMORY PARAMETER FROM DURATIONS TO COUNTS AND REALIZED VOLATILITY

ROHIT DEO AND CLIFFORD M. HURVICH
New York University

PHILIPPE SOULIER
Université Paris X

YI WANG
Wellington Management Company, LLP

We establish sufficient conditions on durations that are stationary with finite variance and memory parameter $d \in [0, \frac{1}{2})$ to ensure that the corresponding counting process $N(t)$ satisfies $\operatorname{Var} N(t) \sim Ct^{2d+1}$ $(C > 0)$ as $t \to \infty$, with the same memory parameter $d \in [0, \frac{1}{2})$ that was assumed for the durations. Thus, these conditions ensure that the memory parameter in durations propagates to the same memory parameter in the counts. We then show that any autoregressive conditional duration ACD(1,1) model with a sufficient number of finite moments yields short memory in counts, whereas any long memory stochastic duration model with $d > 0$ and all finite moments yields long memory in counts, with the same d. Finally, we provide some results about the propagation of long memory to the empirically relevant case of realized variance estimates affected by market microstructure noise contamination.

1. INTRODUCTION

There is a growing literature on long memory in volatility of financial time series. See, e.g., Robinson (1991), Robinson and Henry (1999), Deo and Hurvich (2001), and Hurvich, Moulines, and Soulier (2005). Long memory in volatility, which has been repeatedly found in the empirical literature (see Breidt, Crato, and de Lima, 1998; Baillie, Bollerslev, and Mikkelsen, 1996; Bollerslev and Mikkelsen, 1996; Andersen and Bollerslev, 1997), plays a key role in the forecasting of realized volatility, defined as the sum of squared high-frequency returns (Andersen,

The authors thank the referees for invaluable suggestions that led to more elegant and shorter proofs and a better economic interpretation of the results. They also thank Jushan Bai, Xiaohong Chen, and Raymond Thomas for helpful comments and suggestions. Part of this research was conducted while Deo was at the University of Texas–Austin. Address correspondence to Rohit Deo, New York University, 44 W. 4th Street, New York, NY 10012, USA; e-mail: rdeo@stern.nyu.edu.

 © 2009 Cambridge University Press 0266-4666/09 $15.00

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Bollerslev, Diebold, and Labys, 2001; Barndorff-Nielsen and Shephard, 2006; Deo, Hurvich, and Lu, 2006), and has important implications on option pricing (see Comte and Renault, 1998).

Given the increasing availability of transaction-level data it is of interest to explain phenomena observed at longer time scales from equally spaced returns in terms of more fundamental properties at the transaction level. Engle and Russell (1998) proposed the autoregressive conditional duration (ACD) model to describe the durations between trades, and they briefly explored the implications of this model on volatility of returns in discrete time, though they did not determine the persistence of this volatility, as measured, say, by the decay rate of the autocorrelations of the squared returns. Deo, Hsieh, and Hurvich (2007) proposed the long-memory stochastic duration (LMSD) model and carried out an empirical exploration of the question as to which properties of durations lead to long memory in volatility.

In principle, why should there be a link between the dependence properties of durations and volatility? The durations between transactions determine the number of transactions (counts) in a given interval of time. Thus, any dependence in durations will affect the dependence in the time series of counts. Inspired by the work of Clark (1973), and other related work (Tauchen and Pitts, 1983; Epps and Epps, 1976; Bollerslev and Jubinski, 1999) on the mixture of distributions hypothesis (MDH), financial econometricians now generally accept that dependence in counts is related to dependence in volatility. Overall, then, dependence in durations and volatility are closely linked.

The specification of a model for the durations, together with a model for the transaction-level price changes, constitutes a pure-jump model for the prices. In such a model, it is the transactions that generate price changes. Therefore, the (logarithmic) price change in a given time interval is the accumulation of transaction-level price changes. Furthermore, the total number of transactions up to a given time is simply the counting process induced by the series of intertrade durations. Such models have been considered recently by Oomen (2006) and Rydberg and Shephard (2003), among others. Clearly, then, properties of realized volatility calculated from high-frequency returns generated from such a process would also depend on the properties of durations. It is therefore essential to obtain a better understanding of the theoretical link between durations, counts, and realized volatility so that price process models can be compatible with what is observed in empirical data. This paper provides a study of such a link.

We present first a few basic definitions. The collection of time points $\ldots t_{-1} < t_0 \leq 0 < t_1 < t_2 < \ldots$ at which a transaction (say, a trade of a particular stock on a specific market) takes place comprises a point process, a fact that was exploited by Engle and Russell (1988). These event times $\{t_k\}$ determine a *counting process*,

$$N(t) = \textit{Number of Events in } (0, t].$$

For any fixed time spacing $\Delta t > 0$, one can define the *counts* $\Delta N_s = N(s\Delta t) - N((s-1)\Delta t)$, the number of events in the sth time interval of width Δt, where

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

$s = 1, 2, \ldots$. The event times $\{t_k\}_{k=-\infty}^{\infty}$ also determine the *durations*, given by $\{\tau_k\}_{k=-\infty}^{\infty}$, $\tau_k = t_k - t_{k-1}$. In this paper, we will say that a stationary series in discrete time (either counts or durations) has long memory with memory parameter $d \in (0, \frac{1}{2})$ (or short memory if $d = 0$) if the partial sum of n contiguous values of the series has a variance that behaves as Cn^{2d+1} as $n \to \infty$, where $C > 0$ is a constant. We will say that a stationary counting process $N(t)$ has long memory with memory parameter $d \in (0, \frac{1}{2})$ (or short memory if $d = 0$) if $\operatorname{Var} N(t) \sim Ct^{2d+1}$ as $t \to \infty$ where $C > 0$ is a constant. Note that if $N(t)$ has memory parameter $d \in [0, \frac{1}{2})$, then so does the series of counts, with the same memory parameter, for any value of Δt.

Deo et al. (2007) analyzed transaction-level data on 10 stocks traded on the NYSE. They estimated the memory parameter of the durations between transactions, the counts and squared returns for various choices of Δt, and the daily realized volatility. The estimates were all roughly similar, taking the standard errors and other issues into account, typically lying in the range from 0.3 to 0.45. We establish theoretically that under certain reasonable conditions the memory parameter (be it positive or zero) propagates unchanged from the durations to the counting process and then to realized volatility. Our results have important implications for the choice of an appropriate model for durations.

Here, and in other work involving point processes, it is important to distinguish between two probability measures, P and P^0, that are different in general: under P, the point process is stationary, whereas under P^0 the durations are stationary. We provide more details on these measures in Section 2.

Deo et al. (2007) pointed out, using a theorem of Daley, Rolski, and Vesilo (2000), that if durations are generated by an ACD model and if the durations have tail index $\kappa \in (1, 2)$ under P^0, then the resulting counting process $N(t)$ has long memory with memory parameter $d \geq 1 - \kappa/2$, under P. An open question, however, is the determination of conditions under which the memory parameter of a finite-variance duration sequence propagates unchanged to the counting process. This question is a nontrivial one, because even a time-instantaneous transformation of a Gaussian long-memory process may reduce the memory parameter. See Dittmann and Granger (2002).

In Section 2, we will establish sufficient conditions on durations that are stationary with finite variance and memory parameter $d \in [0, \frac{1}{2})$ under P^0 to ensure that the corresponding counting process $N(t)$ satisfies $\operatorname{Var} N(t) \sim Ct^{2d+1}$ $(C > 0)$ as $t \to \infty$ under P, with the same memory parameter $d \in [0, \frac{1}{2})$ that was assumed for the durations. Thus, these conditions, given in Theorem 1, ensure that the memory parameter in durations propagates to the same memory parameter in counts. Moreover, we show that under a particular model for prices, the same memory parameter propagates further from the counting process to the realized volatility.

We will verify that the sufficient conditions of our Theorem 1 are satisfied for the ACD(1,1) model assuming finite $3 + \epsilon$ moment $(\epsilon > 0)$ of the durations under P^0 and for the LMSD model with any $d \in [0, \frac{1}{2})$ assuming that the multiplying shocks have all moments finite. Thus, any ACD(1,1) model with a sufficient

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

number of finite moments yields short memory in the counting process, whereas any LMSD model with $d > 0$ and all finite moments yields long memory in the counting process. These results for the LMSD and ACD(1,1) models are given in Theorems 2 and 3 in Section 2, respectively. Proposition 1 (in Section 2), which is used in proving Theorem 2, provides a Rosenthal-type inequality for moments of absolute standardized partial sums of durations under the LMSD model and is of interest in its own right.

In Section 3 we present a result (Theorem 4) implying that if counts have memory parameter $d \in [0, \frac{1}{2})$ then further aggregations of these counts to longer time intervals will have a lag-1 autocorrelation that tends to $2^{2d} - 1$ as the level of aggregation grows. Interestingly, this limit is zero if and only if $d = 0$. Thus, one of the important functions of long memory in counts is that it allows the counts to have a nonvanishing autocorrelation even as Δt grows, as was found by Deo et al. (2007) to occur in empirical data. By contrast, short memory in counts implies that counts at long time scales (large Δt) are essentially uncorrelated, in contradiction to what is seen in actual data. To summarize, aggregation ultimately destroys all autocorrelation in counts, if and only if the counts have short memory.

Finally, in Section 4 we provide the link between counts and realized volatility. Using a simple pure-jump model for prices, we show that if the durations have long memory then realized volatility has long memory with the same memory parameter. We also show that if the durations have short memory and finite variance then realized volatility has short memory. Thus, the properties of durations have direct implications on the properties of realized volatility.

The pure-jump model for prices mentioned earlier assumes that the transaction-level returns are independent. However, as we discuss in Section 4, the observed transaction-level returns exhibit autocorrelation at short lags, due to market microstructure noise. Hence we relax the independence assumption to a far weaker summability condition on the autocovariances of the transaction-level returns, and then we replace the realized volatility by the transaction-time-sampling version found to have superior properties by Oomen (2006). We show that this alternative version of realized volatility inherits the same memory parameter as the durations. Thus, we have results for an empirically relevant noise-contaminated realized variance estimate.

2. THEOREMS ON THE PROPAGATION OF THE MEMORY PARAMETER

Both the ACD and LMSD models imply that the doubly infinite sequence of durations $\{\tau_k\}_{k=-\infty}^{\infty}$ is a stationary time series, i.e., there exists a probability measure P^0 under which the joint distribution of any subcollection of the $\{\tau_k\}$ depends only on the lags between the entries. On the other hand, a point process N on the real line is stationary under the measure P if for every positive integer r the joint distribution of $\{N(A_1 + t), \ldots, N(A_r + t)\}$ does not depend on t, where $A_1, \ldots, A_r$ are bounded Borel sets. A fundamental fact about point processes is that in general

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

(a notable exception is the Poisson process) there is no single measure under which both the point process N and the durations $\{\tau_k\}$ are stationary, i.e., in general P and P^0 are not the same. Nevertheless, there is a one-to-one correspondence between the class of measures P^0 that determine a stationary duration sequence and the class of measures P that determine a stationary point process. The measure P^0 corresponding to P is called the *Palm distribution*. The counts are stationary under P, whereas the durations are stationary under P^0.

An economic interpretation of the distinction between P and P^0 is as follows. If the cumulative number of transactions $N(t)$ is calculated from the opening bell (e.g., 9:00 E.T., Wednesday, July 12, 2006), then the appropriate measure is P. By contrast, if $N(t)$ is calculated from the first transaction of that trading day, then the appropriate measure would be P^0. For thinly traded stocks, the difference between these two starting points may be quite large, which helps highlight the difference between the two measures P and P^0 as follows. In general, it may be of economic interest to assess the expected intertrade duration. Perhaps surprisingly, this expectation will be different according to whether the time scale for $N(t)$ is initialized at the opening bell (P) or at the first trade of the day (P^0). In the former case the expectation is larger than in the latter case (cf. eqn. (A.1) in the Appendix), because in the former case the duration is longer than the waiting time to the next event. The more thinly traded the stock is, the larger the difference between the expected values of τ_1 would be under the two measures.

Let E, E^0, Var, Var^0 denote expectations and variances under P and P^0, respectively. Define $\mu = \mathrm{E}^0(\tau_k)$ and $\lambda = 1/\mu$. Our main theorem uses the assumption that P^0 is $\{\tau_k\}$-mixing, defined as follows. Let $\mathcal{N} = \sigma(\{\tau_k\}_{k=-\infty}^{\infty})$ and $\mathcal{F}_n = \sigma(\{\tau_k\}_{k=n}^{\infty})$. Following Nieuwenhuis (1989, p. 597), we say that P^0 is $\{\tau_k\}$-mixing if

$$\lim_{n\to\infty} \sup_{B\in\mathcal{N}\cap\mathcal{F}_n} |P^0(A\cap B) - P^0(A)P^0(B)| = 0$$

for all $A \in \mathcal{N}$.

