A Pure-Jump Transaction-Level Price Model Yielding Cointegration

Clifford M. HURVICH and Yi WANG
Stern School of Business, New York University, New York, NY 10012
(*churvich@stern.nyu.edu*; *ywang2@stern.nyu.edu*)

We propose a new transaction-level bivariate log-price model that yields fractional or standard cointegration. The model provides a link between market microstructure and lower-frequency observations. The two ingredients of our model are a long-memory stochastic duration process for the waiting times, $\{\tau_k\}$, between trades and a pair of stationary noise processes, ($\{e_k\}$ and $\{\eta_k\}$), which determine the jump sizes in the pure-jump log-price process. Our model includes feedback between the disturbances of the two log-price series at the transaction level, which induces standard or fractional cointegration for any fixed sampling interval Δt. We prove that the cointegrating parameter can be consistently estimated by the ordinary least squares estimator, and we obtain a lower bound on the rate of convergence. We propose transaction-level method-of-moments estimators of the other parameters in our model and discuss the consistency of these estimators.

KEY WORDS: Information share; Long-memory stochastic duration; Tick time.

1. INTRODUCTION

We propose a transaction-level, pure-jump model for a bivariate price series in which the intertrade durations are stochastic and enter into the model in a fully endogenous way. The model is flexible and able to capture a variety of stylized facts, including standard or fractional cointegration, persistence in durations, volatility clustering, leverage (i.e., a negative correlation between current returns and future volatility), and nonsynchronous trading effects. In our model, all of these features observed at equally spaced time intervals are derived from transaction-level properties. Thus the model provides a link between market microstructure and lower-frequency observations. This article focuses on the cointegration aspects of the model, presenting theoretical, simulation, and empirical analyses.

Cointegration is a well-known phenomenon that has received considerable attention in economics and econometrics. Under both standard and fractional cointegration, there is a contemporaneous linear combination of two or more time series that is less persistent than the individual series. Under standard cointegration, the memory parameter is reduced from 1 to 0, while under fractional cointegration, the level of reduction need not be an integer. Indeed, in the seminal article of Engle and Granger (1987), both standard and fractional cointegration were allowed for, although the literature has since developed separately for the two cases. Important contributions to the representation, estimation, and testing of standard cointegration models include those of Stock and Watson (1988), Johansen (1988, 1991), and Phillips (1991a). Literature addressing the corresponding problems in fractional cointegration includes works by Dueker and Startz (1998), Marinucci and Robinson (2001), Robinson and Marinucci (2001), Robinson and Yajima (2002), Robinson and Hualde (2003), Velasco (2003), Velasco and Marmol (2004), and Chen and Hurvich (2003a, 2003b, 2006).

A limitation of most existing models for cointegration is that they are based on a particular fixed sampling interval, Δt (e.g., 1 day, 1 month) and thus do not reflect the dynamics at all levels of aggregation. Indeed, Engle and Granger (1987) assumed a fixed sampling interval. It is also possible to build models for cointegration using diffusion-type continuous-time models, such as ordinary or fractional Brownian motion (see Phillips 1991b; Comte and Renault 1996, 1998; Comte 1998), but such models would fail to capture the pure-jump nature of observed asset-price processes.

In this article we propose a pure-jump model for a bivariate log-price series such that any discretization of the process to an equally spaced sampling grid with sampling interval Δt produces fractional or standard cointegration; that is, there exists a contemporaneous linear combination of the two log-price series that has a smaller memory parameter than the two individual series. A key ingredient in our model is a microstructure noise contribution, $\{\eta_k\}$, to the log prices. In the *weak fractional cointegration* case, this noise series is assumed to have memory parameter $d_\eta \in (-\frac{1}{2}, 0)$ in the *strong fractional cointegration* case $d_\eta \in (-1, -\frac{1}{2})$, while in the *standard cointegration* case, $d_\eta = -1$. In all three cases, the reduction of the memory parameter is $-d_\eta$. Due to the presence of the microstructure noise term, the discretized log-price series are not martingales, and the corresponding return series are not linear in an iid sequence, a martingale-difference sequence, or a strong-mixing sequence. This is in sharp contrast to existing discrete-time models for cointegration, most of which assume at least that the series has a linear representation with respect to a strong-mixing sequence.

