Limit Laws in Transaction-Level Asset Price Models*

Alexander Aue† Lajos Horváth‡ Clifford M. Hurvich§

draft of May 26, 2009

Abstract

We consider pure-jump transaction-level models for asset prices in continuous time, driven by point processes. In a bivariate model that admits cointegration, we allow for time deformations to account for such effects as intraday seasonal patterns in volatility, and non-trading periods that may be different for the two assets. Most assumptions are stated directly on the point process, though we provide sufficient conditions on the corresponding inter-trade durations for these assumptions to hold. We obtain the asymptotic distribution of the log-price process. We also obtain the asymptotic distribution of the ordinary least-squares estimator of the cointegrating parameter based on data sampled from an equally-spaced discretization of calendar time, in the case of weak fractional cointegration. Finally, we obtain the limiting distribution of the ordinary least-squares estimator of the autoregressive parameter in a simplified transaction-level univariate model with a unit root.

1 Introduction

The increasingly widespread availability of transaction-level financial price data motivates the development of models to describe such data, as well as theory for widely-used statistics of interest under the assumption of a given transaction-level generating mechanism. We focus here on a bivariate pure-jump model in continuous time for log prices proposed by Hurvich and Wang [19, 20] which yields fractional or standard cointegration. The motivation for using a pure-jump model is that observed price series are step functions, since no change is possible in observed prices during time periods when there are no transactions. Examples of data sets that would fit into the framework of this model include: buy prices and sell prices of a single stock; prices of two different stocks

*Research partially supported by NSF grants DMS 0604670 and DMS 0652420

†Department of Statistics, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA, email: alexaue@wald.ucdavis.edu

‡Department of Mathematics, University of Utah, 155 South 1440 East, Salt Lake City, UT 84112, USA, email: horvath@math.utah.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

within the same industry; stock and option prices of a given company; option prices on a given stock with different degrees of maturity or moneyness; corporate bond prices at different maturities for a given company; Treasury bond prices at different maturities.

Two basic questions that we address here are the asymptotic distribution of the log prices as time $t \to \infty$, and of the usual OLS estimator of the cointegrating parameter based on n observations of the log prices at equally-spaced time intervals as $n \to \infty$. Most of the existing methods for deriving such limit laws (see Phillips and Durlauf [27], Robinson and Marinucci [30]) cannot be applied here because the continuous-time log-price series are not diffusions and because the discretized log-price series are not linear in either an *iid* sequence, a martingale difference sequence or a strong mixing sequence. Nevertheless, it is of interest to know whether and under what conditions the existing limit laws, based, say, on linearity assumptions in discrete time, continue to hold under a transaction-level generating mechanism.

In the model of Hurvich and Wang [19, 20] the price process in continuous time is specified by a counting process giving the cumulative number of transactions up to time t, together with the process of changes in log price at the transaction times. This structure corresponds to the fact that most transaction-level data consists of a time stamp giving the transaction time as well as a price at that time. In such a setting, another observable quantity of interest is the *durations*, i.e., the waiting times between successive transactions of a given asset. There is a growing literature on univariate models for durations, including the seminal paper of Engle and Russell [16] on the autoregressive conditional duration models (ACD), as well as Bauwens and Veredas [4] on the stochastic duration model (SCD), and Deo *et al.* [12] on the long-memory stochastic duration model (LMSD).

It is known from the theory of point processes (see Daley and Vere-Jones [8, 9, 10] or Nieuwenhuis [25]) that the durations are stationary under a measure P^0 (known as the *palm distribution*) if and only if the point process is stationary under a measure P which determines and is uniquely determined by P^0, where in general P and P^0 are different. Deo *et al.* [14] showed that, subject to regularity conditions, if partial sums of centered durations, scaled by $n^{-(d+1/2)}$ with $d \in [0, 1/2)$, satisfy a functional central limit theorem under P^0 then the counting process $N(t)$ has long or short memory (for $d > 0$, $d = 0$, respectively) in the sense that $\text{Var}N(t) \sim Ct^{2d+1}$ under P as $t \to \infty$ (with $C > 0$), and they gave conditions under which this scaling would lead to long memory in volatility. It is important to note that since P and P^0 determine one another, there is no need to start with models for durations, as in most of the existing literature. Instead, as advocated by Bowsher [6], it is possible to start by specifying a model for the point process of events of interest (such as transactions). Such an approach appears to generalize more easily to the multivariate case, since attempting to directly model multivariate durations as a time series is problematic as they are not synchronized in transaction-time, e.g., the kth transactions of IBM and Apple occur at different times, and this time difference is itself a stochastic process with respect to k.

Hurvich and Wang [19, 20] did not derive limit laws for the log price series or the OLS estimator of the cointegrating parameter, but focused instead on properties of variances and covariances for log price series and returns, and on lower bounds on the rate of convergence for the OLS estimator. In this paper, for the model of [19, 20] but under assumptions that are more general than theirs, we obtain the limit law for log prices for standard, weak fractional and strong fractional cointegration, where the cointegrating error is integrated of order d_η with $d_\eta = 0$, $d_\eta \in (1/2, 1)$ or $d_\eta \in (0, 1/2)$, respectively, and for the OLS estimator of the cointegrating parameter in the case of weak fractional cointegration only. In our result on the limit law for log prices, Theorem 4.1, we allow for a stochastic time-varying intensity function in the counting processes. This allows for such effects as dynamic intraday seasonality in volatility (as observed, for example, in Deo *et al.* [13]), as well as fixed non-trading intervals such as holidays and overnight periods. We will also obtain the limiting distribution of the OLS estimator of the autoregressive parameter in a simplified transaction-level univariate model with a unit root.

The remainder of this paper is organized as follows. In Section 2 we write the model for the log price series and state our assumptions on the counting process, the time-deformation functions, and the return shocks. In Section 3, we explain how our assumptions on the counting process can be related to corresponding assumptions on the sequence of durations (inter-arrival times), with focus on some particular duration models that have been proposed in the literature. In Section 4, we provide our main results on the long-run behavior of the log-price process, on the OLS estimator for the cointegrating parameter under weak fractional cointegration (with some remarks on the strong fractional and standard cointegration cases), and on the OLS estimator for univariate unit-root autoregression. Section 5 provides proofs.

2 Transaction-level Model

In this section, we introduce our model and precisely state the assumptions made on the various model components. As in Hurvich and Wang [19, 20], we consider a bivariate pure-jump transaction-level price model that enables cointegration. We define the log-price process $y = (y_1, y_2) = (y(t)\colon t \geq 0)$ by

$$y_1(t) = \sum_{k=1}^{N_1(t)} (e_{1,k} + \eta_{1,k}) + \sum_{k=1}^{N_2(t_{1,N_1(t)})} (\theta e_{2,k} + g_{21}\eta_{2,k}), \tag{2.1}$$

$$y_2(t) = \sum_{k=1}^{N_2(t)} (e_{2,k} + \eta_{2,k}) + \sum_{k=1}^{N_1(t_{2,N_2(t)})} (\theta^{-1} e_{1,k} + g_{12}\eta_{1,k}), \tag{2.2}$$

where for $i = 1, 2$, $N_i(\cdot)$ are counting processes on the real line (see Daley and Vere-Jones [9], page 43) such that, for $t \geq 0$, $N_i(t) := N_i(0, t]$ gives the total number of transactions of Asset i in $(0, t]$, and $t_{i,k}$ is the clock time (calendar time) for the kth transaction of Asset i, with $\cdots t_{i,-1} \leq t_{i,0} \leq 0 < t_{i,1} \leq t_{i,2} \cdots$. The quantity $N_2(t_{1,N_1(t)})$ denotes the number of transactions of Asset 2 between time 0 and the time $t_{1,N_1(t)}$ of the most recent transaction of Asset 1, with an analogous interpretation for $N_1(t_{2,N_2(t)})$. The efficient shock sequences $\{e_{i,k}\}_{k=1}^{\infty}$ model the permanent component and the microstructure noise sequences $\{\eta_{i,k}\}_{k=1}^{\infty}$ model the transitory component of the log-price process. Efficient shock spillover effects are weighted by θ and θ^{-1}, thus yielding cointegration with cointegrating parameter θ, assumed nonzero, while the microstructure spillover is enabled through the quantities g_{12} and g_{21}. A detailed economic justification for this model, derivation of a common-components representation, as well as a comparison with certain discrete-time models, is given in [19, 20]. We shall work in this paper with the following set of assumptions, using the definition of Daley and Vere-Jones [9] (page 47) that a point process is *simple* if the probability is zero that there exists a time t at which more than one event occurs.