Our theorem also uses the notion of α-mixing, also known as strong mixing (see Bradley, 2005, Sect. 2.1).

THEOREM 1. *Let $\{\tau_k\}$ be a duration process such that the following conditions hold:*

(i) $\exists\, d \in [0, \frac{1}{2})$ *such that*

$$Y_n(s) = \frac{\sum_{k=1}^{\lfloor ns\rfloor}(\tau_k - \mu)}{n^{1/2+d}}, \qquad s \in [0,1]$$

converges weakly to $\sigma B_{1/2+d}(\cdot)$ under P^0, where $\sigma > 0$ and $B_{1/2+d}(\cdot)$ is fractional Brownian motion if $0 < d < \frac{1}{2}$ or standard Brownian motion $B_{1/2} = B$ if $d = 0$.

(ii) Consider the d in (i). If $d \in (0, \frac{1}{2})$, then P^0 is $\{\tau_k\}$-mixing. If $d = 0$, then $\{\tau_k\}$ is exponential strong mixing.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

(iii) For the d in (i),

$$\sup_n \mathrm{E}^0 \left| \frac{\sum_{k=1}^n (\tau_k - \mu)}{n^{1/2+d}} \right|^p < \infty \qquad \begin{cases} \text{for all } p > 0, \text{ if } d \in (0, \frac{1}{2}) \\ \text{for } p = 3+\epsilon,\ \epsilon > 0, \text{ if } d = 0. \end{cases}$$

Then the induced counting process $N(t)$ satisfies $\mathrm{Var}\, N(t) \sim Ct^{2d+1}$ under P as $t \to \infty$ where $C > 0$.

Remark 1. Assumption (iii) implies that $Y_n^2(1)$ is uniformly integrable. This, together with (i), implies that $\mathrm{Var}^0[Y_n(1)] \sim C$ for some $C > 0$ and hence that the durations have long (or short) memory with memory parameter $d \in [0, \frac{1}{2})$.

Remark 2. If $d = 0$, the proof of Theorem 1 implies that $\mathrm{Var}^0 N(t) \sim Ct$ (i.e., under the Palm measure P^0) even when Assumption (iii) is weakened to $p = 2+\epsilon$ for $\epsilon > 0$. Nevertheless, we find it preferable to present and prove the result in Theorem 1 in terms of $\mathrm{Var}\, N(t)$, i.e., under the measure P, because it is under this measure that $N(t)$ is a stationary point process. The need for the more stringent condition in our Theorem 1 can be explained as follows. For $\mathrm{E}[\tau_k^{2+\epsilon}] < \infty$, it is necessary in general to have $\mathrm{E}^0[\tau_k^{3+\epsilon}] < \infty$. See equation (1.2.25) of Baccelli and Brémaud (2003), which we use in our proof.

Remark 3. Inspection of the proof of Theorem 1 in the Appendix reveals that if $d > 0$, only $(5-2d)/(1-2d)+\delta$ finite moments are needed, where δ is arbitrarily small. The closer d is to $1/2$, the larger the number of finite moments required.

Remark 4. As pointed out by Nieuwenhuis (1989), if $\{\tau_k\}$ is strong mixing under P^0 then P^0 is $\{\tau_k\}$-mixing. Nevertheless, this weaker form of mixing is essential for our purposes in the case $d > 0$ because even Gaussian long-memory processes are not strong mixing. See Viano, Deniau, and Oppenheim (1995).

We now present two useful models for durations and give results stating that the conditions of Theorem 1 apply for these models, thereby yielding the properties of the variance of the induced counting processes.

2.1. LMSD Process

A latent-variable model for durations is the stochastic duration (SCD) model of Bauwens and Veredas (2004). The model is given by

$$\tau_k = e^{h_k}\epsilon_k,$$
$$h_k = \omega + \beta h_{k-1} + e_k,$$

where $\omega \in \mathbb{R}$, $|\beta| < 1$, the $\{e_k\}$ are $iid\, N(0, \sigma^2)$ and the $\{\epsilon_k\}$ are independent and identically distributed (i.i.d.) with unit mean and positive support. This model is analogous to the widely used stochastic volatility (SV) model for returns (see Harvey, 1998). Bauwens and Veredas (2004) estimated the SCD model to data and found that the autoregressive coefficient β in the latent process $\{h_k\}$ is typically

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

extremely close to 1. Based on a semiparametric analysis, Deo et al. (2007) concluded that the intertrade durations for stock prices have long memory, and so they proposed the LMSD model, in which the $\{h_k\}$ possess long memory. The generalization from the SCD to the LMSD model for durations is analogous to that from the SV to the long memory stochastic volatility (LMSV) model for returns of Harvey (1998) and Breidt et al. (1998).

Following Deo et al. (2007), we now define the LMSD process $\{\tau_k\}_{k=-\infty}^{\infty}$ for $d \in [0, \frac{1}{2})$ as

$$\tau_k = e^{h_k}\epsilon_k,$$

where under P^0 the $\epsilon_k \geq 0$ are i.i.d. with all moments finite, $h_k = \sum_{j=0}^{\infty} b_j e_{k-j}$, the $\{e_k\}$ are i.i.d. Gaussian with zero mean, independent of $\{\epsilon_k\}$, and

$$b_j \sim \begin{cases} Cj^{d-1} & \text{if } d \in (0, \frac{1}{2}) \\ Ca^j,\ |a| < 1 & \text{if } d = 0 \end{cases}$$

$(C \neq 0)$ as $j \to \infty$; for $d = 0$ we further assume that the spectral density of $\{h_k\}$ is bounded away from zero. Any stationary, invertible autoregressive fractionally integrated moving average (ARFIMA) process $\{h_k\}$ with $d \in [0, \frac{1}{2})$ would satisfy the preceding assumptions. Note that for convenience we nest the short-memory case $(d = 0)$ within the LMSD model, so that the allowable values for d in this model are $0 \leq d < 1/2$. In their empirical analysis, Deo et al. (2007) assumed that the $\{h_k\}$ follow an ARFIMA$(1, d, 0)$ process. In the theoretical results of this paper, we do not make any specific parametric assumption on the $\{h_k\}$.

The following theorem establishes that long memory propagates unchanged from LMSD durations to the counting process.

THEOREM 2. *If the durations $\{\tau_k\}$ are generated by the LMSD process with $d \in [0, \frac{1}{2})$, then the conditions of Theorem 1 are satisfied, and therefore the induced counting process $N(t)$ satisfies* $\operatorname{Var} N(t) \sim Ct^{2d+1}$ *under P as $t \to \infty$ where $C > 0$.*

To establish Theorem 2, we will use the following Rosenthal-type inequality, which is of independent interest.

PROPOSITION 1. *For durations $\{\tau_k\}$ generated by the LMSD process with $d \in [0, \frac{1}{2})$, for any fixed positive integer $p \geq 2$, $\mathrm{E}^0\{|y_n - \mathrm{E}^0(y_n)|^p\}$ is bounded uniformly in n, where*

$$y_n = \frac{\sum_{k=1}^{n} \tau_k}{n^{1/2+d}}.$$

2.2. ACD(1,1) Process

The ACD process was introduced by Engle and Russell (1998). Whereas the LMSD model builds dependence in the durations through an unobservable latent

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

variable process, the ACD model is observation driven. The ACD model treats the conditional mean of durations in the same way that the generalized autoregressive conditional heteroskedasticity (GARCH) model treats the conditional variance of returns. The simplest version is the ACD(1,1) process given by $\{\tau_k\}_{k=-\infty}^{\infty}$ as

$$\tau_k = \psi_k \epsilon_k,$$
$$\psi_k = \omega + \alpha \tau_{k-1} + \beta \psi_{k-1}$$

with $\omega > 0$, $\alpha > 0$, $\beta \geq 0$, and $\alpha + \beta < 1$, where under P^0, $\epsilon_k \geq 0$ are i.i.d. with mean 1. We will assume further that under P^0, ϵ_k has a density g_ϵ such that $\int_0^\theta g_\epsilon(x)\,dx > 0, \forall\, \theta > 0$ and $\mathrm{E}^0\left(\tau_k^{3+\epsilon}\right) < \infty$ for some $\epsilon > 0$. This last assumption entails further restrictions on α and β. See Nelson (1990).

Nelson (1990) guarantees the existence of the doubly infinite ACD(1,1) process $\{\tau_k\}_{k=-\infty}^{\infty}$, which in our terminology is stationary under P^0.

The next theorem shows that short memory propagates from ACD(1, 1) durations (with sufficiently many finite moments) to the counting process.

THEOREM 3. *Suppose that the durations $\{\tau_k\}$ are generated by the* ACD(1,1) *model, with the additional assumptions stated previously. Then the conditions of Theorem 1 are satisfied, and therefore the induced counting process $N(t)$ satisfies* $\mathrm{Var}\, N(t) \sim Ct$ *under P as $t \to \infty$ where $C > 0$.*

3. AUTOCORRELATION OF AGGREGATED COUNTS

The following elementary result relates the memory parameter of a stationary process to the lag-1 autocorrelation of partial sums as the level of aggregation grows.

THEOREM 4. *Let $\{X_t\}$ be a stationary process such that* $\mathrm{Var}\left(\sum_{t=1}^{n} X_t\right) \sim Cn^{1+2d}$ *as $n \to \infty$, where $C \neq 0$ and $d \in [0, \frac{1}{2})$. Then*

$$\lim_{n\to\infty} \mathrm{Corr}\left[\sum_{t=1}^{n} X_t, \sum_{t=n+1}^{2n} X_t\right] = 2^{2d} - 1.$$

Proof.

$$\mathrm{Var}\left[\sum_{t=1}^{2n} X_t\right] = 2\,\mathrm{Var}\left[\sum_{t=1}^{n} X_t\right] + 2\,\mathrm{Cov}\left[\sum_{t=1}^{n} X_t, \sum_{t=n+1}^{2n} X_t\right].$$

Thus,

$$\mathrm{Cov}\left[\sum_{t=1}^{n} X_t, \sum_{t=n+1}^{2n} X_t\right] = 0.5\left(\mathrm{Var}\left[\sum_{t=1}^{2n} X_t\right] - 2\mathrm{Var}\left[\sum_{t=1}^{n} X_t\right]\right).$$

The result follows by noting that $\lim_{n\to\infty} n^{-2d-1}\mathrm{Var}\left(\sum_{t=1}^{n} X_t\right) = C$, where $C \neq 0$. □

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

This theorem has an interesting practical interpretation. If we write $X_k = N[k\Delta t] - N[(k-1)\Delta t]$ where $\Delta t > 0$ is fixed, then X_k represents the number of events (count) in a time interval of width Δt, e.g., 1 minute. Thus, $\sum_{k=1}^{n} X_k$ is the number of events in a time interval of length n minutes, e.g., 1 day. The theorem implies that as the level of aggregation (n) increases, the lag-1 autocorrelation of the aggregated counts will approach a nonzero constant if and only if the nonaggregated count series $\{X_k\}$ has long memory. In other words, the only way for a series of counts over a long time period to have nontrivial autocorrelation is for the short-term counts to have long memory. Because in practice long-term counts do have substantial autocorrelation (see Deo et al., 2007), it is important to use only the models for durations that imply long memory in the counting process. Examples of such models include the LMSD model (see Theorem 2) and also include ACD models with infinite variance (see Daley et al., 2000).

4. THE LINK BETWEEN COUNTS AND REALIZED VOLATILITY

To establish a link between counts and volatility it is necessary first to assume a continuous-time model for prices that incorporates the counting process. We will start with a simple pure-jump model (to be generalized subsequently),

$$\log P(t) = \log P(0) + \sum_{j=1}^{N(t)} \xi_j, \tag{1}$$

where $P(t)$ is the price at time t, $N(t)$ is the number of transactions up to time t, and the $\{\xi_j\}$ are i.i.d., independent of $N(\cdot)$, with zero mean, variance $\mu_2 < \infty$, and fourth moment $\mu_4 < \infty$.

Models related to (1) have been considered in the economic literature. Clark (1973) wrote the model

$$\log P(t) = \log P(0) + B(\tilde{N}(t)),$$

where B is Brownian motion and $\tilde{N}(t)$ is a nondecreasing positive stochastic process with independent increments, independent of B. Our model generalizes that of Clark (1973) in that it allows for nonindependent increments of $N(\cdot)$ and non-Gaussian price changes; however in our model $N(\cdot)$ is restricted to be a pure-jump process because it is derived from the underlying duration process. Press (1967) considered a model of form (1) but included an additional continuous component, assumed normality of price changes, and assumed that $N(\cdot)$ is a Poisson process. Oomen (2006) has generalized the pure-jump version of the model of Press (1967) to allow for time-varying intensity of the Poisson process and to allow for nonindependent price changes so as to describe market microstructure effects (see the discussion that follows).