The discretely sampled returns (i.e., the increments in the log-price series) in our model are not martingale differences, because of the microstructure noise term. Instead, for small values of Δt they may exhibit noticeable autocorrelations, as also seen in actual returns over short time intervals. Nevertheless, the returns behave asymptotically like Martingale differences as the sampling interval Δt is increased, in the sense that the lag-k autocorrelation tends to 0 as Δt tends to ∞ for any fixed k. Again, this is consistent with the near-uncorrelatedness observed in actual returns measured over long time intervals.

Journal of Business & Economic Statistics
October 2010, Vol. 28, No. 4
DOI: 10.1198/jbes.2009.07116

The memory parameter of the log prices in our model is 1, in the sense that the variance of the log price increases linearly in t asymptotically as $t \to \infty$. In contrast, the memory parameter of the appropriate contemporaneous linear combination of the two log-price series is reduced to $(1 + d_\eta) < 1$, thereby establishing the existence of cointegration in our model.

To derive the results described herein, we make use of the general theory of point processes, and also rely heavily on the theory developed by Deo et al. (2009) for the counting process $N(t)$ induced by a long-memory duration process. In Section 2 we present our pure-jump model for the bivariate log-price series. Because the two series need not have all of their transactions at the same time points (due to nonsynchronous trading), it is not possible to induce cointegration in the traditional way, that is, by directly imposing in clock time (calendar time) an additive common component for the two series with a memory parameter equal to 1. Instead, the common component is induced indirectly, and incompletely, by means of a feedback mechanism in transaction time between current log-price disturbances of one asset and previous log-price disturbances of the other asset. This feedback mechanism also induces certain end-effect terms, which we explicitly display and handle in our theoretical derivations using the theory of point processes.

The article is organized as follows. In Section 2 we provide economic justification for the model, as well as a transaction-level definition of the information share of a market. We also present some preliminary data analysis results that affirm the potential usefulness of certain flexibilities in the model. In Section 3 we define conditions on the microstructure noise process for both fractional and standard cointegration. These conditions are satisfied by various standard time series models. In Section 4 we present the properties of the log-price series implied by our model. In particular, we show that the log price behaves asymptotically like a martingale as t is increased, and that the discretely sampled returns behave asymptotically like Martingale differences as the sampling interval Δt is increased. In Section 5 we establish that our model has cointegration, by showing that the cointegrating error has memory parameter $(1 + d_\eta)$. We present separate theorems for the weak and strong fractional cointegration and standard cointegration cases. In Section 6 we show that the ordinary least squares (OLS) estimator of the cointegrating parameter θ is consistent, and obtain a lower bound on its rate of convergence. In Section 7 we propose an alternative cointegrating parameter estimator based on the tick-level price series. In Section 8, we propose a method-of-moments estimator for the tick-level model parameters (except the cointegrating parameter θ). The method is based on the observed tick-level returns. In Section 9 we propose a specification test for the transaction-level price model, and in Section 10 we present simulation results on the OLS estimator of θ, the tick-level cointegrating parameter estimator $\tilde{\theta}$, the method-of-moments estimator, and the proposed specification test. In Section 11 we present a data analysis of buy and sell prices of a single stock (Tiffany; TIF), providing evidence of strong fractional cointegration. The cointegrating parameter is estimated by both OLS regression and the alternative tick-level method proposed in Section 7. The proposed specification test is implemented on the data. Interesting results are observed that are consistent with the existing literature about price discovery process in a market-dealer market. We then consider the information content of buy trades versus sell trades in different market environments. In Section 12 we provide some concluding remarks and discuss possible further generalizations of our model and related future work. We provide proofs in the Appendix.