Assumption 2.1. (Counting Processes) *For $i = 1, 2$,*

$$N_i(t) = \tilde{N}_i(t + f_i(t)),$$

where under the measure P, $\tilde{N}_i(\cdot)$ is a simple, stationary and ergodic counting process on $\mathbb{R}$ with intensity $\tilde{\lambda}_i \in (0, \infty)$. The f_i are random or deterministic càdlàg functions such that $t + f_i(t)$ is nondecreasing with probability one, $t^{-1}[t + f_i(t)] \to \gamma_i \in (0, \infty)$ with probability one as $t \to \infty$, and

$$\sup_{t \geq 0} \left| f_i(t) - f_i(t^-) \right| \leq C$$

with probability one, where $C \in (0, \infty)$. Finally, $\tilde{N}_i$ and f_j $(i, j = 1, 2)$ need not be independent.

The functions f_i are 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 $t + f_i(t)$ has a periodic derivative (with period T and with probability one), for example, $f_i(t) = \sin(2\pi t/T)$. Fixed non-trading intervals, say, $t \in [T_1, T_2)$, could be accommodated by taking $f_i(t) = T_1 + f_i(T_1) - t$ for $t \in [T_1, T_2)$ so that $t + f_i(t)$ remains constant for t in this interval, and then taking $f_i(T_2) > T_1 + f_i(T_1) - T_2$ so that $t + f_i(t)$ 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 $[T_1, T_2)$.

The use of the time-varying intensity function f_i renders the counting process N_i nonstationary. We will show, however, in Lemma 5.1 below that N_i satisfies a renewal-type theorem. Since it is possible that f_i has (upward) jumps, the N_i may not be simple even though the $\tilde{N}_i$ are simple.

The counting processes N_i induce associated sequences of durations $\{\tau_{i,k}\}_{k=-\infty}^{\infty}$ defined by $\tau_{i,k} = t_{i,k} - t_{i,k-1}$, thus yielding the duality

$$N_i(t) = \begin{cases} \max\Big\{s : \sum_{k=1}^{s} u_{i,k} \leq t\Big\}, & u_{i,1} \leq t \\ 0, & u_{i,1} > t, \end{cases}$$

where

$$u_{i,k} = \begin{cases} t_{i,1}, & k = 1. \\ \tau_{i,k}, & k \geq 2. \end{cases}$$

Similarly, $\tilde{N}_i$ induces associated durations $\{\tilde{\tau}_{i,k}\}_{k=-\infty}^{\infty}$ satisfying a corresponding duality relation. Our approach here differs somewhat from the univariate duration-based approaches given in the papers Engle and Russell [16], Bauwens and Veredas [4], and Deo *et al.* [14]. These authors start out with the durations $\{\tau_{i,k}\}_{k=-\infty}^{\infty}$ and endow them with certain desirable properties such as stationarity, mixing and ergodicity with respect to the so-called Palm distribution P^0 (see Nieuwenhuis [25] for the definition). How these properties then propagate to the counting processes under P has been addressed in the recent article Deo *et al.* [14]. By contrast, in the current paper we typically start with assumptions on the counting processes $\tilde{N}_i$ rather than on the durations (see Assumption 2.1 above). Nevertheless, we also provide, in Section 3, conditions on the durations $\{\tilde{\tau}_{i,k}\}_{k=-\infty}^{\infty}$ that are sufficient for our assumptions on the counting processes $\tilde{N}_i$ to hold.

Assumption 2.2. (Efficient Shocks) *The efficient shocks $\{e_{i,k}\}_{k=1}^{\infty}$ form independent, identically distributed sequences of random variables with zero mean and finite variance $\sigma_{i,e}^2$.*

Although many of our results would continue to hold if the *iid* assumption above were replaced by a weak-dependence assumption (see Subsection 5.1 below), we maintain the *iid* assumption here in keeping with the economic motivation for the model as provided by Hurvich and Wang ([19]) that in the absence of the microstructure shocks each of the log price series would be a martingale with respect to its own past.

Assumption 2.3. (Microstructure Noise) *The microstructure noise $\{\eta_{i,k}\}_{k=1}^{\infty}$ is a zero mean sequence with memory parameter $d_\eta \in [-1, 0)$ that satisfies moreover one of the following conditions.*

(A) Weak Fractional Cointegration: *If* $d_\eta \in (-\frac{1}{2}, 0)$*, then* $\eta_{i,k} = \sum_{j=0}^{\infty} b_{i,j}\zeta_{i,k-j}$ *where* $b_{i,j} \sim C_i j^{d_\eta - 1}$ *as* $j \to \infty$ *with* $C_i \neq 0$*, and* $\{\zeta_{i,k}\}_{k=-\infty}^{\infty}$ *is iid with zero mean and finite variance. It is also required that* $E[|\zeta_{i,1}|^{2\nu}] < \infty$ *for some* $\nu \geq 1$ *satisfying the condition* $d_\eta \geq \frac{1}{\nu+2} - \frac{1}{2}$*. Finally,* $\{\eta_{i,k}\}$ *has spectral density* $g_{i,\eta}(\lambda)$ *such that, for some* $\beta \in (0, 2]$*,* $g_{i,\eta}(\lambda) = \sigma_{i,\eta}^2 C_\eta \lambda^{-2d_\eta}(1 + \mathcal{O}(\lambda^\beta))$ *holds as* $\lambda \to 0^+$*, where* $C_\eta = \{-4\Gamma(-1 - 2d_\eta)\sin(\pi(d_\eta + 1))\}^{-1}$ *and* $\sigma_{i,\eta}^2 > 0$ *is the long-run variance of* $\{\eta_{i,k}\}$*.*

(B) Strong Fractional Cointegration: *If* $d_\eta \in (-1, -\frac{1}{2})$*, then* $\eta_{i,k} = \varphi_{i,k} - \varphi_{i,k-1}$*,* $k = 1, 2, \ldots$*, where* $\varphi_{i,0} = 0$ *and* $\{\varphi_{i,k}\}_{k=1}^{\infty}$ *is a zero mean, strictly stationary long memory sequence with memory parameter* $d_\varphi = d_\eta + 1 \in (0, \frac{1}{2})$ *in the sense that its autocovariances satisfy* $\mathrm{Cov}(\varphi_{i,k}, \varphi_{i,k+h}) = K_i h^{2d_\varphi - 1} + \mathcal{O}(h^{2d_\varphi - 3})$ *for* $h \geq 1$ *and* $K_i > 0$*.*

(C) Standard Cointegration: *If* $d_\eta = -1$*, then* $\eta_{i,k} = \xi_{i,k} - \xi_{i,k-1}$*,* $k = 1, 2, \ldots$*, where* $\xi_{i,0} = 0$ *and* $\{\xi_{i,k}\}_{k=1}^{\infty}$ *is a zero mean, strictly stationary sequence with exponentially decaying autocovariances,* $|\mathrm{Cov}(\xi_{i,k}, \xi_{i,k+h})| \leq C_\xi e^{-K_\xi h}$ *for* $h \geq 0$ *and* $C_\xi, K_\xi > 0$*.*

Assumption 2.4. *The shocks* $\{e_{1,k}\}_{k=1}^{\infty}$*,* $\{e_{2,k}\}_{k=1}^{\infty}$*,* $\{\eta_{1,k}\}_{k=1}^{\infty}$ *and* $\{\eta_{2,k}\}_{k=1}^{\infty}$ *are mutually independent, and these are independent of the counting processes* N_i*,* $i = 1, 2$*.*

Assumption 2.4 was also made by Hurvich and Wang [19, 20]. Since the trades of Asset 1 are not synchronized in calendar time or in transaction time with those of Asset 2, it seems reasonable to assume that the two efficient shock series are mutually independent, as are the two microstructure shock series. Mutual independence of the efficient and microstructure shock series of a given asset can be justified on economic grounds, and is often made in the econometric literature for calendar-time models. See, e.g., Barndorff-Nielsen *et al.* [3].

We note that (N_1, N_2) together with the sequences of return shocks comprises a marked point process (see Daley and Vere-Jones [10], Nieuwenhuis [26]), where the shocks are marks. However, because Assumption 2.4 states that the shocks are independent of the N_i, it suffices for most of our discussion of point processes theory to focus on the non-marked point process (N_1, N_2). Still, it is worth mentioning that for the marked point process, the independence between (N_1, N_2) and the return shocks holds under both the Palm distribution P^0 and the time-stationary distribution P (see Nieuwenhuis [26]) of the marked point process.

The independence of (N_1, N_2) and the return shocks is restrictive. In particular, it implies that there can be no leverage effect in the returns (for example, a correlation between a return in one time period and a squared return in a subsequent time period). A transaction-level model yielding a leverage effect was proposed (but justified only with simulations) in Hurvich and Wang [20]. Models where the point process need not be independent of the return shocks were discussed in Prigent [29] in the context of option pricing with marked point processes.