The existing literature on realized volatility generally assumes that the logarithmic price process is a diffusion given by

$$\log P(t) = \log P(0) + \int_0^t \sigma(u)\, dW(u),$$

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

where W is a Brownian motion and σ, the instantaneous volatility, is a positive càdlàg process. Though the diffusive and pure-jump frameworks appear to be very different, Oomen (2006) points out that they may have similar implications for realized volatility. Nevertheless, the diffusive models by themselves do not yield a mechanism for generating transaction times or, therefore, durations and counts. We choose to use a pure-jump framework here because our goal in this paper is to link the properties of the observable durations, counts, and realized volatility.

In (1) and henceforth, we adopt the convention that a sum is taken to be zero if the upper limit is less than the lower limit. Because the price changes are independent of each other and of the process $N(\cdot)$, the log prices under (1) are a martingale.

The model (1) implies that the returns r_s at equally spaced clock-time intervals of width $\Delta t > 0$ may be expressed as

$$r_s = \sum_{j=N[(s-1)\Delta t]+1}^{N(s\Delta t)} \xi_j, \qquad s = 1, 2, \ldots. \tag{2}$$

Recall that the counts $\{\Delta N_s\}$ are given by $\Delta N_s = N(s\Delta t) - N[(s-1)\Delta t]$, i.e., the number of events occurring within the given equally spaced intervals of clock time. The following theorem shows that under the model (1) the memory parameter of the durations propagates unchanged to the realized volatility, $\sum_{s=1}^{n} r_s^2$.

THEOREM 5. *For durations $\{\tau_k\}$ satisfying the assumptions of Theorem 1 with $d \in [0, \frac{1}{2})$, we have*

$$\operatorname{Var}\left(\sum_{s=1}^{n} r_s^2\right) \sim Cn^{2d+1}$$

as $n \to \infty$, where $\{r_s\}$ are the returns given by (2).

Remark 5. The conclusions of Theorem 5 would continue to hold if instead of making assumptions on the durations, we assume that $\operatorname{Var} N(t) \sim Ct^{2d+1}$.

Remark 6. It follows from Theorem 5 that if $N(\cdot)$ is a Poisson process, the realized volatility has a variance that is asymptotically proportional to n, i.e., short memory.

The model (1) implies that clock-time returns are uncorrelated, whereas there is considerable empirical evidence to the contrary (see, e.g., Roll, 1984). This autocorrelation is often attributed to microstructure effects such as bid-ask bounce. We therefore consider a generalization that allows returns in tick time to be autocorrelated, which in turn allows for autocorrelation in the clock-time returns. The generalization we consider is in keeping with the literature for both pure-jump models (Oomen 2006) and diffusion models (Bandi and Russell, 2006; Hansen and Lunde, 2006; Zhang, Mykland, and Aït-Sahalia, 2005). The generalization of

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

model (1) takes the form

$$\log P(t) = \log P(0) + \sum_{j=1}^{N(t)} (\xi_j + \eta_j), \tag{3}$$

where $\{\xi_j\}$ satisfy the assumptions given previously and $\{\eta_j\}$ is a stationary, zero-mean process with finite fourth moment, independent of the counting process $N(\cdot)$. We can view the process $\{\eta_j\}$ as representing microstructure noise. Note that we do not require the $\{\eta_j\}$ to be independent of the efficient price shocks $\{\xi_j\}$. Model (3) covers the case considered by Oomen (2006), who assumed that the $\{\eta_j\}$ are a Gaussian qth-order moving average with respect to the difference of an i.i.d. process.

We have so far been unable to derive a result corresponding to Theorem 5 for the variance of the realized volatility based on the clock-time returns generated by the price process in the presence of microstructure noise given by (3). Obtaining such a result would require knowledge of the asymptotic behavior of the variance of the number of zero counts in a sequence of n counts, which is not known as far as we are aware. Although this is unfortunate, there are alternative sampling schemes for the construction of realized volatility that have been proposed in the literature, which have desirable properties. We consider the transaction-time sampling (TTS) scheme proposed by Oomen (2006), under which the returns are measured every K (tick-time) transactions and then the squares of these returns are aggregated to form a realized volatility. This contrasts with calendar time sampling (CTS) (i.e., sampling in fixed intervals of clock time), which results in the realized volatility studied in Theorem 5. Oomen (2006) found that TTS leads to realized volatility that has superior performance relative to that based on CTS.

Define $z_j = \xi_j + \eta_j$ for $j = 1, 2, \ldots$. Sampling every K transactions yields the sampled returns $\tilde{r}_j = z_{(j-1)K+1} + \cdots + z_{jK}$. Aggregating the available squared sampled returns up to time T yields the TTS realized volatility

$$\widetilde{RV}_T = \sum_{j=1}^{\lfloor N(T)/K \rfloor} \tilde{r}_j^2. \tag{4}$$

The following theorem shows that under the model (3) the memory parameter of the durations propagates unchanged to the TTS realized volatility, $\widetilde{RV}_T$.

THEOREM 6. *Let the durations $\{\tau_k\}$ satisfy the assumptions of Theorem 1 with $d \in [0, \frac{1}{2})$. Suppose that the lag-k autocovariances $\{\gamma_k\}$ of $\{\tilde{r}_j^2\}$ satisfy $\sum_{k=-\infty}^{\infty} |\gamma_k| < \infty$ for $d \in (0, \frac{1}{2})$ and the stronger condition $\sum_{k=-\infty}^{\infty} |k||\gamma_k| < \infty$ for $d = 0$. Then*

$$\operatorname{var}\left(\widetilde{RV}_T\right) \sim CT^{2d+1}$$

as $T \to \infty$ for some $C > 0$, where $\widetilde{RV}_T$ is defined in (4).

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

5. DISCUSSION

In their empirical study, Deo et al. (2007) found that the high-frequency data on durations, counts, and realized volatility possess long memory and the memory parameter is reasonably similar across the three series.[1] In this paper, we have established conditions on the durations that guarantee the propagation of the memory parameter, without change, from the durations to the counting process and then to the realized volatility resulting from the return models considered here. Our theoretical results imply that short-memory models, such as the finite-variance ACD model, and the SCD model of Bauwens and Veredas (2004), cannot generate long memory in the resulting counts or realized volatility.

Two duration models that can yield long memory in counts are the LMSD model (see Theorem 2) and ACD models with infinite variance (see Daley et al., 2000). Although these two models could both generate power law decay in the autocovariances of counts, the realizations in general would look very different. Under the infinite-variance ACD model, realizations of the counting process would tend to be dominated by a few very long durations, and hence the resulting series of counts would tend to have long strings of zeros. On the other hand, this behavior would not be displayed by the LMSD model. Nevertheless, there is no known statistical procedure for distinguishing between finite and infinite variance of durations based on runs of zeros in the counts. One could consider estimating the tail index directly from durations. Although there exist semiparametric tail index estimators, they are known to be badly behaved even when the series is independent (Resnick, 1997), to say nothing of strong dependence, as in the case of transaction durations.

We also presented a result implying that the only way for a series of counts aggregated over a long time period to have nontrivial autocorrelation is for the short-term counts to have long memory. In other words, aggregation ultimately destroys all autocorrelation in counts, if and only if the counts have short memory. Deo et al. (2007) found that the empirical lag-1 autocorrelations of the aggregated counts did not decay to zero with increasing aggregation. They also found, in accordance with Theorem 4, that the lag-1 autocorrelations of counts generated from simulated durations generated from the empirically estimated LMSD models did not decay to zero, whereas those from the estimated exponential and Weibull ACD models (whose estimated parameters implied finite variance though this restriction was not imposed in the estimation) did decay to zero, in contradiction to what was observed in the data. This lends further support to the conclusion that LMSD models are more appropriate for intertrade durations than finite-variance ACD models.

In principle, it may be more appropriate to try to compare the LMSD model with alternative long-memory models for durations rather than with the short-memory ACD model. Just as the LMSD model is a long-memory extension of the short-memory SCD model for durations, it is also possible to extend the ACD model to yield an observation-driven, long-memory model for durations.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Following Jasiak (1999), one could consider using a fractionally integrated ACD model that has long memory in the sense that the conditional mean is a linear combination of present and past durations with weights that decay hyperbolically. Such a duration process can be shown to be strictly stationary (under P^0) by a modification of the argument presented by Douc, Roueff, and Soulier (2008) (replacing the conditional variance by the conditional mean), who recently established the existence of a strictly stationary version of the fractionally integrated GARCH (FIGARCH) model of Baillie, Bollerslev, and Mikkelsen (1996), under strong conditions on the driving noise. One potential drawback to this approach, however, is that the durations would have an infinite unconditional mean, implying an extremely heavy tail, and as-yet unknown memory properties of the resulting counts. Another possibility might be to use a modified version of the fractionally integrated exponential GARCH (FIEGARCH) model (Bollerslev and Mikkelsen, 1996) for durations. It seems that such a model would face fewer difficulties in terms of the existence of unconditional moments and would be an interesting subject for future study.

NOTE

1. Nevertheless, the estimates of d for squared returns tend to be smaller than the corresponding estimates for the counts. Deo et al. (2007) conjecture that this apparent phenomenon may be illusory, in keeping with the general observation that the squared returns are a noisy proxy for volatility.

REFERENCES

Andersen, T.G. & T. Bollerslev (1997) Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance* 52, 975–1005.

Andersen, T.G., T. Bollerslev, F.X. Diebold, & P. Labys (2001) The distribution of exchange rate volatility. *Journal of the American Statistical Association* 96, 42–55.

Andrasfai, B. (1977) *Introductory Graph Theory*. Adam Hilger.

Baccelli, F. & P. Brémaud (2003) *Elements of Queueing Theory*. Springer-Verlag.

Baillie, R., T. Bollerslev, & H. Mikkelsen (1996) Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics* 74, 3–30.

Bandi, F.M. & J.R. Russell (2006) Separating microstructure noise from volatility. *Journal of Financial Economics* 79, 655–692.

Barndorff-Nielsen, O. & N. Shephard (2006) Impact of jumps on returns and realised variances: Econometric analysis of time-deformed Lévy processes. *Journal of Econometrics* 131, 217–252.

Bauwens, L. & D. Veredas (2004) The stochastic conditional duration model: A latent variable model for the analysis of financial durations. *Journal of Econometrics* 119, 381–412.

Billingsley, P. (1986) *Probability and Measure*, 2nd ed. Wiley.

Bollerslev, T. & D. Jubinski (1999) Equity trading volume and volatility: Latent information arrivals and common long-run dependencies. *Journal of Business & Economic Statistics* 17, 9–21.

Bollerslev, T. & H.O. Mikkelsen (1996) Modeling and pricing long memory in stock market volatility. *Journal of Econometrics* 73, 151–184.

Bradley, R. (2005) Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys* 2, 107–144.