2. A PURE–JUMP MODEL FOR LOG PRICES

Before describing our model, we provide some background on transaction-level modeling. Currently, a wealth of transaction-level price data is available, and for such data, the (observed) price remains constant between transactions. If there is a diffusion component underlying the price, it is not directly observable. Thus pure-jump models for prices provide a potentially appealing alternative to diffusion-type models. The compound Poisson process proposed by Press (1967) is a pure-jump model for the logarithmic price series under which innovations to the log price are iid, and these innovations are introduced at random time points, determined by a Poisson process. The model was generalized by Oomen (2006), who introduced an additional innovation term to capture market microstructure.

An informative and directly observable quantity in transaction-level data is the duration $\{\tau_k\}$ between transactions. In a seminal article focusing on durations and, to some extent, on the induced price process, Engle and Russell (1998) documented a key empirical fact that durations are strongly autocorrelated, quite unlike the iid exponential duration process implied by a Poisson transaction process, and they proposed the autoregressive conditional duration (ACD) model, which is closely related to the generalized autoregressive conditional heteroscedasticity model of Bollerslev (1986). Deo, Hsieh, and Hurvich (2006) presented empirical evidence that durations, as well as transaction counts, squared returns, and realized volatility, have long memory, and introduced the long-memory stochastic duration (LMSD) model, which is closely related to the long-memory stochastic volatility model of Breidt, Crato, and de Lima (1998) and Harvey (1998). The LMSD model is $\tau_k = e^{h_k}\epsilon_k$, where $\{h_k\}$ is a Gaussian long-memory series with memory parameter $d_\tau \in (0, \frac{1}{2})$, the $\{\epsilon_k\}$ are iid positive random variables with mean 1, and $\{h_k\}$ and $\{\epsilon_k\}$ are mutually independent.

Deo et al. (2009) demonstrated that long memory in durations propagates to long memory in the counting process $N(t)$, where $N(t)$ counts the number of transactions in the time interval $(0, t]$. In particular, if the durations are generated by an LMSD model with memory parameter $d_\tau \in (0, \frac{1}{2})$, then $N(t)$ is long-range count–dependent with the same memory parameter, in the sense that $\operatorname{var} N(t) \sim Ct^{2d_\tau+1}$ as $t \to \infty$. This long-range count dependence then propagates to the realized volatility, as studied by Deo et al. (2009).

We now describe the tick-time return interactions that yield cointegration in our model. Suppose that there are two assets, 1 and 2, and that each log price is affected by two types of disturbances when a transaction occurs. These disturbances are the value shocks, $\{e_{i,k}\}$, and the microstructure noise, $\{\eta_{i,k}\}$, for asset $i = 1, 2$. The subscript i, k pertains to the kth transaction of asset i. The value shocks are iid and represent permanent contributions to the intrinsic log value of the assets, which in the absence of feedback effects is a martingale with respect to full information, both public and private (see Amihud and Mendelson 1987; Glosten 1987). The microstructure shocks represent

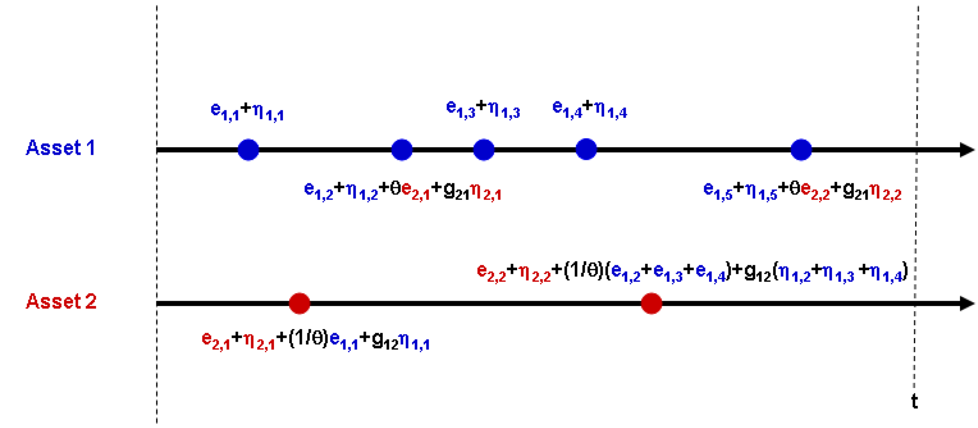

Figure 1. Changes in log prices. The online version of this figure is in color.