Hurvich and Wang [19, 20] assumed $f_i \equiv 0$. They also assumed that the durations satisfied the conditions of Theorem 1 of Deo *et al.* [12] which entails finite moments of all orders for the durations under P^0. The conditions stated in this section allow $f_i \neq 0$ and also allow the durations to have infinite moment of order $1+\epsilon$ under P^0, for any $\epsilon > 0$.

3 From Durations to Counting Processes

Many existing transaction-level models start with assumptions on the durations $\{\tau_k\}_{k=-\infty}^{\infty}$ instead of imposing assumptions directly the corresponding counting processes. Duration-based approaches in econometrics have originally been used to examine the impact of past unemployment on current levels in Lancaster [23]. The first attempt in modeling tick-by-tick data on the basis of durations has led to the autoregressive conditional duration (ACD) model of Engle and Russell [16]. More recent contributions in the literature include the stochastic conditional duration (SCD) model of Bauwens and Veredas [4] and the long memory stochastic duration (LMSD) model of Deo *et al.* [12]. In this section, we first provide the general link between the counting-process-based assumptions imposed in this paper and the duration-based modeling in the above mentioned papers, followed by sufficient criteria for ACD, SCD and LMSD duration sequences to fit into the present framework of Assumption 2.1.

Let $N(\cdot)$ be a counting process on $\mathbb{R}$ satisfying the conditions of Assumption 2.1 with time-deformation function $f \equiv 0$, (that is, for notational convenience, we will often suppress the distinction between N and $\tilde{N}$ in the remainder of this section) and let moreover $\tau = \{\tau_k\}_{k=-\infty}^{\infty}$ be the sequence of associated durations. Except for the case of a Poisson process, N and τ are not stationary with respect to the same measure. It follows from Iglehart and Whitt [21] that functional limit theorems for counting processes and associated partial sums of durations are essentially equivalent under P. If N is stationary with respect to a measure P, then one can also construct the so-called Palm measure P^0 under which the durations τ are stationary using for example the results in [9, 10, 25]. In duration-based approaches assumptions are stated under P^0 and not under P. In general properties holding under one measure do not translate one-to-one to the other (see the discussions in [14]). We have, however, the following theorem whose proof can be found in Baccelli and Brémaud [2].

Theorem 3.1. *The counting process N is ergodic with respect to P if and only if the durations τ are ergodic with respect to the Palm distribution P^0.*

To validate the ergodicity part of Assumption 2.1 for the ACD, SCD and LMSD models, it suffices according to Theorem 3.1 to establish sufficient conditions ensuring ergodicity of τ under

P^0. We do this in the examples below.

Sufficient conditions on a duration sequence (stationary with finite mean under P^0) for the corresponding counting process $\tilde{N}$ to satisfy the conditions of Assumption 2.1 (i.e., stationarity, ergodicity and simplicity under P), are that the durations are positive with P^0-probability 1, and that the durations are ergodic under P^0. To justify this statement, we note that the positivity, ergodicity, stationarity and finite mean of durations under P^0 implies that $\tilde{N}$, under P, is ergodic, orderly and strictly stationary (by Daley and Vere-Jones [8], Theorem 12.3.II) and the latter two properties imply that $\tilde{N}$ is simple under P (by Daley and Vere-Jones [9], Proposition 3.3.VI).

Since the ACD, SCD and LMSD durations are positive with P^0-probability 1 and have a finite mean under P^0, it will follow that these models have corresponding counting processes $\tilde{N}$, stationary under P, satisfying Assumption 2.1, as long as we can verify that the durations are ergodic under P^0, which we do presently.

Example 3.1. The ACD model proposed in Engle and Russell [16] can be rewritten as

$$\tau_k = \psi_k \varepsilon_k, \qquad \psi_k = g(\varepsilon_{k-1}, \varepsilon_{k-2}, \ldots), \qquad k \in \mathbb{Z}, \tag{3.1}$$

where $\{\varepsilon_k\}_{k=-\infty}^{\infty}$ is a sequence of independent, identically distributed random variables under P^0 satisfying $\varepsilon_0 > 0$ P^0-almost surely, and $E^0[\varepsilon_0] = 1$ with E^0 denoting expectation with respect to P^0, whereas $g > 0$ is a measurable function. It is clear that $\{\tau_k\}_{k=-\infty}^{\infty}$ is then strictly stationary under P^0 provided such a solution to the equations (3.1) exists. From Section 9.5 of Grimmett and Stirzaker [17], it follows that stationary ACD durations are ergodic under P^0 if $E^0[\tau_0] < \infty$. In a simple and more specific form, one can let (see [16])

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

where the parameters involved satisfy $\omega > 0$ and $\alpha, \beta \geq 0$. The P^0-strictly stationary solution is then determined by $\psi_k = \omega \sum_{j=1}^{\infty} \prod_{i=1}^{j-1} (\alpha\varepsilon_{k-i} + \beta)$ and exists if $E^0[\ln(\alpha\varepsilon_0 + \beta)] < 0$, following arguments given in [5]. As in [1] one can derive that now $E^0[\tau_0]$ is finite if $\alpha + \beta < 1$ which, on account of Jensen's inequality, is also sufficient for $E^0[\ln(\alpha\varepsilon_0 + \beta)] < 0$ to hold. Theorem 3.1 consequently implies that the ergodicity part of Assumption 2.1 is satisfied for the ACD model, as long as $\alpha + \beta < 1$.

Example 3.2. In contrast to the ACD model which is observation driven, the SCD model of Bauwens and Veredas [4] is based on a latent variable. It is defined via the equations

$$\tau_k = \psi_k \varepsilon_k, \qquad \ln \psi_k = \omega + \beta \ln \psi_{k-1} + w_k, \qquad k \in \mathbb{Z}, \tag{3.2}$$

where $\omega \in \mathbb{R}$ and $|\beta| < 1$. The sequence $\{\varepsilon_k\}_{k=-\infty}^{\infty}$ consists of independent, identically distributed random variables under P^0 satisfying $\varepsilon_0 > 0$ P^0-almost surely, and $E^0[\varepsilon_0] = 1$. This sequence is P^0-independent of $\{w_k\}_{k=-\infty}^{\infty}$ which is itself independent, identically distributed under P^0. If $E^0[\ln|w_0|] < 0$, then a strictly stationary solution to the ψ-equations exists and is given by the series $\sum_{j=0}^{\infty} \beta^j(\omega + w_{k-j})$. If w_0 possesses P^0-finite moments of all order, then $E[\tau_0] < \infty$ and the same arguments as in Example 3.1 imply here that the ergodicity part of Assumption 2.1 holds as well.

Example 3.3. The LMSD model was introduced as an extension of the SCD model that can capture long memory in the durations. Following [12], the LMSD is given by replacing ψ_k in (3.2) with the linear process specification

$$\tau_k = \psi_k \varepsilon_k, \qquad \ln \psi_k = \sum_{j=0}^{\infty} b_j w_{k-j}, \qquad k \in \mathbb{Z}, \tag{3.3}$$

where $\{w_k\}_{k=-\infty}^{\infty}$ are independent, identically distributed normal random variables with zero mean under P^0, independent of $\{\varepsilon_k\}_{k=-\infty}^{\infty}$ which is assumed to possess all moments and to be positive P^0-almost surely. Long memory in durations is enabled through the coefficients $b_j \sim Cj^{d-1}$, where $C \neq 0$ is a constant and $d \in (0, 1/2)$. It is also possible to nest the short memory case within this framework (see [14] for details). Deo *et al.* [14] have shown in their Lemma 3 that $\{\tau_k\}_{k=-\infty}^{\infty}$ is τ_k-mixing and therefore weak mixing and ergodic under P^0 (see, for example, Choe [7], page 133). It follows from Theorem 3.1 that the associated counting process is ergodic under P. Assumption 2.1 is consequently satisfied for the LMSD model.

4 Main results

4.1 The long-run behavior of the bivariate log-price process

With the assumptions made in Section 2, the long-run behavior of the bivariate process $y = (y_1, y_2)$ can be determined. The following theorem shows that the log-prices are approximately integrated. Even though independence is assumed between the various shock series, the log-price process $y = (y(t)\colon t \geq 0)$ exhibits a nontrivial variance-covariance structure which is determined by a complex interplay of the model parameters.