Breidt, F.J., N. Crato, & P. de Lima (1998) The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* 83, 325–348.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Burkholder, D.L. (1966) Martingale transforms. *Annals of Mathematical Statistics* 37, 1494–1505.
Carrasco, M. & X. Chen (2002) Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* 18, 17–39.
Chung, K.L. (1974) *A Course in Probability Theory*, 2nd ed. Academic Press.
Clark, P.K. (1973) A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* 41, 135–155.
Comte, F. & E. Renault (1998) Long memory in continuous-time stochastic volatility models. *Mathematical Finance* 8, 291–323.
Daley, D.J., T. Rolski, & R. Vesilo (2000) Long-range dependent point processes and their Palm-Khinchin distributions. *Advances in Applied Probability* 32, 1051–1063.
Daley, D.J. & D. Vere-Jones (2003) *An Introduction to the Theory of Point Processes*, 2nd ed. Springer-Verlag.
Deo, R., M. Hsieh, & C.M. Hurvich (2007) Long Memory in Intertrade Durations, Counts and Realized Volatility of NYSE Stocks. Preprint, Stern School of Business, New York University.
Deo, R.S. & C.M. Hurvich (2001) On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory* 17, 686–710.
Deo, R., C. Hurvich, & Y. Lu (2006) Forecasting realized volatility using a long memory stochastic volatility model: Estimation, prediction and seasonal adjustment. *Journal of Econometrics* 131, 29–58.
Dittmann, I. & C.W.J. Granger (2002) Properties of nonlinear transformations of fractionally integrated processes. *Journal of Econometrics* 110, 113–133.
Douc, R., F. Roueff, & P. Soulier (2008) On the existence of some ARCH(∞) processes. *Stochastic Processes and Their Applications* 118, 755–761.
Doukhan, P. (1994) *Mixing*. Lecture Notes in Statistics 85. Springer-Verlag.
Engle, R. & J. Russell (1998) Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* 66, 1127–1162.
Epps, T.W. & M.L. Epps (1976) The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distribution hypothesis. *Econometrica* 44, 305–321.
Hansen, P.R. & A. Lunde (2006) Realized variance and market microstructure noise. *Journal of Business & Economic Statistics* 24, 127–161.
Harvey, A.C. (1998) Long memory in stochastic volatility. In J. Knight & S. Satchell (eds.), *Forecasting Volatility in Financial Markets*, pp. 307–320. Butterworth-Heinemann.
Hurvich, C.M., E. Moulines, & P. Soulier (2005) Estimating long memory in volatility. *Econometrica* 73, 1283–1328.
Ibragimov, I.A. & Y.V. Linnik (1971) *Independent and Stationary Sequences of Random Variables*, ed. J.F.C. Kingman. Wolters-Noordhoff.
Iglehart, D.L. & W. Whitt (1971) The equivalence of functional central limit theorems for counting processes and associated partial sums. *Annals of Mathematical Statistics* 42, 1372–1378.
Jasiak, J. (1999) Persistence in Intertrade Durations. Working paper, York University.
Nelson, D. (1990) Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* 6, 318–334.
Nieuwenhuis, G. (1989) Equivalence of functional limit theorems for stationary point processes and their palm distributions. *Probability Theory and Related Fields* 81, 593–608.
Oomen, R.C.A. (2006) Properties of realized variance under alternative sampling schemes. *Journal of Business & Economic Statistics* 24, 219–237.
Press, S.J. (1967) A compound events model for security prices. *Journal of Business* 40, 317–335.
Resnick, S.I. (1997) Heavy tail modeling and teletraffic data. With discussion and a rejoinder by the author. *Annals of Statistics* 25, 1805–1869.
Robinson, P.M. (1991) Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* 47, 67–84.
Robinson, P.M. & M. Henry (1999) Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels. *Econometric Theory* 15, 299–336.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Roll, R. (1984) A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance* 39, 1127–1139.

Rydberg, T.H. & N. Shephard (2003) Dynamics of trade-by-trade price movements: Decomposition and models. *Journal of Financial Econometrics* 1, 2–25.

Surgailis, D. & M.-C. Viano (2002) Long memory properties and covariance structure of the EGARCH model. *European Series in Applied and Industrial Mathematics: Probability and Statistics* 6, 311–329.

Tauchen, G. & M. Pitts (1983) The price variability-volume relationship on speculative markets. *Econometrica* 51, 485–505.

Tsay, R. (2002) *Analysis of Financial Time Series*. Wiley.

Viano, M.-C. C. Deniau, & G. Oppenheim (1995) Long-range dependence and mixing for discrete time fractional processes. *Journal of Time Series Analysis* 16, 323–338.

Yokoyama, R. (1980) Moment bounds for stationary mixing sequences. *Zeitschrift Für Wahrscheinlichskeitstheorie und Verwandte Gebiete* 52, 45–57.

Zhang, L., P.A. Mykland, & Y. Aït-Sahalia (2005) A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association* 100, 1394–1411.

APPENDIX: Proofs

Let P denote the stationary distribution of the point process N on the real line and let P^0 denote the corresponding Palm distribution. The distribution P determines and is completely determined by the stationary distribution P^0 of the doubly infinite sequence $\{\tau_k\}_{k=-\infty}^{\infty}$ of durations. Note that the counting process N is stationary under P and the durations are stationary under P^0, but in general there is no single distribution under which both the counting process and the durations are stationary. For more details on the correspondence between P and P^0, see Daley and Vere-Jones (2003), Baccelli and Brémaud (2003), and Nieuwenhuis (1989).

Following the standard notation for point processes on the real line (see, e.g., Nieuwenhuis, 1989, p. 594), we assume that the event times $\{t_k\}_{k=-\infty}^{\infty}$ satisfy

$$\ldots < t_{-1} < t_0 \leq 0 < t_1 < t_2 < \ldots.$$

Let

$$u_k = \begin{cases} t_1 & \text{if } k = 1 \\ \tau_k & \text{if } k \geq 2. \end{cases}$$

Here, the random variable $t_1 > 0$ is the time of occurrence of the first event following $t = 0$. For $t > 0$, define the count on the interval $(0, t]$, $N(t) := N(0, t]$, by

$$\begin{aligned} N(t) &= \max\{s : \sum_{i=1}^{s} u_i \leq t\}, \qquad u_1 \leq t \\ &= 0, \qquad\qquad\qquad\qquad\quad u_1 > t. \end{aligned}$$

Throughout the paper, the symbol $\Longrightarrow$ denotes weak convergence in the space $D[0, 1]$.

Proof of Theorem 1.

The proof proceeds by establishing the following facts.

(A) Y_n converges weakly under P to Brownian motion ($d = 0$) or fractional Brownian motion ($d > 0$).

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

(B) The standardized counting process converges weakly under P to a multiple of the same limit obtained in (A).

The theorem then follows by applying the uniform integrability of the standardized counting process established in Lemma 2 later in this Appendix, in conjunction with Theorem 25.12 of Billingsley (1986, p. 348).

We start by proving (A). By assumption (i), $Y_n \Longrightarrow \sigma B_{1/2+d}$ under P^0, where $\sigma > 0$. First, we will apply Theorem 6.3 of Nieuwenhuis (1989) to the durations $\{\tau_k\}_{k=-\infty}^{\infty}$ to conclude that $Y_n \Longrightarrow \sigma B_{1/2+d}$ under P. Because the $\{\tau_k\}_{k=-\infty}^{\infty}$ are stationary under P^0 and are generated by the shift to the first event following time zero (see Nieuwenhuis, 1989, p. 600), and because it follows from (ii) and Remark 4 that P^0 is $\{\tau_k\}$-mixing, his Theorem 6.3 applies. It follows that $Y_n \Longrightarrow \sigma B_{1/2+d}$ under P.

We next establish (B). Define

$$\tilde{Y}_n(s) = \frac{\sum_{k=1}^{\lfloor ns \rfloor}(u_k - \mu)}{n^{1/2+d}}, \qquad s \in [0,1].$$

Note that for all s, $\tilde{Y}_n(s) = Y_n(s) + n^{-(1/2+d)}(u_1 - \tau_1)$. From Baccelli and Brémaud (2003, eqn. (1.4.2), p. 33), for any measurable function h,

$$\mathrm{E}[h(\tau_1)] = \lambda \mathrm{E}^0[\tau_1 h(\tau_1)]. \tag{A.1}$$

Because $u_1 \leq \tau_1$, and because assumption (iii) implies that τ_1 has finite variance under P^0, using $h(x) = x$ in (A.1), it follows that $n^{-(1/2+d)}(u_1 - \tau_1)$ is $o_p(1)$ under P. Thus, $\tilde{Y}_n \Longrightarrow \sigma B_{1/2+d}$ under P.

Let

$$Z(t) = \frac{N(t) - t/\mu}{t^{1/2+d}}. \tag{A.2}$$

By Iglehart and Whitt (1971, Thm. 1), it follows that $Z(t) \xrightarrow{\mathcal{L}} \tilde{C} B_{1/2+d}(1)$ under P as $t \to \infty$, where $\tilde{C} > 0$.

Finally, by Lemma 2, $Z^2(t)$ is uniformly integrable under P, and hence by Theorem 25.12 of Billingsley (1986, p. 348), $\lim_t \mathrm{Var}[Z(t)] = \tilde{C}^2 \mathrm{Var}[B_{1/2+d}(1)]$. The theorem is proved. ■

Proof of Theorem 2. We simply verify that the conditions of Theorem 1 hold for this process.

Assume first that $d > 0$. By definition $\{\tau_k\}$ is stationary under P^0, and by Lemma 3, P^0 is $\{\tau_k\}$ mixing. By Surgailis and Viano (2002), $Y_n \Longrightarrow \sigma B_{1/2+d}$ under P^0, where $\sigma > 0$, and by Proposition 1, $\sup_n \mathrm{E}^0 \left| (\sum_{k=1}^{n}(\tau_k - \mu))/n^{1/2+d} \right|^p < \infty$ for all p. Thus, the result is proved for $d > 0$. If $d = 0$, the proof follows along the same lines as the proof of Theorem 3, because in this case under our assumptions on the model for $d = 0$ the process is exponential α-mixing (see, e.g., Doukhan, 1994, Cor. 1, Sect. 2.1, p. 58), and the weak convergence of Y_n follows again from Surgailis and Viano (2002) with $d = 0$. ■

Proof of Theorem 3. We simply verify that the conditions of Theorem 1 hold for this process.

By Lemma 3 in this Appendix, $\{\tau_k\}$ is exponential α-mixing, and hence strong mixing, and thus by Nieuwenhuis (1989), P^0 is $\{\tau_k\}$-mixing. Furthermore, because all moments of

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

τ_k exist up to order $3+\epsilon, \epsilon > 0$, we can apply results from Doukhan (1994, Thm. 1, Sect. 1.5, p. 46) to obtain

$$Y_n \Rightarrow CB, \tag{A.3}$$

if $\frac{1}{n}\mathrm{Var}(\sum_{k=1}^{n}\tau_k) \to C^2 > 0$, as $n \to \infty$.

It is well known that the GARCH(1,1) model can be represented as an ARMA(1,1) model; see Tsay (2002). Similarly, the ACD(1,1) model can also be reformulated as an ARMA(1,1) model,

$$\tau_k = \omega + (\alpha+\beta)\tau_{k-1} + (\eta_k - \beta\eta_{k-1}), \tag{A.4}$$

where $\eta_k = \tau_k - \psi_k$ is white noise with finite variance because $\mathrm{E}(\tau_k^{3+\epsilon}) < \infty$. The autoregressive and moving average parameters of the resulting ARMA(1,1) model are $(\alpha+\beta)$ and β, respectively.

It is also known that for any stationary invertible autoregressive moving average (ARMA) model $\{z_k\}$, $n\mathrm{Var}(\bar{z}) \to 2\pi f_z(0)$, where $f_z(0)$ is the spectral density of $\{z_k\}$ at zero frequency. For an ARMA(1,1) process, $f_z(0) > 0$ if the moving average coefficient is less than 1. Here, because $0 \le \beta < 1$, we obtain $\frac{1}{n}\mathrm{Var}(\sum_{k=1}^{n}\tau_k) = n\mathrm{Var}(\bar{\tau}) \to 2\pi f_\tau(0) > 0$, as $n \to \infty$. Therefore (A.3) follows.

Define $y_n = \frac{1}{\sqrt{n}}\sum_{k=1}^{n}\tau_k$. Because all moments of τ_k are bounded up to order $3+\epsilon$ ($\epsilon > 0$) under P^0, by Yokoyama (1980), we obtain

$$\mathrm{E}^0\{|y_n - \mathrm{E}^0(y_n)|^{3+\epsilon}\} \le K < \infty, \qquad \epsilon > 0 \tag{A.5}$$

uniformly in n, provided that $\{\tau_k\}$ is exponential α-mixing, which is proved in Lemma 3.