the remaining contributions to the observed log prices, along similar lines as the noise process considered by Amihud and Mendelson (1987), reflecting transitory price fluctuations due to, for example, liquidity impact of orders. We assume that the mth tick-time return of asset 1 incorporates not only its own current disturbances $e_{1,m}$ and $\eta_{1,m}$, but also weighted versions of all intervening disturbances of asset 2 that were originally introduced between the $(m-1)$th and mth transactions of asset 1. The weight for the value shocks, denoted by θ, may differ from the weight for the microstructure noise, denoted by g_{21} (the impact from asset 2 to asset 1). We similarly define the mth tick-time return of asset 2, but the weight for the value shocks from asset 1 to asset 2 is $(1/\theta)$, and the corresponding weight for the microstructure noise is denoted by g_{12}. The choice of the second impact coefficient $(1/\theta)$ is necessary for the two log-price series to be cointegrated. In general, if the two series are not cointegrated, then this constraint is not required.

Figure 1 illustrates the mechanism by which tick-time returns are generated in our model. All disturbances originating from asset 1 are shown in blue, and all disturbances originating from asset 2 are shown in red. When the first transaction of asset 1 occurs, a value shock $e_{1,1}$ and a microstructure disturbance $\eta_{1,1}$ are introduced. The first transaction of asset 2 follows in clock time, and because the first transaction of asset 1 occurred before it, the return for this transaction is $(e_{2,1}+\eta_{2,1}+\frac{1}{\theta}e_{1,1}+g_{12}\eta_{1,1})$, that is, the sum of the first value shock of asset 2, $e_{2,1}$, the first microstructure disturbance of asset 2, $\eta_{2,1}$, and a feedback term from the first transaction of asset 1 whose disturbances are $e_{1,1}$ and $\eta_{1,1}$, weighted by the corresponding feedback impact coefficients $\frac{1}{\theta}$ and g_{12}. In the figure, both log-price processes evolve until time t. Notice that the third return of asset 1 contains no feedback term from asset 2, because there is no intervening transaction of asset 2. The second return of asset 2 includes its own current disturbances $(e_{2,2}, \eta_{2,2})$ as well as six weighted disturbances $(e_{1,2}, e_{1,3}, e_{1,4}, \eta_{1,2}, \eta_{1,3}$, and $\eta_{1,4})$ from asset 1, because there are three intervening transactions of asset 1.

At a given clock time t, most of the disturbances of asset 1 are incorporated into the log price of asset 2 and vice-versa. There is an *end effect*, however. The problem is readily seen in the figure: because the fifth transaction of asset 1 occurred after the last transaction of asset 2 before time t, the most recent asset 1 disturbances $e_{1,5}$ and $\eta_{1,5}$ are not incorporated in the log price of asset 2 at time t. Eventually, at the next transaction of asset 2, which will occur after time t, these two disturbances will be incorporated. But this end effect may be present at any given time t. We handle this end effect explicitly in all derivations in this article.

Our model for the log prices for all nonnegative real t is then given by

$$\begin{aligned}\log P_{1,t} &= \sum_{k=1}^{N_1(t)}(e_{1,k}+\eta_{1,k}) + \sum_{k=1}^{N_2(t_{1,N_1(t)})}(\theta e_{2,k}+g_{21}\eta_{2,k}),\\ \log P_{2,t} &= \sum_{k=1}^{N_2(t)}(e_{2,k}+\eta_{2,k}) + \sum_{k=1}^{N_1(t_{2,N_2(t)})}\left(\frac{1}{\theta}e_{1,k}+g_{12}\eta_{1,k}\right),\end{aligned} \tag{1}$$