Theorem 4.1. *If Assumptions 2.1–2.4 are satisfied, then as $n \to \infty$,*

$$\left(\frac{1}{\sqrt{n}} y(nu)\colon u \in [0,1]\right) \xrightarrow{d} B_y = (B_y(u)\colon u \in [0,1]),$$

where $\xrightarrow{d}$ signifies convergence in the Skorohod space $\mathcal{D}^2[0,1]$ and B_y is a bivariate Brownian motion with 2×2 covariance matrix $\Sigma = (\Sigma_{i,j}\colon i,j = 1,2)$ given by the entries

$$\Sigma_{1,1} = \lambda_1 \sigma_{1,e}^2 + \theta^2 \lambda_2 \sigma_{e,2}^2, \qquad \Sigma_{2,2} = \frac{\lambda_1 \sigma_{1,e}^2}{\theta^2} + \lambda_2 \sigma_{2,e}^2 \qquad \textit{and} \qquad \Sigma_{1,2} = \frac{\lambda_1 \sigma_{1,e}^2}{\theta} + \theta \lambda_2 \sigma_{2,e}^2 = \Sigma_{2,1},$$

where $\lambda_i = \tilde{\lambda}_i \gamma_i = a.s.\text{-}\lim_t t^{-1} N_i(t)$ is the asymptotic intensity of the counting process N_i.

Hurvich and Wang [19, 20] have in their Theorem 1 computed the long-run variances of $y_1(t)$ and $y_2(t)$ which are given as $\Sigma_{1,1}t$ and $\Sigma_{2,2}t$, respectively. Our theorem yields the variances as well as the covariances in the limiting distribution of $(t^{-1/2}y(t) \colon t \geq 0)$. More importantly, our theorem provides the limiting distribution itself for the (normalized) log-price process y which, in turn, can be used for asymptotic statistical inference.

4.2 The OLS estimator for the cointegrating parameter

In this section, we derive the asymptotic behavior of the ordinary least-squares estimator (OLS) of the cointegrating parameter θ. To do so, we assume that the log-price series are observed at integer multiples of Δt. We will work here, without loss of generality, with $\Delta t = 1$ in order to keep the notation simple. Then (2.1) and (2.2) become

$$y_{1,j} = \sum_{k=1}^{N_1(j)} (e_{1,k} + \eta_{1,k}) + \sum_{k=1}^{N_2(t_{1,N_1(j)})} (\theta e_{2,k} + g_{21}\eta_{2,k}), \tag{4.1}$$

$$y_{2,j} = \sum_{k=1}^{N_2(j)} (e_{2,k} + \eta_{2,k}) + \sum_{k=1}^{N_1(t_{2,N_2(j)})} (\theta^{-1} e_{1,k} + g_{12}\eta_{1,k}). \tag{4.2}$$

Regressing $y_{1,1}, \ldots, y_{1,n}$ on $y_{2,1}, \ldots, y_{2,n}$ without intercept, we obtain the OLS estimator for the cointegration parameter θ as

$$\hat{\theta}_n = \frac{\sum_{j=1}^n y_{2,j} y_{1,j}}{\sum_{j=1}^n y_{2,j}^2}. \tag{4.3}$$

Hurvich and Wang [19, 20] have shown in their Theorem 6 that $\hat{\theta}_n$ is weakly consistent for θ and obtained a lower bound on the rate of convergence in the case of weak fractional, strong fractional and standard cointegration. The exact limit distributions, however, were not given. We fill in this gap next for weak fractional cointegration.

Theorem 4.2. *Under Assumptions 2.1–2.4 with the restrictions that $E^0[\tau_{i,k}^2] < \infty$ and $f_i \equiv 0$ for $i = 1, 2$, and $d_\eta \in (-\frac{1}{2}, 0)$,*

$$n^{-d_\eta}(\hat{\theta}_n - \theta) \xrightarrow{\mathcal{D}} \frac{\sigma \int_0^1 B(u) B_H(u) du}{\Sigma_{2,2}^{1/2} \int_0^1 B^2(u) du} \qquad (n \to \infty),$$

where $\xrightarrow{\mathcal{D}}$ indicates convergence in distribution, $\sigma^2 = \lambda_1^{2H} \sigma_{\eta,1}^2 [1 - \theta g_{12}]^2 + \lambda_2^{2H} \sigma_{\eta,2}^2 [g_{21} - \theta]^2$, $\Sigma_{2,2}$ is defined in Theorem 4.1, $B = (B(u) \colon u \in [0,1])$ is a standard Brownian motion, $B_H = (B_H(u) \colon u \in [0,1])$ is a Type I fractional Brownian motion with Hurst parameter $H = d_\eta + \frac{1}{2}$, and B, B_H are mutually independent.

The result in Theorem 4.2 is similar to that obtained in Proposition 6.5, formula (6.8) of Robinson and Marinucci [30], under their Assumption 6.1, for which a sufficient condition (their formula (6.5)) was verified in Marinucci and Robinson [24] to hold for weak (but not strong) fractional cointegration in the case where the process is linear with respect to *iid* innovations. The restrictions in Theorem 4.2 that $E^0[\tau_{i,k}^2] < \infty$ and $f_i \equiv 0$ for $i = 1, 2$ are needed only for proving Lemma 5.4, which handles certain end effects. We conjecture that if the durations have infinite variance under P^0 (holding the f_i at zero), the limiting distribution of the OLS estimator under weak fractional cointegration would be affected.

The reason we are currently unable to generalize Theorem 4.2 to the strong fractional cointegration case is the same as given by Robinson and Marinucci [30] (page 963) for not covering this case, namely that here the continuous mapping theorem cannot be applied, and that the process with lower memory parameter cannot be approximated by a semimartingale.

We are also unable to provide conditions under which $n(\hat{\theta}_n - \theta)$ converges in distribution in the standard cointegration case. Phillips and Durlauf [27] establish in their Theorem 4.1 that $n(\hat{\theta}_n - \theta)$ converges in distribution to a functional of a bivariate Brownian motion, under assumptions that may not hold for the process $\{(y_{1,j}, y_{2,j})'\}_{j=1}^n$ given in (4.1) and (4.2). Inspection of their proof suggests that in order for this limit result to hold, it is necessary that the bivariate process $\{z_j\}_{j=1}^n$ with $z_j = (y_{1,j} - \theta y_{2,j}\,,\, y_{2,j} - y_{2,j-1})'$ have the property that as $n \to \infty$,

$$\left(\frac{1}{\sqrt{n}} M \sum_{j=1}^{\lfloor nu \rfloor} z_j \colon u \in [0,1] \right) \xrightarrow{d} B_z = (B_z(u) \colon u \in [0,1]), \tag{4.4}$$

where M is a fixed 2×2 matrix and B_z is a bivariate standard Brownian motion. We are unable to provide conditions on the point process and other elements of our model such that (4.4) would hold. In particular, we note that the first component of z_j is

$$\begin{aligned} y_{1,j} - \theta y_{2,j} = & \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k} - \theta \sum_{k=N_2(t_{1,N_1(j)})+1}^{N_2(j)} e_{2,k} \\ & + \xi_{1,N_1(j)} I\{N_1(j) > 0\} - \theta g_{12} \cdot \xi_{1,N_1(t_{2,N_2(j)})} I\{N_1(t_{2,N_2(j)}) > 0\} \\ & - \theta \cdot \xi_{2,N_2(j)} I\{N_2(j) > 0\} + g_{21} \cdot \xi_{2,N_2(t_{1,N_1(j)})} I\{N_1(j) > 0\}. \end{aligned} \tag{4.5}$$

In the special case where $f_i \equiv 0$, for $i = 1, 2$ the $\{\xi_{i,k}\}$ are *iid* with zero mean and finite variance, the $\{\tau_{i,k}\}$ under P^0 are *iid* and restricted to the positive integers, with a distribution in the domain of attraction of an α-stable distribution with $\alpha \in (1, 2)$ and the $\{\tau_{i,k}\}$ are independent of the $\{\xi_{i,k}\}$, it is known that (see, for example, Proposition 3.3 of Hsieh *et al.* [18]) the finite dimensional distributions of $\ell(n)^{-1}n^{-1/\alpha}\sum_{j=1}^{\lfloor nu\rfloor}\xi_{i,N_i(j)}$ converge weakly to those of an α-stable Lévy process as $n\to\infty$ for $u\in[0,1]$ where $\ell(n)$ is slowly varying as $n\to\infty$. Even though the $\{\xi_{i,k}\}$ are *iid*, the fact that there may be long sets of contiguous integers for which $\xi_{i,N_i(j)}$ is constant, due to the heavy tails in the $\{\tau_{i,k}\}$, leads to a non-Gaussian limit for suitably normalized partial sums of the $\{\xi_{i,N_i(j)}\}$. Thus, in view of (4.5) it seems that (4.4) would not hold in this special case. Barring heavy tails in the distributions of the $\{\tau_{i,k}\}$ (as we do in Theorem 4.2) it is still not clear whether and under what conditions (4.4) would hold.