Therefore, we can apply Theorem 1 to the ACD(1,1) model, and the result follows. ■

Proof of Theorem 5. It follows from the proof of Lemma 2 that, for any fixed Δt and integer $m \ge 1$, under the assumptions of Theorem 1,

$$\mathrm{E}\{(\Delta N_s)^m\} = \mathrm{E}\{[N(\Delta t)]^m\} = C < \infty. \tag{A.6}$$

Denote $RV = \sum_{s=1}^{n} r_s^2$. By the law of total variance,

$$\mathrm{var}(RV) = \mathrm{E}\{\mathrm{var}[RV|N(\cdot)]\} + \mathrm{var}\{\mathrm{E}[RV|N(\cdot)]\}.$$

First, consider $\mathrm{Var}\{\mathrm{E}[RV|N(\cdot)]\}$. Because $\{\xi_i\}$ is i.i.d., we have

$$\mathrm{E}[r_s^2|N(\cdot)] = \mu_2 \Delta N_s,$$

where $\mu_2 = \mathrm{E}(\xi_i^2)$. Thus, by Theorem 1,

$$\mathrm{var}\{\mathrm{E}[RV|N(\cdot)]\} = \mathrm{var}\left\{\sum_{s=1}^{n}\mathrm{E}[r_s^2|N(\cdot)]\right\} = \mathrm{var}\{\mu_2 N(n\Delta t)\} \sim Cn^{2d+1}. \tag{A.7}$$

Next, we consider $\mathrm{E}\{\mathrm{var}[RV|N(\cdot)]\}$. Because $\mathrm{E}[r_s^4|N(\cdot)] \le \mu_4 \Delta N_s^4$, where $\mu_4 = \mathrm{E}(\xi_i^4)$, we obtain

$$\mathrm{var}[r_s^2|N(\cdot)] = \mathrm{E}[r_s^4|N(\cdot)] - \{\mathrm{E}[r_s^2|N(\cdot)]\}^2 \le \mu_4 \Delta N_s^4 - (\mu_2 \Delta N_s)^2. \tag{A.8}$$

Because $\{\xi_i\}$ is i.i.d., $\{r_s^2\}$ is serially independent, conditional on the counting process, $N(\cdot)$. By (A.6), (A.8), and the stationarity of $\{\Delta N_s\}$, we obtain

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

$$\mathrm{E}\{\mathrm{var}[RV|N(\cdot)]\} = \mathrm{E}\{\sum_{s=1}^{n} \mathrm{var}[r_s^2|N(\cdot)]\} = \sum_{s=1}^{n} \mathrm{E}\{\mathrm{var}[r_s^2|N(\cdot)]\} = nK. \tag{A.9}$$

where K is a positive constant.

The result follows from (A.7) and (A.9). ■

Proof of Proposition 1. We present the proof for the case $0 < d < \frac{1}{2}$. The proof for the case $d = 0$ follows along similar lines. Also, we assume here that p is a positive even integer. The result for all positive odd integers follows by Hölder's inequality.

Let $\tilde{y}_n = y_n - \mathrm{E}^0(y_n)$ and $\mu_\epsilon = \mathrm{E}^0(\epsilon_k)$. Write $\tilde{y}_n = \tilde{y}_{n1} + \tilde{y}_{n2}$, where

$$\tilde{y}_{n1} = \frac{\mu_\epsilon \sum_{k=1}^{n} (e^{h_k} - \mathrm{E}^0 e^{h_k})}{n^{d+1/2}}, \qquad \tilde{y}_{n2} = \frac{\sum_{k=1}^{n} e^{h_k}(\epsilon_k - \mu_\epsilon)}{n^{d+1/2}}.$$

Because $\tilde{y}_{n2}$ is a martingale, by Burkholder (1966, Thm. 9), for $1 < p < \infty$, there exists a finite and positive constant C_p depending only on p, such that

$$\mathrm{E}^0|\tilde{y}_{n2}|^p \le C_p \mathrm{E}^0 \left\{ \left[\sum_{k=1}^{n} \left(\frac{1}{n^{d+1/2}} e^{h_k}(\epsilon_k - \mu_\epsilon) \right)^2 \right]^{p/2} \right\}$$

$$= C_p \cdot \frac{1}{n^{(d+1/2)p}} \cdot \mathrm{E}^0 \left\{ \left[\sum_{k=1}^{n} e^{2h_k}(\epsilon_k - \mu_\epsilon)^2 \right]^{p/2} \right\}.$$

Using Minkowski's inequality and the independence between $\{h_k\}$ and $\{\epsilon_k\}$, we have

$$\mathrm{E}^0 \left\{ \left[\sum_{k=1}^{n} e^{2h_k}(\epsilon_k - \mu_\epsilon)^2 \right]^{p/2} \right\} \le \left(\sum_{k=1}^{n} \left\{ \mathrm{E}^0 \left([e^{2h_k}(\epsilon_k - \mu_\epsilon)^2]^{p/2} \right) \right\}^{2/p} \right)^{p/2}$$

$$= n^{p/2} \mathrm{E}^0(e^{ph_0}) \mathrm{E}^0(|\epsilon_0 - \mu_\epsilon|^p) = C n^{p/2}.$$

Thus $\mathrm{E}^0|\tilde{y}_{n2}|^p = O(n^{-pd}) = O(1)$. Therefore, by Minkowski's inequality, it is sufficient to show that $\mathrm{E}^0|\tilde{y}_{n1}|^p$ is bounded uniformly in n for all positive even integers $p \ge 2$.

Because $p \ge 2$ is even and $\mathrm{E}^0(\tilde{y}_{n1})^p$ can be expressed as a linear combination of the products of the joint cumulants of $\tilde{y}_{n1}$ of order $2, \ldots, p$, we have

$$0 \le \mathrm{E}^0|\tilde{y}_{n1}|^p = \mathrm{E}^0(\tilde{y}_{n1}^p) = \sum_{\pi} \left[c_\pi \prod_{j \in \pi} \mathrm{cum}(\underbrace{\tilde{y}_{n1}, \ldots, \tilde{y}_{n1}}_{j \text{ terms}}) \right]$$

$$\le \sum_{\pi} \left[|c_\pi| \prod_{j \in \pi} \left| \mathrm{cum}(\underbrace{\tilde{y}_{n1}, \ldots, \tilde{y}_{n1}}_{j \text{ terms}}) \right| \right],$$

where π ranges over the additive partitions of n and c_π is a finite constant depending on π.

Denote

$$y_{n1} = \frac{\mu_\epsilon \sum_{k=1}^{n} e^{h_k}}{n^{d+1/2}}.$$

Because the first-order cumulant of $\tilde{y}_{n1}$ is zero and for all integers $m \ge 2$, the mth-order cumulant of $\tilde{y}_{n1}$ is equal to that of y_{n1}, it suffices to show that the absolute value of the mth-order cumulant of y_{n1} is bounded uniformly in n under P^0, for all $m \in \{2, \ldots, p\}$.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

We first consider the second-and the third-order cumulants.

For the second-order cumulant ($m = 2$),

$$|\mathrm{cum}(y_{n1}, y_{n1})| = \mu_\epsilon^2 \left| \mathrm{cum}\left(\frac{\Sigma_{k=1}^n e^{h_k}}{n^{d+\frac{1}{2}}}, \frac{\Sigma_{s=1}^n e^{h_s}}{n^{d+\frac{1}{2}}} \right) \right| \leq \frac{\mu_\epsilon^2}{n^{2d+1}} \sum_{k=1}^n \sum_{s=1}^n \left| \mathrm{cum}(e^{h_k}, e^{h_s}) \right|.$$

By Surgailis and Viano (2002, Cor. 5.3),

$$\left| \mathrm{cum}\left(e^{h_k}, e^{h_s}\right) \right| = e^{\sigma_h^2} \left| e^{r_{|k-s|}} - 1 \right|,$$

where $r_{|k-s|} = \mathrm{cov}(h_k, h_s)$ and $\sigma_h^2 = \mathrm{Var}(h_k)$.

From our assumptions on $\{b_j\}$ in the definition of the LMSD process, it follows that $r_s \sim K s^{2d-1}$, as $s \to \infty$, where $K > 0$, so that

$$\begin{aligned} \sum_{k=1}^n \sum_{s=1}^n |e^{r_{|k-s|}} - 1| &\leq 2 \sum_{k=1}^n \sum_{s>k}^n |e^{r_{|k-s|}} - 1| + n|e^{r_0} - 1| \\ &\leq Kn \sum_{j=1}^n j^{2d-1} + n|e^{r_0} - 1| = O(n^{2d+1}). \end{aligned}$$

Thus $|\mathrm{cum}(y_{n1}, y_{n1})|$ is $O(1)$.

Next, for the third-order cumulant ($m = 3$), we have

$$\begin{aligned} |\mathrm{cum}(y_{n1}, y_{n1}, y_{n1})| &= \frac{\mu_\epsilon^3}{n^{3d+3/2}} \left| \sum_{k=1}^n \sum_{s=1}^n \sum_{u=1}^n \mathrm{cum}\left(e^{h_k}, e^{h_s}, e^{h_u}\right) \right| \\ &\leq \frac{\mu_\epsilon^3}{n^{3d+3/2}} \sum_{k=1}^n \sum_{s=1}^n \sum_{u=1}^n \left| \mathrm{cum}\left(e^{h_k}, e^{h_s}, e^{h_u}\right) \right|. \end{aligned}$$

By Surgailis and Viano (2002, Cor. 5.3),

$$\begin{aligned} &\sum_{k=1}^n \sum_{s=1}^n \sum_{u=1}^n \left| \mathrm{cum}\left(e^{h_k}, e^{h_s}, e^{h_u}\right) \right| \\ &\quad \leq \sum_{k=1}^n \sum_{s=1}^n \sum_{u=1}^n e^{\frac{3}{2}\sigma_h^2} |e^{r_{|k-s|}} - 1||e^{r_{|k-u|}} - 1||e^{r_{|s-u|}} - 1| \\ &\quad + \sum_{k=1}^n \sum_{s=1}^n \sum_{u=1}^n e^{\frac{3}{2}\sigma_h^2} |e^{r_{|k-s|}} - 1||e^{r_{|k-u|}} - 1| + \sum_{k=1}^n \sum_{s=1}^n \sum_{u=1}^n e^{\frac{3}{2}\sigma_h^2} |e^{r_{|k-s|}} - 1||e^{r_{|s-u|}} - 1| \\ &\quad + \sum_{k=1}^n \sum_{s=1}^n \sum_{u=1}^n e^{\frac{3}{2}\sigma_h^2} |e^{r_{|k-u|}} - 1||e^{r_{|s-u|}} - 1|. \end{aligned}$$

The last three summations are actually the same because of symmetry: we can simply relabel the indexes in the last summation by $s \leftrightarrow u$. As for the first summation, because

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

$|r_{|k-u|}| = |\mathrm{cov}(h_k, h_u)| \le \sigma_h^2 = \mathrm{Var}(h_k)$, we have $|e^{r_{|k-u|}} - 1| \le \left(e^{\sigma_h^2} + 1\right) < \infty$. So

$$\begin{aligned}
&\sum_{k=1}^{n}\sum_{s=1}^{n}\sum_{u=1}^{n}\left|\mathrm{cum}(e^{h_k}, e^{h_s}, e^{h_u})\right| \\
&\quad \le K\sum_{k=1}^{n}\sum_{s=1}^{n}\sum_{u=1}^{n} e^{\frac{3}{2}\sigma_h^2}|e^{r_{|k-s|}} - 1||e^{r_{|s-u|}} - 1| \\
&\qquad +3\sum_{k=1}^{n}\sum_{s=1}^{n}\sum_{u=1}^{n} e^{\frac{3}{2}\sigma_h^2}|e^{r_{|k-s|}} - 1||e^{r_{|s-u|}} - 1| \\
&\quad \le K\sum_{k=1}^{n}\sum_{s=1}^{n}\sum_{u=1}^{n}|e^{r_{|k-s|}} - 1||e^{r_{|s-u|}} - 1| \qquad \text{(for some } K > 0\text{)}. \\
&\quad = O(n^{4d+1}).
\end{aligned}$$

The last step follows from Lemma 1. So $\mu_\epsilon^3/n^{3d+\frac{3}{2}}\sum_{k=1}^{n}\sum_{s=1}^{n}\sum_{u=1}^{n}|\mathrm{cum}(e^{h_k}, e^{h_s}, e^{h_u})|$ converges to zero because $(4d+1) < (3d+3/2)$.

Next, we prove that the mth-order joint cumulant, which satisfies

$$|\mathrm{cum}(\underbrace{y_{n1}, \ldots, y_{n1}}_{m \text{ terms}})| \le \frac{\mu_\epsilon^m}{n^{m(d+\frac{1}{2})}}\sum_{k_1=1}^{n}\cdots\sum_{k_m=1}^{n}\left|\mathrm{cum}\left(e^{h_{k_1}}, \ldots, e^{h_{k_m}}\right)\right|, \tag{A.10}$$

converges to zero for all $m > 2$.

By Corollary 5.3 of Surgailis and Viano (2002), the absolute value of the mth joint cumulant, $\left|\mathrm{cum}\left(e^{h_{k_1}}, \ldots, e^{h_{k_m}}\right)\right|$, is bounded by a summation taken over all connected graphs with m vertices. Each entry of the summation is a product of terms of the form $\left|e^{r_{|k_i-k_j|}} - 1\right|$ along the edges that connect vertices k_i and k_j of a connected m-vertex graph.

For a graph with m vertices, we need at least $(m-1)$ edges to connect them. It is known (see Andrasfai, 1977, Ch. 2) that any connected m-vertex graph with $(m-1)$ edges may be represented as a *tree*. Let $W_{\{k_i,\ldots,k_j\}} < \infty$ be the total number of trees with vertices labeled by $k_i, k_{i+1}, \ldots, k_j$.