where $t_{i,k}$ is the clock time for the kth transaction of asset i and $N_i(t)$ $(i=1,2)$ are counting processes that count the total number of transactions of asset i up to time t. Later we impose specific conditions on $\{e_{i,k}\}$, $\{\eta_{i,k}\}$, and $N_i(t)$. Note that (1) implies that $\log P_{1,0}=\log P_{2,0}=0$, the same standardization used by Stock and Watson (1988) and others. The quantity $N_2(t_{1,N_1(t)})$ represents the total number of transactions of asset 2 occurring up to the time $(t_{1,N_1(t)})$ of the most recent transaction of asset 1. An analogous interpretation holds for the quantity $N_1(t_{2,N_2(t)})$.

To exhibit the various components of our model, we rewrite (1) as

$$\begin{aligned}\log P_{1,t} = &\underbrace{\left(\sum_{k=1}^{N_1(t)}e_{1,k}+\sum_{k=1}^{N_2(t)}\theta e_{2,k}\right)}_{\text{common component}}\\ &+\underbrace{\left(\sum_{k=1}^{N_1(t)}\eta_{1,k}+\sum_{k=1}^{N_2(t)}g_{21}\eta_{2,k}\right)}_{\text{microstructure component}}\\ &-\underbrace{\sum_{k=N_2(t_{1,N_1(t)})+1}^{N_2(t)}(\theta e_{2,k}+g_{21}\eta_{2,k}),}_{\text{end effect}}\end{aligned} \tag{2}$$

$$\log P_{2,t} = \underbrace{\left(\sum_{k=1}^{N_1(t)} \frac{1}{\theta} e_{1,k} + \sum_{k=1}^{N_2(t)} e_{2,k}\right)}_{\text{common component}} + \underbrace{\left(\sum_{k=1}^{N_1(t)} g_{12}\eta_{1,k} + \sum_{k=1}^{N_2(t)} \eta_{2,k}\right)}_{\text{microstructure component}} \underbrace{- \sum_{k=N_1(t_{2,N_2(t)})+1}^{N_1(t)} \left(\frac{1}{\theta} e_{1,k} + g_{12}\eta_{1,k}\right)}_{\text{end effect}}.$$

The common component is a martingale and thus is $I(1)$. We show that the microstructure components are $I(1+d_\eta)$, so these components are less persistent than the common component. The end-effect terms are random sums over time periods that are $O_p(1)$ as $t \to \infty$ [see (A.10) to (A.12)] and thus are negligible compared with all other terms. Because both $\log P_{1,t}$ and $\log P_{2,t}$ are $I(1)$ (see Theorem 1) and the linear combination $\log P_{1,t} - \theta \log P_{2,t}$ is $I(1+d_\eta)$ as defined in Section 3, the log-price series are cointegrated (see Theorems 3, 4, and 5).

Frijns and Schotman (2006) considered a mechanism for generating quotes in tick time that is similar to the mechanism shown in Figure 1; however, they condition on durations, whereas we endogenize them in our model (1). Furthermore, their model implies standard cointegration, with a cointegrating parameter known to be 1 and a single value shock component.

Throughout the article, unless noted otherwise, we make the following assumptions for our theoretical results. The duration processes $\{\tau_{i,k}\}$ of asset i $(i = 1, 2)$, are assumed to have long memory with memory parameters $d_{\tau_1}, d_{\tau_2} \in (0, \frac{1}{2})$, to reflect the empirically observed persistence in durations and the resulting realized volatility. Specifically, the $\{\tau_{i,k}\}$ are assumed to satisfy the assumptions in theorem 1 of Deo et al. (2009), which are very general and would allow, for example, the LMSD model of Deo, Hsieh, and Hurvich (2006).

We assume that the $\{e_{i,k}\}$ are mutually independent iid processes with mean 0 and variance $\sigma^2_{i,e}$ $(i = 1, 2)$. We also assume that the $\{\eta_{i,k}\}$ are mutually independent, with mean 0 and memory parameter d_{η_i}. For notational convenience, we set $d_{\eta_1} = d_{\eta_2} = d_\eta$ in our theoretical results. All theorems will continue to hold, however, when d_{η_1} and d_{η_2} are distinct, simply by replacing d_η with $d^*_\eta = \max(d_{\eta_1}, d_{\eta_2})$. For Theorem 6, which establishes the consistency of the OLS estimator of θ, we further assume $\{e_{i,k}\}$ to be $\mathrm{N}(0, \sigma^2_{i,e})$.