4.3 The OLS Estimator for Univariate Unit-Root Autoregression

Consider a marked point process $N(\cdot)$ on $\mathbb{R}$, with durations $\{\tau_k\}_{k=-\infty}^{\infty}$ and marks (representing weakly-dependent transaction-level returns) $\{u_k\}_{k=-\infty}^{\infty}$, independent of $N(\cdot)$. We assume that the process is simple, stationary and ergodic with finite mean intensity λ under the time-stationary measure P (see Nieuwenhuis [26], page 686) with $E[u_k]=0$ and $\operatorname{Var}u_k<\infty$. We assume in this subsection that the time-deformation function (f) is identically equal to zero. We also assume that as $n\to\infty$,

$$\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{\lfloor nu\rfloor}u_j\colon u\in[0,1]\right)\xrightarrow{d}B_0,\tag{4.6}$$

where $B_0=CW$, C is a positive constant (the long-run variance of $\{u_k\}$) and W is a standard Brownian motion on $[0,1]$. Sufficient conditions for (4.6) are provided in Phillips and Durlauf [27]. It follows from (4.6) and the proof of Theorem 4.1 that under the time-stationary measure P as $n\to\infty$,

$$\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{N(\lfloor nu\rfloor)}u_j\colon u\in[0,1]\right)\xrightarrow{d}B,\tag{4.7}$$

where $B=C\lambda W$ on $[0,1]$ and $\lambda=E[N(1)]=1/E^0[\tau_k]\in(0,\infty)$ is the intensity.

Define a discrete-time log-price process $\{y_j\}_{j=0}^{\infty}$ by

$$y_j=\sum_{k=1}^{N(j)}u_k,\qquad j=0,1,2,\ldots,$$

where $N(j)=N(0,j]$. Thus, $y_0=0$, and

$$y_j=y_{j-1}+U_j,\qquad j=1,2,\ldots,$$

where

$$U_j=\sum_{k=N(j-1)+1}^{N(j)}u_k,\qquad j\in\mathbb{Z}.$$

Note that $\{U_j\}_{j=-\infty}^{\infty}$ is strictly stationary under P. It follows from the independence of $\{u_k\}$ and $N(\cdot)$ together with the assumption $\operatorname{Var} u_k < \infty$ that $E[U_j^2] < \infty$. Now consider the OLS estimator of the $AR(1)$ parameter (without intercept),

$$\hat{a} = \frac{\sum_{j=1}^{n} y_{j-1} y_j}{\sum_{j=1}^{n} y_{j-1}^2} = 1 + \frac{\sum_{j=1}^{n} y_{j-1} U_j}{\sum_{j=1}^{n} y_{j-1}^2} \ .$$

Since

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nu \rfloor} U_j = \frac{1}{\sqrt{n}} \sum_{k=1}^{N(\lfloor nu \rfloor)} u_k = \frac{1}{\sqrt{n}} y_{\lfloor nu \rfloor},$$

it follows from (4.7) that

$$\left(\frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor nu \rfloor} U_j \colon u \in [0,1] \right) \xrightarrow{d} B \tag{4.8}$$

under P, where $B = C\lambda W$ on $[0,1]$. For $k \geq 1$, let $S_k = \sum_{j=1}^{k} U_j$ and $X_n(r) = n^{-1/2} S_{\lfloor nr \rfloor}$ for $r \in [0,1]$. Arguing as in Phillips [28] (page 254), we have

$$\frac{1}{n} \sum_{j=1}^{n} y_{j-1} U_j = \frac{1}{2} \left(X_n^2(1) - \frac{1}{n} \sum_{j=1}^{n} U_j^2 \right).$$

Since P is ergodic, $\{U_j^2\}$ is ergodic under P. It follows from the ergodic theorem that since $E[U_j^2] < \infty$ there exists an $\omega_0^2 > 0$ such that $(1/n) \sum_{j=1}^{n} U_j^2 \to \omega_0^2$ almost surely under P. It then follows from (4.8) that

$$\frac{1}{n} \sum_{j=1}^{n} y_{j-1} U_j \xrightarrow{\mathcal{D}} \frac{1}{2} \left(C^2 \lambda^2 W^2(1) - \omega_0^2 \right)$$

as $n \to \infty$, under P. This, together with (4.8) and the continuous mapping theorem implies that

$$n(\hat{a} - 1) \xrightarrow{\mathcal{D}} \frac{\int_0^1 B dB + \frac{1}{2}(C^2 \lambda^2 - \omega_0^2)}{\int_0^1 B^2(r) dr} \ . \tag{4.9}$$

Except for the constant λ^2, (4.9) is identical to the result (10) of Phillips [28] obtained under strong mixing conditions.

5 Proofs

5.1 Properties of the efficient shocks and the microstructure noise

We provide some limit results for the partial sums of the efficient shocks and the microstructure noise when the upper summation index in the partial sums is nonrandom. To this end, we introduce the relevant notation first. If $(z_k\colon k \geq 1)$ is a sequence of random variables, we denote in the following by $S_z = (S_z(t)\colon t \geq 0)$ its partial sum process, where $S_z(t) = z_1 + \ldots + z_{\lfloor t \rfloor}$.

First, we study the weak convergence of the bivariate partial sum process $S_e = (S_{e,1}, S_{e,2})$. It follows then from Assumptions 2.2 and 2.4 that, for any $T > 0$,

$$\left(\frac{1}{\sqrt{n}} S_e(nu)\colon u \in [0,T]\right) \xrightarrow{d} B_e = (B_e(u)\colon u \in [0,T]), \tag{5.1}$$

where B_e stands for a bivariate Brownian motion with 2×2 covariance matrix $\Sigma_e = \mathrm{diag}(\sigma_{1,e}^2, \sigma_{2,e}^2)$. Observe that the components of B_e are independent, due to the independence of $(e_{1,k}\colon k \geq 1)$ and $(e_{2,k}\colon k \geq 1)$ imposed in Assumption 2.4. The limit result (5.1) holds also under dependence, but Σ_e is in general not diagonal.

Next, we study the weak convergence of the microstructure noise. Let $d_\eta \in (-\frac{1}{2}, 0)$ and set $S_\eta = (S_{\eta,1}, S_{\eta,2})$. In the case of weak fractional cointegration, Theorem 2 in Davydov [11] implies in combination with part (A) of Assumption 2.3 and Assumption 2.4 that

$$\left(\frac{1}{n^{d_\eta + 1/2}} S_\eta(nu)\colon u \in [0,T]\right) \xrightarrow{d} B_\eta = (B_\eta(u)\colon u \in [0,T]), \tag{5.2}$$

where B_η denotes a Type I fractional Brownian motion with Hurst parameter $H = d_\eta + \frac{1}{2} \in (0, \frac{1}{2})$. We have $\Sigma_\eta = \mathrm{diag}(\sigma_{1,\eta}^2, \sigma_{2,\eta}^2)$, in view of Assumption 2.4.

It is important to notice that, again in view of Assumption 2.4, the weak convergence results for the efficient shocks in (5.1) and for the microstructure noise in (5.2) hold also jointly. We formulate this as a proposition.

Proposition 5.1. *If Assumptions 2.1–2.4 are satisfied and $d_\eta \in (-\frac{1}{2}, 0)$, then*

$$\left(\frac{1}{\sqrt{n}} S_e(nu), \frac{1}{n^H} S_\eta(nu)\colon u \in [0,T]\right) \xrightarrow{d} B_{e,\psi} = (B_e, B_\eta) \qquad (n \to \infty), \tag{5.3}$$

where the Gaussian limit process $B_{e,\eta}$ possesses the variance-covariance matrix $\Sigma_{e,\eta} = diag(\Sigma_e, \Sigma_\eta)$, and $H = d_\eta + \frac{1}{2}$.

Another approach to deriving the limit theorems of Section 2 is to start with prescribing the joint convergence in (5.3), potentially allowing for dependence between the various shock sequences. Then, the covariance matrices in the earlier displays (5.1) and (5.2) have to be adjusted accordingly. Though technically feasible, there is no apparent economic explanation for using this more complicated approach and we therefore focus on the set of assumptions given in Section 2.

5.2 Properties of the counting processes

In this subsection we collect several key results concerning the counting processes N_i. By assumption 2.1, we have that $N_i(t) = \tilde{N}_i(t + f_i(t))$, where $\tilde{N}_i$ is stationary and ergodic. Therefore (see displays (12.2.3) and (12.2.4) of Daley and Vere-Jones [10]), $\tilde{N}_i$ satisfies the renewal theorem

$$\frac{\tilde{N}_i(t)}{t} \to \tilde{\lambda}_i \qquad \text{a.s.} \qquad (t \to \infty), \tag{5.4}$$

The latter limit result translates as follows to the time-deformed counting processes N_i used to build the bivariate transaction-level returns $(y(t)\colon t \geq 0)$.