If a connected m-vertex graph used in applying Corollary 5.3 of Surgailis and Viano (2002) has more than $(m-1)$ edges, it is not a tree, and there will be more than $(m-1)$ terms of the form $\left|e^{r_{|k_i-k_j|}} - 1\right|$ being multiplied together in the m-fold summation in (A.10). But, for all k_i, k_j, $|r_{|k_i-k_j|}| = |\mathrm{cov}(h_{k_i}, h_{k_j})| \le \sigma_h^2 = \mathrm{Var}(h_{k_i})$, and so $|e^{r_{|k_i-k_j|}} - 1|$ $\le \left(e^{\sigma_h^2} + 1\right) < \infty$, and for any connected m-vertex graph with more than $(m-1)$ edges, there exists an m-vertex subgraph that has a tree representation. So we can retain a product of $(m-1)$ terms of the form $\left|e^{r_{|k_i-k_j|}} - 1\right|$ in the m-fold summation in (A.10) and move remaining terms out of the summation, bounding each by $\left(e^{\sigma_h^2} + 1\right)$. The resulting product of $(m-1)$ terms of the form $\left|e^{r_{|k_i-k_j|}} - 1\right|$ is itself a product over the edges of an m-vertex tree.

In all, $\left|\mathrm{cum}\left(e^{h_{k_1}}, \ldots, e^{h_{k_m}}\right)\right|$ is bounded by a constant times a summation over the set $G_{\{k_1,\ldots,k_m\}}$ of all $W_{\{k_1,\ldots,k_m\}}$ trees. Each entry of the summation is a product of terms of the form $\left|e^{r_{|k_i-k_j|}} - 1\right|$ being multiplied over the $(m-1)$ edges of the tree. Thus, we have

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

$$\sum_{k_1=1}^{n} \cdots \sum_{k_m=1}^{n} \left|\operatorname{cum}(e^{h_{k_1}}, \ldots, e^{h_{k_m}})\right|$$

$$\leq K \sum_{k_1=1}^{n} \cdots \sum_{k_m=1}^{n} \left\{ \sum_{G_{\{k_1,\ldots,k_m\}}} \prod_{(k_i,k_j)\in\Omega(G_{\{k_1,\ldots,k_m\}})} |e^{r_{|k_i-k_j|}} - 1| \right\}, \qquad (K > 0)$$

$$= K \sum_{G_{\{k_1,\ldots,k_m\}}} \sum_{k_1=1}^{n} \cdots \sum_{k_m=1}^{n} \underbrace{\left\{ \prod_{(k_i,k_j)\in\Omega(G_{\{k_1,\ldots,k_m\}})} |e^{r_{|k_i-k_j|}} - 1| \right\}}_{(m-1) \text{ terms}},$$

where $\Omega(G_{\{k_1,\ldots,k_m\}})$ is the set of edges of the graph indexed by $G_{\{k_1,\ldots,k_m\}}$.

By Lemma 1, each entry of the summation over $G_{\{k_1,\ldots,k_m\}}$ is of order $O\left(n^{2dm-2d+1}\right)$. Also, this summation is taken over a finite number of graphs ($W_{\{k_1,\ldots,k_m\}} < \infty$); therefore

$$\sum_{k_1=1}^{n} \cdots \sum_{k_m=1}^{n} \left|\operatorname{cum}\left(e^{h_{k_1}}, \ldots, e^{h_{k_m}}\right)\right| = O\left(n^{2dm-2d+1}\right).$$

Because the normalization term in (A.10) is of order $O\left(n^{m(d+1/2)}\right)$, and $m(d+1/2) > 2dm - 2d + 1$, we have $|\operatorname{cum}(y_{n1}, \ldots, y_{n1})|$ converges to zero for all $m > 2$. ■

LEMMA 1. *For any $M > 2$ and $0 < d < 1/2$,*

$$\underbrace{\sum_{k_1=1}^{n} \cdots \sum_{k_{M=1}}^{n}}_{M\text{-fold}} \underbrace{\left\{ \prod_{(k_i,k_j)\in\Omega(G)} \left|e^{r_{|k_i-k_j|}} - 1\right| \right\}}_{(M-1)\ terms} = O(n^{2d(M-1)+1}), \qquad \textbf{(A.11)}$$

where $\Omega(G)$ is the set of edges of G, G is any connected M-vertex graph with vertices $\{k_1, \ldots, k_M\}$ and $(M-1)$ edges; $r_{|k_i-k_j|} = \operatorname{cov}(h_{k_i}, h_{k_j})$, $1 \leq i \leq M$, $1 \leq j \leq M$, $\{h_{k_i}\}$ is a long-memory process with memory parameter d.

Proof. Because G is a connected graph with M vertices and $(M-1)$ edges, it can be represented as a tree (see Andrasfai, 1977, Ch. 2). The tree representation is not unique. Fix a particular representation. Then there is one vertex with no parent, called the *root*. A vertex with both a parent and a child is called a *node*. A vertex with no child is called a *leaf*.

We proceed iteratively. First, select any leaf vertex. By definition of a leaf, the corresponding index only appears once in the product, and so the sum on this index can be evaluated for this term only, holding the other terms fixed. Because $r_s \sim Cs^{2d-1}$ as $s \to \infty$, we have for any fixed integer i with $1 \leq i \leq n$, $\sum_{j=1}^{n} |e^{r_{|i-j|}} - 1| = O(n^{2d})$.

It follows that the sum on the first index is $O(n^{2d})$. Next, delete the leaf just used from the tree. The resulting graph is again a tree. Repeat the process of selecting a leaf, performing the corresponding sum, and deleting the leaf until only the root remains. The M-fold sum in (A.11) is now bounded by a constant times the sum of n terms each of which is $O\left(n^{2d(M-1)}\right)$. Thus, the sum in (A.11) is $O\left(n^{2d(M-1)+1}\right)$. □

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

LEMMA 2. *For durations* $\{\tau_k\}$ *satisfying the assumptions of Theorem 1, there exists* $\delta > 0$ *such that*

$$\sup_{t \geq 16} \mathrm{E}\left[|Z(t)|^{2+\delta}\right] < \infty,$$

where $Z(t)$ *is defined in equation (A.2).*

Proof. Select any $\delta \in (0, \epsilon)$. Let $\theta = 1/(2+\delta)$ and assume without loss of generality that $\mu = 1$ and hence that $\lambda = \frac{1}{\mu} = 1$. By Chung (1974, Thm. 3.2.1, p. 42), $\mathrm{E}[|Z(t)|^{2+\delta}] \leq 1 + \sum_{s=1}^{\infty} P[|Z(t)|^{2+\delta} \geq s]$. Thus, it suffices to show that

$$\sup_{t \geq 16} \sum_{s=1}^{\infty} P\left[|Z(t)|^{2+\delta} \geq s\right] < \infty. \tag{A.12}$$

Note that for any positive real k,

$$N(t) \geq k \Longleftrightarrow \sum_{i=1}^{\lceil k \rceil} u_i \leq t. \tag{A.13}$$

For any $s \geq 1$, we have

$$P\left[|Z(t)|^{2+\delta} \geq s\right] = P\left[Z(t) \leq -s^{\theta}\right] + P\left[Z(t) \geq s^{\theta}\right]. \tag{A.14}$$

Using the relationship (A.13), we obtain

$$P\left[Z(t) \leq -s^{\theta}\right] = P\left[N(t) \leq t - s^{\theta} t^{1/2+d}\right]$$

$$\leq P\left[N(t) < t - s^{\theta} t^{1/2+d} + 1\right] = \begin{cases} P(\sum_{i=1}^{\lceil \upsilon(t,s) \rceil} u_i > t), & s \leq a(t); \\ 0, & s > a(t), \end{cases} \tag{A.15}$$

where $a(t) = t^{(1/2-d)(2+\delta)}$ and $\upsilon(t,s) = t - s^{\theta} t^{1/2+d} + 1$. Similarly,

$$P\left[Z(t) \geq s^{\theta}\right] = P\left[N(t) \geq t + s^{\theta} t^{1/2+d}\right] = P\left(\sum_{i=1}^{\lceil g(t,s) \rceil} u_i \leq t\right), \tag{A.16}$$

where $g(t,s) = t + s^{\theta} t^{1/2+d}$.

Next, we show that both (A.15) and (A.16) are summable in s, uniformly in t. We treat the cases $d \in (0, 1/2)$ and $d = 0$ separately, because in condition (iii) of Theorem 1, we assume $p = 3 + \epsilon$ for $d = 0$, which is much weaker than what we assume for $d \in (0, 1/2)$.

Case I. $d \in (0, 1/2)$.

First, we consider (A.15). Suppose first that $s \leq a(t)$, so that $\upsilon(s,t) \geq 1$. Let

$$W = \frac{\sum_{i=1}^{\lceil \upsilon(t,s) \rceil} u_i - \lceil \upsilon(t,s) \rceil}{\lceil \upsilon(t,s) \rceil^{1/2+d}}.$$

Then, for $t \geq 16$, because $s^{\theta} t^{1/2+d} - 2 \geq \frac{1}{2} s^{\theta} t^{1/2+d}$ and $\lceil \upsilon(t,s) \rceil \leq \upsilon(t,s)$,

$$P\left(\sum_{i=1}^{\lceil \upsilon(t,s) \rceil} u_i > t\right) = P\left(\frac{\sum_{i=1}^{\lceil \upsilon(t,s) \rceil} u_i - \lceil \upsilon(t,s) \rceil}{\lceil \upsilon(t,s) \rceil^{1/2+d}} > \frac{t - \lceil \upsilon(t,s) \rceil}{\lceil \upsilon(t,s) \rceil^{1/2+d}}\right)$$

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

$$\leq P\left(W > \frac{t - v(t,s) - 1}{[v(t,s)+1]^{1/2+d}}\right)$$

$$\leq \mathrm{E}(|W|^{2+\epsilon}) \cdot \frac{(t - s^{\theta} t^{1/2+d} + 2)^{(2+\epsilon)(1/2+d)}}{\left(s^{\theta} t^{1/2+d} - 2\right)^{2+\epsilon}}$$

$$\leq \mathrm{E}(|W|^{2+\epsilon}) \cdot \frac{(2t)^{(2+\epsilon)(1/2+d)}}{\left(\frac{1}{2} s^{\theta} t^{1/2+d}\right)^{2+\epsilon}} = C \frac{\mathrm{E}(|W|^{2+\epsilon})}{s^{(2+\epsilon)/(2+\delta)}}, \tag{A.17}$$

where $C > 0$ is a constant.

For $s > a(t)$, $P[Z(t) \leq -s^{1/(2+\delta)}] = 0$.

Because $0 < \delta < \epsilon$, the right-hand side of (A.17) is summable in s, uniformly in t, provided that $\sup_{t \geq 1, s \geq 1} \mathrm{E}(|W|^{2+\epsilon}) < \infty$, which we show next.

Define

$$B_1 = \frac{u_1 - 1}{\lceil v(t,s) \rceil^{1/2+d}}, \qquad B_2 = \frac{\sum_{i=2}^{\lceil v(t,s) \rceil} (\tau_i - 1)}{\lceil v(t,s) \rceil^{1/2+d}},$$

so that $W = B_1 + B_2$. By Minkowski's inequality,

$$\mathrm{E}[|W|^{2+\epsilon}] \leq \left[\left(\mathrm{E}|B_1|^{2+\epsilon}\right)^{1/(2+\epsilon)} + \left(\mathrm{E}|B_2|^{2+\epsilon}\right)^{1/(2+\epsilon)}\right]^{2+\epsilon}.$$

Because $u_1 \leq \tau_1$, using $h(x) = (x+1)^{2+\epsilon}$ in $\mathrm{E}[h(\tau_1)] = \mathrm{E}^0[\tau_1 h(\tau_1)]$ (see Baccelli and Brémaud, 2003, eqn. (1.4.2), p. 33), and because our assumptions imply that $\mathrm{E}^0[\tau_k^{3+\epsilon}] < \infty$, we obtain

$$\sup_{t \geq 1, s \geq 1} \mathrm{E}|B_1|^{2+\epsilon} < \infty.$$

From Baccelli and Brémaud (2003, eqn. (1.2.25), p. 20), for any measurable function h,

$$\mathrm{E}[h(\tau_2, \ldots, \tau_n)] = \mathrm{E}^0[\tau_1 h(\tau_2, \ldots, \tau_n)].$$

This, together with Hölder's inequality, yields

$$\mathrm{E}|B_2|^{2+\epsilon} = \mathrm{E}^0\left(\tau_1 |B_2|^{2+\epsilon}\right) \leq \left[\mathrm{E}^0(\tau_1^{3+\epsilon})\right]^{1/(3+\epsilon)} \left[\mathrm{E}^0|B_2|^{3+\epsilon}\right]^{(2+\epsilon)/(3+\epsilon)}. \tag{A.18}$$

By condition (iii) of Theorem 1, $\sup_{t \geq 1, s \geq 1} \mathrm{E}^0|B_2|^{3+\epsilon} < \infty$; thus we have $\sup_{t \geq 1, s \geq 1} \mathrm{E}|B_2|^{2+\epsilon} < \infty$, and hence

$$\sup_{t \geq 1, s \geq 1} \mathrm{E}|W|^{2+\epsilon} = \mathrm{E}\left|\frac{\sum_{i=1}^{\lceil v(t,s) \rceil} u_i - \lceil v(t,s) \rceil}{\lceil v(t,s) \rceil^{1/2+d}}\right|^{2+\epsilon} < \infty. \tag{A.19}$$

Therefore,

$$\sup_{t \geq 1} \sum_{s=1}^{\infty} P[Z(t) \leq -s^{\theta}] = \sup_{t \geq 1} \sum_{s=1}^{\infty} P\left(\sum_{i=1}^{\lceil v(t,s) \rceil} u_i > t\right) < \infty. \tag{A.20}$$

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Next, we consider (A.16). Defining

$$U = \frac{\sum_{i=1}^{\lceil g(t,s)\rceil} u_i - \lceil g(t,s)\rceil}{\lceil g(t,s)\rceil^{1/2+d}},$$

an argument similar to that in (A.17) gives

$$P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) \le \mathrm{E}(|U|^m)\frac{\left(t + s^\theta t^{1/2+d}\right)^{m(1/2+d)}}{\left(s^\theta t^{1/2+d} - 1\right)^m}$$

for any $m > 0$.