We assume that $\{\tau_{1,k}\}$ and $\{\tau_{2,k}\}$ are independent of all disturbance series $\{e_{1,k}\}$, $\{e_{2,k}\}$, $\{\eta_{1,k}\}$, and $\{\eta_{2,k}\}$, which we assume to be mutually independent. But we do not require that $N_1(\cdot)$ and $N_2(\cdot)$ be mutually independent, nor that $\{\tau_{1,k}\}$ and $\{\tau_{2,k}\}$ be mutually independent. This is in keeping with recent literature suggesting that feedback occurs between the counting processes (see, e.g., Nijman, Spierdijk, and Soest 2004; Bowsher 2007; and references therein).

2.1 Economic Justification for the Model

Here we provide some economic rationale for the transaction-level return interactions leading to model (1). This supplements our earlier discussion around Figure 1 on the formal mechanism for price formation. After a brief data analysis affirming the potential usefulness of certain flexibilities of the model, we compare and contrast the model with a clock-time model proposed by Hasbrouck (1995), and then propose a transaction-level generalization of Hasbrouck's definition of the information share of a market.

The model (1) is potentially economically appropriate for pairs of measured prices that are both affected by the same information shocks (i.e., value shocks), possibly in different ways. Examples include buy prices and sell prices of a single stock, prices of two classified stocks (with different voting rights) from a given company, prices of two different stocks 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, and Treasury bond prices at different maturities.

The fundamental (value) prices at time t are an accumulation of information shocks. If we ignore the end effects, then these fundamental prices may be thought of as the common components in (2). More precisely, the fundamental prices may be obtained by setting the microstructure shocks in (1) to 0. For definiteness, consider the example of buy prices (asset 1) and sell prices (asset 2) of a single stock. Information shocks may be generated on either the buy side or the sell side. According to the model, each buy transaction generates its own information shock, as does each sell transaction. Furthermore, these shocks spill over from the side of the market in which they originated to the other side. Clearly, shocks originating from the sell side of the market cannot be impounded into the buy price until there is a transaction on the buy side. Similarly, in the absence of information arrivals (transactions) on the sell side, any string of intervening information shocks from the buy side will render the most recent sell price stale, until the intervening buy-side shocks are actually impounded into the sell-side price at the next sell transaction.

Shocks spilling over from the buy side to the sell side are weighted by $1/\theta$, while those spilling from the sell side to the buy side are weighted by θ. When $\theta = 1$, shocks spill over from one side to the other in an identical way, and there is just a single fundamental price, shared by both the buy side and sell side. In general, as can be seen from (1), the fundamental (log) price for the buy side is an accumulation of information shocks from both the buy side and the sell side, with the sell-side shocks weighted by θ. Ignoring end effects, as can be seen from (2), the common component on the buy side is proportional to the common component on the sell side, and the constant of proportionality is θ.

Analogous interactions take place on the microstructure shocks, such that a microstructure shock originating on the buy side spills over to the sell side with weight g_{12}, and the opposite spillover occurs with weight g_{21}. Even in the absence of spillover of the microstructure shocks ($g_{12} = g_{21} = 0$), the difference between the buy price and θ times the sell price is (except for end effects) an accumulation of microstructure shocks. It seems to be in accordance with the economic connotation of the term "microstructure" that the microstructure shocks be transitory, that is, the aggregate of microstructure shocks be stochastically of smaller order than the aggregate of fundamental shocks, as $t \to \infty$. This will occur if and only if the microstructure shocks have a smaller memory parameter than the fundamental shocks ($d_\eta < 0$), as we assume. Cointegration arises as a consequence of the spillover of the fundamental shocks, together with the assumption $d_\eta < 0$. The spillover of the fundamental shocks induces the common component, while the assumption $d_\eta < 0$ ensures that the cointegrating error arising from the microstructure is less persistent than the common component.