Lemma 5.1. *If $N_i(\cdot)$, $i = 1, 2$, are the counting processes of Assumption 2.1, then*

$$\frac{N_i(t)}{t} \to \lambda_i = \tilde{\lambda}_i \gamma_i \qquad P\text{-a.s.} \qquad (t \to \infty) \tag{5.5}$$

Proof. Since $N_i(t) = \tilde{N}_i(t + f_i(t))$, we obtain

$$\frac{N_i(t)}{t} = \frac{\tilde{N}_i(t + f_i(t))}{t} = \frac{\tilde{N}_i(t + f_i(t))}{t + f_i(t)} \frac{t + f_i(t)}{t}.$$

Since $t + f_i(t) \to \infty$ with probability one, the renewal theorem (5.4) implies that the first term on the right-hand side converges with probability one to $\tilde{\lambda}_i$. The second term satisfies $t^{-1}[t + f_i(t)] \to \gamma_i$ with probability one on account of Assumption 2.1. This is the assertion. □

The following proposition contains an auxiliary result needed to exchange asymptotically the counting process value at a large enough time t, $N_i(t)$, with the corresponding average rate $\lambda_i t$. We formulate it in terms of an arbitrary stochastic process $(Z(t)\colon t \geq 0)$.

Proposition 5.2. *Let $(Z(t)\colon t \geq 0)$ be a stochastic process satisfying $Z(s) \leq Z(t)$ if $s \leq t$ and the renewal theorem $t^{-1}Z(t) \to \lambda \geq 0$ with probability one as $t \to \infty$. Then,*

$$\sup_{0 \leq u \leq 1} \left| \frac{Z(nu)}{n} - \lambda u \right| \to 0 \qquad P\text{-a.s.} \qquad (n \to \infty).$$

Proof. Let $\epsilon > 0$. Choose $\delta > 0$ such that $\lambda\delta \leq \epsilon$. Now the monotonicity of the stochastic process $(Z(t)\colon t \geq 0)$ implies that

$$\sup_{0 \leq u \leq \delta} \left| \frac{Z(nu)}{n} - \lambda u \right| \leq \frac{Z(n\delta)}{n} + \lambda\delta$$

and therefore, taking the limit superior on both side of this equation and applying the renewal theorem for $(Z(t)\colon t \geq 0)$, we get

$$\limsup_{n \to \infty} \sup_{0 \leq u \leq \delta} \left| \frac{Z(nu)}{n} - \lambda u \right| \leq 2\lambda\delta \leq 2\epsilon \qquad \text{a.s.}$$

In the preceding, we can choose $\delta < 1$ without loss of generality. In the following, we still need to consider the limit superior of the supremum taken over $u \in [\delta, 1]$ to complete the proof. To this end, note that there is a random variable t_0 such that $-\epsilon t \le Z(t) - \lambda t \le \epsilon t$ if $t \ge t_0$. This implies for $n \ge n_0 = \delta^{-1} t_0$ and $\delta \le u \le 1$ that $-\epsilon n u \le Z(nu) - \lambda n u \le \epsilon n u$ and thus

$$-\epsilon n \le Z(nu) - \lambda n u \le \epsilon n, \qquad u \in [\delta, 1].$$

Since the lower and upper bounds are independent of u, the latter inequalities imply that

$$\limsup_{n\to\infty} \sup_{\delta \le u \le 1} \left| \frac{Z(nu)}{n} - \lambda u \right| \le \epsilon \qquad \text{a.s.},$$

which also completes the proof. $\square$

Next, we will show that the processes N_i satisfy the assumptions of Proposition 5.2.

Lemma 5.2. *If $N_i(\cdot)$ are the counting processes of Assumption 2.1, then it holds with probability one that*

$$\sup_{0 \le u \le 1} \left| \frac{N_i(nu)}{n} - \lambda_i u \right| \to 0 \qquad (n \to \infty).$$

Proof. Since $\tilde{N}_i$ is a counting process and $s + f_i(s) \le t + f_i(t)$ as long as $s \le t$ by Assumption 2.1, we have with probability one that

$$N_i(s) = \tilde{N}_i(s + f_i(s)) \le \tilde{N}_i(t + f_i(t)) = N_i(t), \qquad s \le t.$$

We have shown in Lemma 5.1 that N_i satisfies a renewal theorem, so that the assertion of this lemma follows from Proposition 5.2. $\square$

The definition of $y_1(t)$ and $y_2(t)$ also contains the quantities $N_2(t_{1,N_1(t)})$ and $N_1(t_{2,N_2(t)})$ whose limiting behavior needs to be determined for the asymptotics in Theorems 4.1 and 4.2. Recall that $t_{i,k}$ denotes the clock-time of the kth event of N_i.

Lemma 5.3. *If $N_i(\cdot)$ are the counting processes of Assumption 2.1, then it holds with probability one that*

$$\sup_{0 \le u \le 1} \left| \frac{N_2(t_{1,N_1(nu)})}{n} - \lambda_2 u \right| = o(1) \qquad (n \to \infty).$$

This statement remains true if the roles of the indices 1 and 2 are interchanged.

Proof. Since N_1 and N_2 are nondecreasing in clock-time t and $t_{1,k}$ is nondecreasing in tick-time k, we get that $N_2(t_{1,N_1(s)}) \le N_2(t_{1,N_1(t)})$ whenever $s \le t$. Note that clearly $t_{1,k} \to \infty$ with probability one, as $k \to \infty$. Observe next that

$$1 \le \frac{N_1(t_{1,k})}{k} \le 1 + \frac{1}{k}\left[\tilde{N}_1(t_{1,k} + f_1(t_{1,k})) - \tilde{N}_1(t_{1,k}^- + f_1(t_{1,k}^-))\right] = 1 + o(1) \qquad \text{a.s.}$$

as $k \to \infty$, using the definition of N_1 and the boundedness requirement on the jumps of f_1 imposed in Assumption 2.1. Hence, by Lemma 5.2,

$$\frac{t_{1,k}}{k} \sim \frac{t_{1,k}}{N_1(t_{1,k})} \to \frac{1}{\lambda_1} \qquad \text{a.s.} \qquad (k \to \infty),$$

where $a_k \sim b_k$ indicates that the ratio of a_k and b_k tends to one as $k \to \infty$. Now we arrive at the asymptotic relation

$$\frac{N_2(t_{1,N_1(t)})}{t} = \frac{N_2(t_{1,N_1(t)})}{t_{1,N_1(t)}} \frac{t_{1,N_1(t)}}{N_1(t)} \frac{N_1(t)}{t} \to \lambda_2 \frac{1}{\lambda_1} \lambda_1 = \lambda_2 \qquad \text{a.s.}$$

The assertion of the lemma follows therefore again from Proposition 5.2. □

5.3 Proofs of Theorems

First, we separate the efficient shocks components from the microstructure noise and thus decompose $y = r + s$, where the processes $s = (s(t) \colon t \geq 0)$ and $r = (r(t) \colon t \geq 0)$ are, for $t \geq 0$, given by $s(t) = (s_1(t), s_2(t))$ and $r(t) = (r_1(t), r_2(t))$ with

$$s_1(t) = \sum_{k=1}^{N_1(t)} e_{1,k} + \theta \sum_{k=1}^{N_2(t_{1,N_1(t)})} e_{2,k}, \qquad s_2(t) = \sum_{k=1}^{N_2(t)} e_{2,k} + \theta^{-1} \sum_{k=1}^{N_1(t_{2,N_2(t)})} e_{1,k},$$

$$r_1(t) = \sum_{k=1}^{N_1(t)} \eta_{1,k} + g_{21} \sum_{k=1}^{N_2(t_{1,N_1(t)})} \eta_{2,k}, \qquad r_2(t) = \sum_{k=1}^{N_2(t)} \eta_{2,k} + g_{12} \sum_{k=1}^{N_1(t_{2,N_2(t)})} \eta_{1,k}.$$

To apply the weak convergence results collected in Proposition 5.1, it is necessary to replace the random limits in the partial sums that define $s_i(t)$ and $r_i(t)$ with deterministic ones. These key steps have been established in the previous subsection.