For $t \ge 16$, because $s^\theta t^{1/2+d} - 2 \ge \frac{1}{2} s^\theta t^{1/2+d}$ and $t^{1/2+d} < t$, we obtain for all $m > 0$

$$P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) \le \mathrm{E}(|U|^m)\frac{\left(2t + s^\theta t^{1/2+d}\right)^{m(1/2+d)}}{\left(\frac{1}{2} s^\theta t^{1/2+d}\right)^m} \le C\frac{\mathrm{E}(|U|^m)}{s^{\theta m(1/2-d)}}, \tag{A.21}$$

where C is a constant.

Because $d \in (0, 1/2)$, we can choose m sufficiently large so that (A.21) is summable in s. By a similar argument as in the proof of (A.19), we have for this same value of m that $\sup_{t\ge1, s\ge1} \mathrm{E}(|U|^m) < \infty$. Therefore,

$$\sup_{t\ge1}\sum_{s=1}^{\infty} P\left[Z(t) \ge s^\theta\right] = \sup_{t\ge1}\sum_{s=1}^{\infty} P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) < \infty. \tag{A.22}$$

Case II. $d = 0$.

The bound for $P\left[Z(t) \le -s^\theta\right]$ for Case II follows along similar lines as in Case I, replacing d by 0 in the proof. Next, we obtain a bound for $P\left[Z(t) \ge s^\theta\right]$. Because $d = 0$, $g(t,s) = t + s^\theta t^{1/2}$.

Let integer $s_0 = \lfloor K t^{(2+\delta)/2} \rfloor$ for some K to be chosen later. Consider

$$\sup_{t\ge1}\sum_{s=1}^{\infty} P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) \le \sup_{t\ge1}\sum_{s=1}^{s_0} P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) + \sup_{t\ge1}\sum_{s=s_0}^{\infty} P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right). \tag{A.23}$$

The first term on the right-hand side of (A.23) involves summation in s from 1 to s_0. By (A.21) with $m = 2 + \epsilon$, we obtain

$$P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) \le \mathrm{E}(|U|^{2+\epsilon})\frac{\left(t + s^\theta t^{1/2} + 1\right)^{(2+\epsilon)/2}}{\left(\frac{1}{2} s^\theta t^{1/2}\right)^{2+\epsilon}}. \tag{A.24}$$

As in the proof of (A.19), we obtain $\sup_{t\ge1, s\ge1} \mathrm{E}(|U|^{2+\epsilon}) < \infty$.

Because

$$\left(t + s^\theta t^{1/2} + 1\right)^{(2+\epsilon)/2} < C\left(t^{(2+\epsilon)/2} + s^{\theta(2+\epsilon)/2} t^{(2+\epsilon)/4}\right) \tag{A.25}$$

we obtain

$$P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) \le C\left[\frac{1}{s^{\theta(2+\epsilon)}} + \frac{1}{s^{\theta(2+\epsilon)/2} t^{(2+\epsilon)/4}}\right]. \tag{A.26}$$

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

The first term on the right-hand side of (A.26) is summable in s because $\theta(2+\epsilon) = \frac{(2+\epsilon)}{(2+\delta)} > 1$. As for the second term, because $s_0 = \lfloor Kt^{(2+\delta)/2} \rfloor \leq Kt^{(2+\delta)/2}$, we obtain

$$\frac{1}{t^{(2+\epsilon)/4}} \sum_{s=1}^{Kt^{(2+\delta)/2}} \frac{1}{s^{\theta(2+\epsilon)/2}} \leq \frac{C}{t^{(2+\epsilon)/4}} s^{-\theta(2+\epsilon)/2+1} \Big|_{s=\lfloor Kt^{(2+\delta)/2} \rfloor}$$
$$\leq \frac{C}{t^{(2+\epsilon)/4}} \left(Kt^{(2+\delta)/2}\right)^{-\theta(2+\epsilon)/2+1} = Ct^{(\delta-\epsilon)/2},$$

which is bounded uniformly in t. It follows that

$$\sum_{s=1}^{s_0} P\left(\sum_{i=1}^{\lceil g(t,s) \rceil} u_i \leq t\right) \leq C_1 + C_2,$$

where the constants C_1 and C_2 are free of t. Hence

$$\sup_{t \geq 1} \sum_{s=1}^{s_0} P\left(\sum_{i=1}^{\lceil g(t,s) \rceil} u_i \leq t\right) < \infty. \tag{A.27}$$

We now consider $P\left(\sum_{i=1}^{\lceil g(t,s) \rceil} u_i \leq t\right)$ for $s \geq s_0$. By equation (1.2.25) of Baccelli and Brémaud (2003)

$$P\left[Z(t) \geq s^\theta\right] = P\left(\sum_{i=1}^{\lceil g(t,s) \rceil} u_i \leq t\right)$$
$$= \mathrm{E}^0\left[\int_0^\infty I\ \{0 \leq x \leq \tau_1\} I\left\{\tau_1 - x + \sum_{i=2}^{\lceil g(t,s) \rceil} \tau_i \leq t\right\} dx\right]$$
$$\leq \mathrm{E}^0\left[\int_0^\infty I\{0 \leq x \leq \tau_1\} I\left\{\sum_{i=2}^{\lceil g(t,s) \rceil} \tau_i \leq t\right\} dx\right] = \mathrm{E}^0 \tau_1 I\left\{\sum_{i=2}^{\lceil g(t,s) \rceil} \tau_i \leq t\right\}. \tag{A.28}$$

We bound (A.28) by Hölder's inequality:

$$\mathrm{E}^0 \tau_1 I\left\{\sum_{i=1}^{\lceil g(t,s) \rceil} \tau_i \leq t\right\} \leq (\mathrm{E}^0 \tau_1^\alpha)^{1/\alpha} \left[P^0\left(\sum_{i=1}^{\lceil g(t,s) \rceil} \tau_i \leq t\right)\right]^{1/\beta}, \tag{A.29}$$

where $1/\alpha + 1/\beta = 1$, $\alpha > 0$, $\beta > 0$ and the values of α and β will be chosen later.

We now show that the term $\left[P^0\left(\sum_{i=1}^{\lceil g(t,s) \rceil} \tau_i \leq t\right)\right]^{1/\beta}$ on the right-hand side of (A.29) is bounded and summable in s. Note that, for all nonnegative integers i and all $a \geq 0$, $0 \leq \mathrm{E}^0(e^{-a\tau_i}) < \infty$ because, by the dominated convergence theorem, $\lim_{a \to 0} \mathrm{E}^0(e^{-a\tau_i}) = 1$ and $\lim_{a \to \infty} \mathrm{E}^0(e^{-a\tau_i}) = 0$. Hence

$$\exists\ 0 < a_0 < \infty, \quad \text{such that } \mathrm{E}^0(e^{-a_0\tau_i}) = e^{-1},$$

where a_0 is free of i because $\{\tau_i\}$ are identically distributed under P^0.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Then,

$$P^0\left(\sum_{i=1}^{\lceil g(t,s)\rceil}\tau_i \le t\right) = P^0\left(e^{((-1)/\lceil g(t,s)\rceil)\sum_{i=1}^{\lceil g(t,s)\rceil} a_0\tau_i} \ge e^{-a_0((t)/\lceil g(t,s)\rceil)}\right). \tag{A.30}$$

Because by Jensen's inequality,

$$e^{((-1)/\lceil g(t,s)\rceil)\sum_{i=1}^{\lceil g(t,s)\rceil} a_0\tau_i} \le \frac{1}{\lceil g(t,s)\rceil}\sum_{i=1}^{\lceil g(t,s)\rceil} e^{-a_0\tau_i}$$

we conclude from (A.30) that

$$\begin{aligned} P^0\left(\sum_{i=1}^{\lceil g(t,s)\rceil}\tau_i \le t\right) &\le P^0\left(\frac{1}{\lceil g(t,s)\rceil}\sum_{i=1}^{\lceil g(t,s)\rceil} e^{-a_0\tau_i} \ge e^{-a_0((t)/\lceil g(t,s)\rceil)}\right) \\ &= P^0\left(\frac{1}{\lceil g(t,s)\rceil}\sum_{i=1}^{\lceil g(t,s)\rceil} x_i \ge e^{-1}\left(e^{1-a_0((t)/\lceil g(t,s)\rceil)}-1\right)\right), \end{aligned} \tag{A.31}$$

where $x_i = e^{-a_0\tau_i} - e^{-1}$.

Now, we choose $K > \max(2, 2a_0^{2+\delta})$. Because $s \ge s_0 = \lfloor Kt^{(2+\delta)/2}\rfloor$,

$$\begin{aligned} \lceil g(t,s)\rceil - a_0 t &\ge \left(t + s^\theta t^{1/2}\right) - a_0 t \ge s^\theta t^{1/2} - a_0 t \ge \left(Kt^{(2+\delta)/2}-1\right)^\theta t^{1/2} - a_0 t \\ &\ge \left(\frac{1}{2}Kt^{(2+\delta)/2}\right)^\theta t^{1/2} - a_0 t = \left[\left(\frac{K}{2}\right)^\theta - a_0\right] t \qquad \text{(because } t \ge 1\text{)} \\ &\ge \gamma > 0. \end{aligned}$$

Hence, for every $s \ge s_0$, $e^{1-((a_0 t)/\lceil g(t,s)\rceil)} - 1 > 0$.

Also, for fixed t, as $s \to \infty$, $e^{1-((a_0 t)/\lceil g(t,s)\rceil)}$ is monotonically nondecreasing. Thus,

$$\begin{aligned} \inf_{s\ge s_0} e^{1-((a_0 t)/\lceil g(t,s)\rceil)} - 1 &\ge e^{1-((a_0 t)/\lceil g(t,s_0)\rceil)} - 1 \ge e^{1-((a_0 t)/g(t,s_0)-t)} - 1 \\ &\ge e^{1-a_0(\frac{2}{K})^\theta} - 1 = \xi, \end{aligned}$$

using the fact that, for $s_0 = \lfloor Kt^{(2+\delta)/2}\rfloor \ge Kt^{(2+\delta)/2} - 1$,

$$g(t,s_0) - t = s_0^\theta t^{1/2} \ge \left(Kt^{(2+\delta)/2}-1\right)^\theta t^{1/2} \ge \left(\frac{1}{2}Kt^{(2+\delta)/2}\right)^\theta t^{1/2} = \left(\frac{K}{2}\right)^\theta t$$

and for $K > 2a_0^{2+\delta}$, $0 < a_0\left(\frac{2}{K}\right)^\theta < 1$, so that $\xi > 0$.