Two questions that might be raised in the context of model (1) are whether there are situations in which the two prices are affected by information in different ways, so that the cointegrating parameter is not equal to 1, and whether it is helpful in practice to allow for fractional cointegration as opposed to standard cointegration. To address these questions, we briefly present some results of a preliminary data analysis. We considered clock-time option best-available-bid prices and underlying best-available-bid prices for IBM on the NYSE at 390 1-minute intervals from 9:30 a.m. to 4 p.m. on May 31, 2007. We originally analyzed 74 different options, but removed 5 from consideration because they had either at least one zero bid price during the day or a constant bid price throughout the day. For the remaining 69 options, we regressed the log stock bid price on the log option bid price, and constructed a semiparametric GPH estimator (Geweke and Porter-Hudak 1983) of the memory parameter of the residuals. The least squares regression slopes ranged from −0.21 to 0.39, with a mean of 0.04 and a standard deviation of 0.13. This suggests that information affected the two prices in different ways for all 69 options. For the GPH estimators, we used $390^{0.5}$ for the number of frequencies; this resulted in an approximate standard error for the GPH estimator of 0.19. The GPH estimator for the log stock bid price was 1.02. The GPH estimators for the residuals ranged from 0.05 to 1.14, with a mean of 0.55 and a standard deviation of 0.28. Of the 69 sets of residuals, 62 yielded a GPH estimator <1, with 54 <0.75, 42 <0.6, and 18 between 0.4 and 0.6. These results suggest the presence of cointegration in most cases, and also imply that the cointegration in some of these cases may be fractional instead of standard.

It is instructive to compare and contrast our model (1) with the clock-time model of Hasbrouck (1995), in which a single security is traded on several markets and different market prices share an identical random-walk component. To facilitate comparisons with the bivariate model (1), suppose that there are two markets. Then for all nonnegative integers j, the clock-time log stock prices at time j on the two different markets, are given by Hasbrouck's model as

$$\begin{aligned} \log P_{1,j} &= \log P_{1,0} + \sum_{s=1}^{j} (\psi_1 \tilde{e}_{1,s} + \psi_2 \tilde{e}_{2,s}) + v_{1,j}, \\ \log P_{2,j} &= \log P_{2,0} + \sum_{s=1}^{j} (\psi_1 \tilde{e}_{1,s} + \psi_2 \tilde{e}_{2,s}) + v_{2,j}, \end{aligned} \quad (3)$$

where $\log P_{1,0}$ and $\log P_{2,0}$ are constants, $(\tilde{e}_{1,s}, \tilde{e}_{2,s})'$ is a mean-0 vector of serially uncorrelated disturbances with covariance matrix Ω, $\psi = (\psi_1, \psi_2)$ are the weights for $\tilde{e}_{1,s}, \tilde{e}_{2,s}$, and $\{(v_{1,j}, v_{2,j})'\}$ is a mean-0 stationary bivariate time series. The quantity $\tilde{e}_{i,s}$ $(i = 1, 2)$ may be considered the fundamental shock originating from the ith market. Hasbrouck (1995) estimated the model on data using a 1-second sampling interval.

Both models (1) and (3) induce a common component, as well as cointegration. Both contain spillover of the fundamental shocks from one market to the other. In model (3) the spillover is the same in both directions, so the common components are identical and the cointegrating parameter is 1. In contrast, in model (1) the cointegrating parameter need not be equal to 1. In model (3) the cointegrating error is $I(0)$, while in model (1) the cointegrating error is allowed to be $I(1+d_\eta)$ for any d_η with $-1 \le d_\eta < 0$.