Proof of Theorem 4.1. The assertion that $\lambda_i = \tilde{\lambda}_i \gamma_i = \text{a.s.-}\lim_t N_i(t)$ is proved in Lemma 5.1. In view of (5.1) we show first that the microstructure noise components are less persistent than the efficient shocks components for all $d_\eta \in [-1, 0)$, that is,

$$\frac{1}{\sqrt{n}} \sup_{u \in [0,1]} |r(nu)| = o_P(1) \qquad (n \to \infty), \tag{5.6}$$

where $|\cdot|$ denotes Euclidean norm. The asymptotic behavior of y is therefore completely determined by s. To prove (5.6), we start with letting $d_\eta \in (-\frac{1}{2}, 0)$. Utilizing the weak convergence result (5.2) and the Skorohod-Dudley-Wichura representation theorem (see Shorack and Wellner [31], page 47), we obtain that, for every $n \geq 1$ and $T > 0$, there exists a process $B_\eta^{(n)} \in \mathcal{C}^2[0, T]$, the set of stochastic processes in $\mathbb{R}^2$ with a.s. continuous sample paths, such that $B_\eta^{(n)} \stackrel{\mathcal{D}}{=} B_\eta$ and

$$\sup_{u \in [0,T]} \left| \frac{1}{n^H} S_\eta(nu) - B_\eta^{(n)}(u) \right| = o(1) \qquad \text{a.s.} \qquad (n \to \infty), \tag{5.7}$$

where $H = d_\eta + \frac{1}{2}$. We focus on $r_1(t)$ in the following. The same arguments provide a similar result also for $r_2(t)$. Note that $r_1(t) = S_{\eta,1}(N_1(t)) + g_{21} S_{\eta,2}(N_2(t_{1,N_1(t)}))$ and that consequently

$$
\begin{aligned}
&\sup_{u\in[0,1]} \left| \frac{1}{n^H} r_1(nu) - \Big[B^{(n)}_{\eta,1}(\lambda_1 u) + g_{21} B^{(n)}_{\eta,2}(\lambda_2 u) \Big] \right| \qquad (5.8) \\
&\quad \le \sup_{u\in[0,1]} \left| \frac{1}{n^H} r_1(nu) - \Big[B^{(n)}_{\eta,1}\Big(\frac{N_1(nu)}{n}\Big) + g_{21} B^{(n)}_{\eta,2}\Big(\frac{N_2(t_{1,N_1(nu)})}{n}\Big) \Big] \right| \\
&\qquad + \sup_{u\in[0,1]} \left| \Big[B^{(n)}_{\eta,1}\Big(\frac{N_1(nu)}{n}\Big) + g_{21} B^{(n)}_{\eta,2}\Big(\frac{N_2(t_{1,N_1(nu)})}{n}\Big) \Big] - \Big[B^{(n)}_{\eta,1}(\lambda_1 u) + g_{21} B^{(n)}_{\eta,2}(\lambda_2 u) \Big] \right| \\
&\quad = o(1) \qquad \text{a.s.}
\end{aligned}
$$

as $n \to \infty$, using display (5.7) for the first term and the continuity of the fractional Brownian motion sample paths in combination with Lemmas 5.2 and 5.3 for the second. Since $H < \frac{1}{2}$, (5.6) is established for the weak fractional cointegration case. In the case of strong fractional cointegration (standard cointegration), we have that $r_1(t) = \varphi_{1,N_1(t)} + g_{21}\varphi_{2,N_2(t_{1,N_1(t)})}$ and $r_2(t) = \varphi_{2,N_2(t)} + g_{12}\varphi_{1,N_1(t_{2,N_2(t)})}$ ($r_1(t) = \xi_{1,N_1(t)} + g_{21}\xi_{2,N_2(t_{1,N_1(t)})}$ and $r_2(t) = \xi_{2,N_2(t)} + g_{12}\xi_{1,N_1(t_{2,N_2(t)})}$) are each the sum of two random variables and thus (5.6) continues to hold.

To complete the proof of Theorem 4.1, we now derive the limit for $s(t)$ and do so only for its first component $s_1(t) = S_{e,1}(N_1(t)) + \theta S_{e,2}(N_2(t_{1,N_1(t)}))$. The Skorohod-Dudley-Wichura representation theorem and (5.1) yield that, for every $n \ge 1$ and $T > 0$, there is a process $B^{(n)}_e \in \mathcal{C}^2[0,T]$ such that $B^{(n)}_e \stackrel{\mathcal{D}}{=} B_e$ and

$$
\sup_{u\in[0,T]} \left| \frac{1}{\sqrt{n}} S_e(nu) - B^{(n)}_e(u) \right| = o(1) \qquad \text{a.s.} \qquad (n\to\infty). \tag{5.9}
$$

Therefore,

$$
\begin{aligned}
&\sup_{u\in[0,1]} \left| \frac{1}{\sqrt{n}} s_1(nu) - \Big[B^{(n)}_{e,1}(\lambda_1 u) + \theta B^{(n)}_{e,2}(\lambda_2 u) \Big] \right| \qquad (5.10) \\
&\quad \le \sup_{u\in[0,1]} \left| \frac{1}{\sqrt{n}} s_1(nu) - \Big[B^{(n)}_{e,1}\Big(\frac{N_1(nu)}{n}\Big) + \theta B^{(n)}_{e,2}\Big(\frac{N_2(t_{1,N_1(nu)})}{n}\Big) \Big] \right| \\
&\qquad + \sup_{u\in[0,1]} \left| \Big[B^{(n)}_{e,1}\Big(\frac{N_1(nu)}{n}\Big) + \theta B^{(n)}_{e,2}\Big(\frac{N_2(t_{1,N_1(nu)})}{n}\Big) \Big] - \Big[B^{(n)}_{e,1}(\lambda_1 u) + \theta B^{(n)}_{e,2}(\lambda_2 u) \Big] \right| \\
&\quad = o(1) \qquad \text{a.s.}
\end{aligned}
$$

as $n \to \infty$, applying the same arguments as in (5.8) replacing display (5.7) with (5.9) and fractional with standard Brownian motion. Similarly, one obtains that

$$
\sup_{u\in[0,1]} \left| \frac{1}{\sqrt{n}} s_2(nu) - \Big[B^{(n)}_{e,2}(\lambda_2 u) + \theta^{-1} B^{(n)}_{e,1}(\lambda_1 u) \Big] \right| = o(1) \qquad \text{a.s.} \qquad (n\to\infty). \tag{5.11}
$$

The distributions of $B_{y,1}(u) = B_{e,1}^{(n)}(\lambda_1 u) + \theta B_{e,2}^{(n)}(\lambda_2 u)$ and $B_{y,2}(u) = B_{e,2}^{(n)}(\lambda_2 u) + \theta^{-1} B_{e,1}^{(n)}(\lambda_1 u)$ do not depend on n and direct computations show that $B_y = (B_{y,1}, B_{y,2})$ is a bivariate Brownian motion with covariance matrix Σ as given in Theorem 4.1. This completes the proof. $\square$

To give the proof of Theorem 4.2, we need one further lemma which deals with the end effects.

Lemma 5.4. *If the assumptions of Theorem 4.2 are satisfied, then*

$$E_1(n) := \frac{1}{n^{3/2}} \sum_{j=1}^{n} y_{2,j} \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k} = \mathcal{O}_P(1) \qquad (n \to \infty),$$

$$E_2(n) := \frac{1}{n^{3/2}} \sum_{j=1}^{n} y_{2,j} \sum_{k=N_2(t_{1,N_1(j)})+1}^{N_2(j)} e_{2,k} = \mathcal{O}_P(1) \qquad (n \to \infty).$$

Proof: We only provide the proof of the first statement, the second follows analogously. First, we estimate

$$|E_1(n)| \le \frac{1}{\sqrt{n}} \max_{1 \le j \le n} |y_{2,j}| \cdot \frac{1}{n} \sum_{j=1}^{n} \left| \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k} \right| = F_1(n) \cdot F_2(n),$$

where the first term $F_1(n)$ converges in distribution to $\Sigma_{2,2} \sup_t |B_{y,2}(t)|$ by Theorem 4.1 and the continuous mapping theorem. To complete the proof, it remains to be shown that $F_2(n) = \mathcal{O}_P(1)$. If we define

$$x_j = \sum_{k=N_1(t_{2,N_2(j)})+1}^{N_1(j)} e_{1,k},$$

then $F_2(n) = \frac{1}{n} \sum_{j=1}^{n} |x_j|$. Note that $E[|x_j|] < CE[N_1(j) - N_1(t_{2,N_2(j)})]$ with some $C > 0$. So we consider $E[N_1(j) - N_1(t_{2,N_2(j)})]$, which is the expected number of transactions of Asset 1 after the most recent transaction of Asset 2 up to time j. Let

$$B_{2,j} = \inf\{s > 0 : N_2(j) - N_2(j - s) > 0\}$$

be the backward recurrence time for Asset 2 at time j. Clearly, $B_{2,j} = j - t_{2,N_2(j)}$. Exploiting the P-stationarity of N_2 and display (3.1.7) of Daley and Vere-Jones [9] we obtain

$$E[N_1(j) - N_1(t_{2,N_2(j)})] = E[-N(-B_{2,j})] = E[-N_1(-B_{2,0})]. \tag{5.12}$$