Therefore, (A.31) becomes

$$\begin{aligned} P^0\left(\sum_{i=1}^{\lceil g(t,s)\rceil}\tau_i \le t\right) &\le P^0\left(\frac{1}{\lceil g(t,s)\rceil}\sum_{i=1}^{\lceil g(t,s)\rceil} x_i \ge e^{-1}\xi\right) \\ &\le \frac{1}{\lceil g(t,s)\rceil^{m/2}}\mathrm{E}^0\left|\frac{1}{\lceil g(t,s)\rceil^{1/2}}\sum_{i=1}^{\lceil g(t,s)\rceil} x_i\right|^m \frac{1}{(e^{-1}\xi)^m}, \end{aligned} \tag{A.32}$$

for any $m > 0$.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Note that $\mathrm{E}^0\left(x_i^k\right) < \infty$ for all positive integers k. Also $\{x_i\}$ is strong mixing because $\{\tau_i\}$ is; hence by Yokoyama (1980), for any $m > 0$

$$\mathrm{E}^0\left|\frac{1}{\lceil g(t,s)\rceil^{1/2}}\sum_{i=1}^{\lceil g(t,s)\rceil} x_i\right|^m < C,$$

and (A.32) yields

$$P^0\left(\sum_{i=1}^{\lceil g(t,s)\rceil}\tau_i \le t\right) \le \frac{C}{(t+s^\theta t^{1/2}-2)^{m/2}} \le \frac{C}{s^{m\theta}} \qquad \text{(because } t \ge 2\text{)}.$$

Thus,

$$\left[P^0\left(\sum_{i=1}^{\lceil g(t,s)\rceil}\tau_i < t\right)\right]^{1/\beta} \le \frac{C}{s^{m\theta/\beta}}. \tag{A.33}$$

Now, on the right-hand side of (A.29), we can choose $\alpha = 3+\epsilon > 0$ and $\beta = \alpha/(\alpha-1) > 0$. Given this choice of α, β, we then choose m sufficiently large in (A.33) so that it is summable in s. Therefore, the right-hand side of (A.29) is summable in s, uniformly in t.

This then implies

$$\sup_{t\ge 16}\sum_{s=s_0}^{\infty} P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i < t\right) < \infty. \tag{A.34}$$

In all, by (A.27) and (A.34), we obtain

$$\sup_{t\ge 16}\sum_{s=1}^{\infty} P\left[Z(t) \ge s^\theta\right] = \sup_{t\ge 16}\sum_{s=1}^{\infty} P\left(\sum_{i=1}^{\lceil g(t,s)\rceil} u_i \le t\right) < \infty. \tag{A.35}$$

□

LEMMA 3. *Under the LMSD model with memory parameter* $d \in \left[0, \frac{1}{2}\right)$, P^0 *is* $\{\tau_k\}$*-mixing. The durations* $\{\tau_k\}$ *generated by the* ACD*(1,1) model are exponential* α*-mixing.*

Proof. Under P^0, $\{h_k\}$ is a stationary Gaussian process with a log spectral density having an integral on $[-\pi, \pi]$ that is greater than $-\infty$, so that the innovation variance is positive. Because Gaussian processes are time reversible, it follows that we can represent $h_k = \sum_{j=0}^{\infty} a_j w_{k+j}$ where $\sum a_j^2 < \infty$ and $\{w_k\}$ is an i.i.d. Gaussian sequence. Arguing as in the proof of Theorem 17.3.1 of Ibragimov and Linnik (1971, pp. 311–312), replacing $\{\ldots w_{k-1}, w_k\}$ by $\{w_k, w_{k+1}, \ldots\}$, it follows that P^0 is $\{h_k\}$-mixing. Because the $\{\epsilon_k\}$ are i.i.d. it follows that P^0 is also $\{\epsilon_k\}$-mixing. Because for any process $\{\xi_k\}$, P^0 is $\{\xi_k\}$-mixing if and only if the future tail σ-field of $\{\xi_k\}$ is trivial (see, e.g., Nieuwenhuis, 1989, eqn. (3.3)), it follows from Lemma 4, which appears next in this Appendix, that P^0 is $\{\tau_k\}$-mixing, where $\tau_k = e^{h_k}\epsilon_k$.

For the ACD(1,1) model, by Proposition 17 of Carrasco and Chen (2002), $\{\tau_k\}$ is exponential β-mixing (or also called absolutely regular) if $\{\tau_0, \psi_0\}$ are initialized from the stationary distribution. Their result still holds for a doubly infinite sequence $\{\tau_k\}, k \in (-\infty, \infty)$. It is well known that β-mixing implies α-mixing (or strong mixing) (see Bradley, 2005, Sect. 2.1). Therefore, $\{\tau_k\}$ is also exponential α-mixing, which further implies $\{\tau_k\}$-mixing of P^0 for the ACD(1,1) model; see Nieuwenhuis, 1989, eqn. (3.5). □

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

LEMMA 4. *Let* $\{\xi_s\}$ *and* $\{\zeta_s\}$ *be two* independent *processes whose future tail* σ*-fields are trivial. Then the future tail* σ*-field of the process* $\{\xi_s, \zeta_s\}$ *is trivial.*

Proof. Define $\mathcal{S}_t = \sigma(\xi_s, s \geq t)$, $\mathcal{T}_t = \sigma(\zeta_s, s \geq t)$, and $\mathcal{U}_t = \sigma(\xi_s, \zeta_s, s \geq t)$. As pointed out by Ibragimov and Linnik (1971, p. 303) (for regularity), to prove that $\mathcal{U}_\infty$ is trivial, it suffices to prove that for all $\mathcal{U}_0$-measurable zero mean random variables η such that $\mathrm{E}[\eta^2] \leq 1$, $\mathrm{E}[\eta \mid \mathcal{U}_t]$ converges to zero in quadratic mean. By standard arguments, it suffices to prove this for a random variable η that can be expressed as $\eta = \eta_1 \eta_2$ with η_1 $\mathcal{S}_0$-measurable and η_2 $\mathcal{T}_0$-measurable and, without loss of generality, both with zero mean. Then, by independence of $\{\xi_s\}$ and $\{\zeta_s\}$,

$$\mathrm{E}[\eta \mid \mathcal{U}_t] = \mathrm{E}[\eta_1 \mid \mathcal{S}_t] \times \mathrm{E}[\eta_2 \mid \mathcal{T}_t].$$

Because $\mathcal{S}_\infty$ and $\mathcal{T}_\infty$ are trivial, both terms on the right-hand side of the preceding expression tend to zero in quadratic mean (q.m). By independence, their product also tends to zero in q.m. □

Proof of Theorem 6. We will separately consider the cases $d \in \left(0, \frac{1}{2}\right)$ and $d = 0$.

Case I. $d \in \left(0, \frac{1}{2}\right)$.

By the law of total variance, $\mathrm{var}(\widetilde{RV}_T) = \mathrm{E}\{\mathrm{var}[\widetilde{RV}_T | N(\cdot)]\} + \mathrm{var}\{\mathrm{E}[\widetilde{RV}_T | N(\cdot)]\}$. First, we consider $\mathrm{Var}\{\mathrm{E}[\widetilde{RV}_T | N(\cdot)]\}$. We have

$$\mathrm{E}[\widetilde{RV}_T | N(\cdot)] = \mathrm{E}\left[\sum_{j=1}^{\lfloor N(T)/K \rfloor} \tilde{r}_j^2 | N(\cdot)\right] = \mu_{\tilde{r}^2} \left\lfloor \frac{N(T)}{K} \right\rfloor, \qquad \text{where } \mu_{\tilde{r}^2} = \mathrm{E}(\tilde{r}_j^2). \quad \textbf{(A.36)}$$

By Theorem 1, we know that $\mathrm{Var}\left[\frac{N(T)}{K}\right] \sim CT^{2d+1}$, which goes to infinity as T increases. Because $\left\lfloor \frac{N(T)}{K} \right\rfloor = \frac{N(T)}{K} - A$, where the random variable A can take only K finite values in $[0, 1)$ and hence $\mathrm{var}(A)$ must be finite, by the Cauchy–Schwartz inequality, we obtain

$$\mathrm{var}\left\lfloor \frac{N(T)}{K} \right\rfloor \sim CT^{2d+1}. \quad \textbf{(A.37)}$$

Using (A.37) in (A.36), it follows that

$$\mathrm{var}\{\mathrm{E}[\widetilde{RV}_T | N(\cdot)]\} \sim C_1 T^{2d+1} \qquad (C_1 > 0). \quad \textbf{(A.38)}$$

Next, we consider $\mathrm{var}\{\mathrm{E}[\widetilde{RV}_T | N(\cdot)]\}$. Then

$$0 \leq \mathrm{var}[\widetilde{RV}_T | N(\cdot)] = \mathrm{var}\left[\sum_{j=1}^{\lfloor N(T)/K \rfloor} \tilde{r}_j^2 | N(\cdot)\right] = \sum_{|j| < \lfloor \frac{N(T)}{K} \rfloor} \left(\left\lfloor \frac{N(T)}{K} \right\rfloor - |j| \right) \gamma_j$$

$$\leq \left\lfloor \frac{N(T)}{K} \right\rfloor \sum_{|j| < \lfloor \frac{N(T)}{K} \rfloor} |\gamma_j| \quad \leq \quad C_2 \left\lfloor \frac{N(T)}{K} \right\rfloor,$$

where $C_2 > 0$. Because $\mathrm{E}\left\lfloor \frac{N(T)}{K} \right\rfloor = O(T)$,

$$\mathrm{E}\{\mathrm{var}[\widetilde{RV}_T | N(\cdot)]\} \leq C_2 \mathrm{E}\left\lfloor \frac{N(T)}{K} \right\rfloor = O(T); \quad \textbf{(A.39)}$$

therefore, by (A.38) and (A.39), we obtain $\mathrm{var}(\widetilde{RV}_T) \sim C_1 T^{2d+1}$.

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press

Case II. $d = 0$.

Equation (A.38) still holds when $d = 0$ because Theorem 1 includes the $d = 0$ case. Next, we consider $\mathrm{E}\{\mathrm{var}[\widetilde{RV}_T|N(\cdot)]\}$. Again,

$$\begin{aligned}\mathrm{var}[\widetilde{RV}_T|N(\cdot)] &= \sum_{|j|<\lfloor\frac{N(T)}{K}\rfloor}\left(\left\lfloor\frac{N(T)}{K}\right\rfloor - |j|\right)\gamma_j \\ &= \left\lfloor\frac{N(T)}{K}\right\rfloor\sum_{j=-\infty}^{\infty}\gamma_j - \left\lfloor\frac{N(T)}{K}\right\rfloor\sum_{|j|\geq\lfloor\frac{N(T)}{K}\rfloor}\gamma_j - \sum_{|j|<\lfloor\frac{N(T)}{K}\rfloor}|j|\gamma_j \\ &= \left\lfloor\frac{N(T)}{K}\right\rfloor 2\pi f_{\tilde{r}^2}(0) - \left\lfloor\frac{N(T)}{K}\right\rfloor\sum_{|j|\geq\lfloor\frac{N(T)}{K}\rfloor}\gamma_j - \sum_{|j|<\lfloor\frac{N(T)}{K}\rfloor}|j|\gamma_j,\end{aligned}$$

where $f_{\tilde{r}^2}(0)$ is the spectral density of $\{\tilde{r}_j^2\}$ at zero frequency.

We now consider the behavior of the three terms on the right-hand side of the preceding expression. First, by the renewal theorem,

$$\mathrm{E}\left\{\left\lfloor\frac{N(T)}{K}\right\rfloor 2\pi f_{\tilde{r}^2}(0)\right\} \sim \frac{2\pi f_{\tilde{r}^2}(0)\lambda}{K}T.$$

Next,

$$\begin{aligned}\left|\mathrm{E}\left\{\left\lfloor\frac{N(T)}{K}\right\rfloor\sum_{|j|\geq\lfloor\frac{N(T)}{K}\rfloor}\gamma_j\right\}\right| &\leq \mathrm{E}\left\{\left\lfloor\frac{N(T)}{K}\right\rfloor\sum_{|j|\geq\lfloor\frac{N(T)}{K}\rfloor}\frac{1}{|j|}|j||\gamma_j|\right\} \\ &\leq \sum_{j=-\infty}^{\infty}|j||\gamma_j| = O(1).\end{aligned}$$

Finally, the third term is bounded by

$$\left|\mathrm{E}\left\{\sum_{|j|<\lfloor\frac{N(T)}{K}\rfloor}|j|\gamma_j\right\}\right| \leq \sum_{j=-\infty}^{\infty}|j||\gamma_j| = O(1).$$

Thus, $\mathrm{E}\left\{\mathrm{var}[\widetilde{RV}_T|N(\cdot)]\right\} \sim 2\pi f_{\tilde{r}^2}(0)(\lambda T/K)$. This together with (A.38) proves the theorem for the case of $d = 0$. ■

https://doi.org/10.1017/S0266466608090294 Published online by Cambridge University Press