In the right-hand equality, we have used that, since N_2 is a P-stationary point process, $B_{2,j}$ has the same distribution as $B_{2,0}$ which does not depend on j (see pages 58–59 of [9] for a detailed discussion). By Example 3.4.1 of [9], $B_{2,0} \stackrel{\mathcal{D}}{=} T_{2,0}$, where $T_{2,0}$ denotes the forward recurrence time for Asset 2 at time $j = 0$. By equation (3.4.17) of [9], $E[T_{2,0}] = \frac{1}{2}\lambda_2 E^0[\tau_{2,0}^2]$, which is finite by assumption on $\{\tau_{2,k}\}_{k=-\infty}^{\infty}$. Going back to (5.12) we find that $E[N_1(j) - N_1(t_{2,N_2(j)})] < \infty$ and consequently $E[|x_j|] < \infty$. Since $\{x_j\}_{j=-\infty}^{\infty}$ is a strictly stationary process under P, an application of the ergodic theorem given on page 209 of Dudley [15] implies that $F_2(n)$ converges with P-probability one to a random variable with finite mean. Hence, $F_2(n) = \mathcal{O}_P(1)$. The proof is complete. $\square$

Proof of Theorem 4.2. Observe that

$$\hat{\theta}_n - \theta = \frac{\sum_{j=1}^n y_{2,j}(y_{1,j} - \theta y_{2,j})}{\sum_{j=1}^n y_{2,j}^2}.$$

First, we consider the denominator part of $\hat{\theta}_n - \theta$. It follows from Theorem 4.1 that the partial sum-type process $\frac{1}{\sqrt{n}} y_2(n\,\cdot)$ converges weakly in $\mathcal{D}[0,1]$ to $\Sigma_{2,2}^{1/2} B$, where $B = (B(u)\colon u \in [0,1])$ is a standard Brownian motion. Thus, the continuous mapping theorem and standard arguments imply that

$$\frac{1}{n^2} \sum_{j=1}^n y_{2,j}^2 = \frac{1}{n} \sum_{j=1}^n \left(\frac{1}{\sqrt{n}} y_{2,j} \right)^2 \xrightarrow{\mathcal{D}} \Sigma_{2,2} \int_0^1 B^2(u) du \qquad (n \to \infty). \tag{5.13}$$

As for the numerator part, we obtain from the definition of $y_{1,j}$ and $y_{2,j}$ in (4.1) and (4.2) that

$$\begin{aligned} A_n &:= \sum_{j=1}^n y_{2,j}(y_{1,j} - \theta y_{2,j}) \\ &= \sum_{j=1}^n y_{2,j}(r_{1,j} - \theta r_{2,j}) + \sum_{j=1}^n y_{2,j}(s_{1,j} - \theta s_{2,j}) \\ &= \sum_{j=1}^n y_{2,j}(r_{1,j} - \theta r_{2,j}) + n^{3/2}(E_1(n) - \theta E_2(n)). \end{aligned}$$

Since $2 + d_\eta > 3/2$, Lemma 5.4 first implies that $n^{3/2-2-d_\eta}(E_1(n) - \theta E_2(n)) = o_P(1)$. As indicated above, Theorem 4.1 implies that $\frac{1}{\sqrt{n}} y_2(n\,\cdot)$ converges weakly in $\mathcal{D}[0,1]$ to $\Sigma_{2,2}^{1/2} B$. Relation (5.8) and the corresponding statement for r_2 yield moreover that the microstructure process $\frac{1}{n^H}[r_1(n\,\cdot) - \theta r_2(n\,\cdot)]$ converges weakly in $\mathcal{D}[0,1]$ to σB_H, where the variance parameter σ^2 is defined in Theorem 4.2 and $B_H = (B_H(u)\colon u \in [0,1])$ denotes a fractional Brownian motion with Hurst parameter $H = d_\eta + \frac{1}{2}$. An application of Theorem 2.2 in Kurtz and Protter [22] leads now to

$$\begin{aligned} \frac{1}{n^{2+d_\eta}} A_n &= \frac{1}{n} \sum_{j=1}^n \left(\frac{1}{\sqrt{n}} y_{2,j} \right) \left(\frac{1}{n^H} [r_{1,j} - \theta r_{2,j}] \right) + o_P(1) \\ &\xrightarrow{\mathcal{D}} \Sigma_{2,2}^{1/2} \sigma \int_0^1 B(u) B_H(u) du, \end{aligned} \tag{5.14}$$

as $n \to \infty$. By Assumption 2.4 and (5.6), $y_{2,j}$ and $r_{1,j} - \theta r_{2,j}$ are moreover asymptotically independent, thereby rendering B and B_H independent. In view of (5.13) and (5.14), the proof of Theorem 4.2 is complete. □

References

[1] Aue, A., Berkes, I. and L. Horváth (2006). Strong approximation for the sums of squares of augmented GARCH sequences. *Bernoulli* **12**, 583–608.

[2] Baccelli, F. and P. Brémaud (1987). *Palm Probabilities and Stationary Queues. Lecture Notes in Statistics* **41**. Springer-Verlag, Berlin.

[3] Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and N. Shephard (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica* **76**, 1481–1536.

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

[5] Bougerol, P. and N. Picard (1992). Stationarity of GARCH processes and of some nonnegative time series. *Journal of Econometrics* **52**, 115–127.

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

[7] Choe, G.H. (2005). *Computational Ergodic Theory.* Springer-Verlag, New York.

[8] Daley, D.J. and D. Vere-Jones (1988). *An Introduction to the Theory of Point Processes (1st ed.).* Springer-Verlag, New York.

[9] Daley, D.J. and D. Vere-Jones (2003). *An Introduction to the Theory of Point Processes, Vol. I (2nd ed.).* Springer-Verlag, New York.

[10] Daley, D.J. and D. Vere-Jones (2008). *An Introduction to the Theory of Point Processes, Vol. II (2nd ed.).* Springer-Verlag, New York.

[11] Davydov, Y.A. (1970). The invariance principle for stationary processes. *Theory of Probability and Its Applications* **15**, 487–498.

[12] Deo, R., Hsieh, M. and C. Hurvich (2007). Long memory in intertrade durations, counts and realized volatility of NYSE stocks. Working Paper, Stern School of Business, New York University.

[13] Deo, R., Hurvich, C. and Y. Lu (2006). Forecasting realized volatility using a long memory stochastic volatility model: estimation, prediction and seasonal adjustment. *Journal of Econometrics* **131**, 29–58.

[14] Deo, R., Hurvich, C., Soulier, P. and Y. Wang (2009). Conditions for the propagation of memory parameter from durations to counts and realized volatility. *Econometric Theory*, to appear.

[15] Dudley, R.M. (1989). *Real Analysis and Probability.* Wadsworth & Brooks/Cole, Pacific Grove, CA.

[16] Engle, R. and J. Russell (1998). Autoregressive conditional durations: a new model for irregularly spaced transaction data. *Econometrica* **66**, 127–1162.

[17] Grimmett, G. and D. Stirzaker (2001). *Probability and Random Processes (3rd ed.).* Oxford University Press, Oxford.

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

[19] Hurvich, C.M. and Y. Wang (2009). A pure-jump transaction-level price model yielding cointegration. *Journal of Business and Economic Statistics*, to appear.

[20] Hurvich, C.M. and Y. Wang (2009). A pure-jump transaction-level price model yielding cointegration, leverage, and nonsynchronous trading effects. Working Paper, Stern School of Business, New York University.

[21] Iglehart, DL. and W. Whitt (1971). The equivalence of functional central limit theory for counting processes and associated partial sums. *The Annals of Mathematical Statistics* **42**, 1372–1378.

[22] Kurtz, T.G. and P. Protter (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *The Annals of Probability* **19**, 1035–1070.

[23] Lancaster, T. (1979). Econometric methods for the duration of unemployment. *Econometrica* **47**, 939–956.

[24] Marinucci, D. and P.M. Robinson (2000). Weak convergence of multivariate fractional processes. *Stochastic Processes and their Applications* **86**, 103–120.

[25] Nieuwenhuis, G. (1989). Equivalence of functional limit theorems for stationary point processes and their palm distributions. *Probability Theory and Related Fields* **81**, 593–608.

[26] Nieuwenhuis, G. (1998). Ergodicity conditions and Cesàro limit results for marked point processes. *Stochastic Models* **14**, 681–714.

[27] Phillips, P.C.B. and S. Durlauf (1986). Multiple time series regression with integrated processes. *The Review of Economic Studies* **53**, 473–495.

[28] Phillips, P.C.B. (1988). Weak convergence to the matrix stochastic integral $\int_0^1 B\,dB'$. *Journal of Multivariate Analysis* **24**, 252–264.

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

[30] Robinson, P.M. and D. Marinucci (2001). Narrowband analysis of nonstationary processes. *The Annals of Statistics* **29**, 947–986.

[31] Shorack, G.R. and J.A. Wellner (1986). *Empirical processes with applications to statistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics.* Wiley, New York.