In model (3) the contemporaneous correlation between the fundamental shocks originating from the two markets is allowed to be nonzero, (i.e., Ω is allowed to be nondiagonal), whereas in model (1) the two fundamental shock series are assumed to be independent. Note, however, that in the transaction-level model (1), the kth transactions of the two assets will (almost surely) occur at different clock times, so any correlation between the two fundamental shocks $e_{1,k}$ and $e_{2,k}$ would not be contemporaneous in clock time. This provides a motivation for our assumption that $\{e_{1,k}\}$ and $\{e_{2,k}\}$ are mutually independent. An economic motivation for this assumption stems from the following remarks of Hasbrouck (1995, p. 1183): "In practice, market prices usually change sequentially: a new price is posted in one market, and then the other markets respond. If the observation interval is so long that the sequencing cannot be determined, however, the initial change and the response will appear to be contemporaneous. Therefore, one obvious way of minimizing the correlation is to shorten the interval of observation." Because model (1) is defined in continuous time, the interval of observation is effectively zero, so at least under the idealized assumptions that there are no truly simultaneous transactions on the two markets and that the time stamps for the transactions are exact, the assumption of mutual independence would be economically reasonable.

In the remainder of this section, we discuss the information share, originally defined by Hasbrouck (1995) to measure how market information that drives stock prices is distributed across different exchanges. Hasbrouck (1995) defined the information share of market i based on model (3) as $S_i = (\psi_i^2 \Omega_{ii})/(\psi \Omega \psi')$, which is the proportional contribution from market i to the total fundamental innovation variance. Only the random-walk component is used in constructing the information share, because this is the only permanent component. As Hasbrouck (1995) discussed, because Ω may not be diagonal, only a bound for the information share can be estimated. Here we propose a transaction-level generalization of the concept of information share based on model (1), which is directly estimable because of our assumption of mutual independence of the transaction-level fundamental disturbance series. The information share proposed by Hasbrouck (1995) measures how the price discovery of one security is fulfilled across difference exchanges. In contrast, our information share instead measures how the price-driving information of a security is distributed between buy versus sell trades in a market. The ideas are similar. Indeed, as Hasbrouck (1995) noted, his model can be extended to model bid and ask price dynamics. Nevertheless, as discussed earlier, our model ultimately is a tick-level model that differs from existing clock-time models, including that of Hasbrouck (1995).

For model (1), we define the information share as follows. For a given clock-time sampling interval Δt, the information share of asset i is given by

$$S_{1,C} = \frac{\operatorname{var}(\sum_{k=N_1((j-1)\Delta t)+1}^{N_1(j\Delta t)} e_{1,k})}{\operatorname{var}(\sum_{k=N_1((j-1)\Delta t)+1}^{N_1(j\Delta t)} e_{1,k} + \theta \sum_{k=N_2((j-1)\Delta t)+1}^{N_2(j\Delta t)} e_{2,k})} = \frac{\lambda_1 \sigma_{1,e}^2}{\lambda_1 \sigma_{1,e}^2 + \theta^2 \lambda_2 \sigma_{2,e}^2}, \tag{4}$$

$$S_{2,C} = \frac{\theta^2 \lambda_2 \sigma_{2,e}^2}{\lambda_1 \sigma_{1,e}^2 + \theta^2 \lambda_2 \sigma_{2,e}^2},$$

where λ_i is the intensity of the counting process $N_i(\cdot)$ (see Daley and Vere-Jones 2003) and represents the intensity of trading (level of market activity) of asset i. The ultimate expressions for $S_{i,C}$ do not depend on the sampling interval Δt. Note that only the common component in (2) is used to evaluate the information share, as was done by Hasbrouck (1995). As $\lambda_1/\lambda_2 \to \infty$, $S_{1,C}$ approaches 1 and $S_{2,C}$ approaches 0. This is consistent with general intuition; an actively traded security should reveal more information than a thinly traded one. Indeed, Hasbrouck (1995) found that for the 30 Dow Jones stocks, the preponderance of the price discovery occurs at the NYSE and the majority of the transactions occur on the NYSE. The information share $S_{i,C}$ can be estimated using the transaction-level method of moments, as described in Section 8. We present estimates of $S_{i,C}$ computed from transaction-level data in Section 11.

3. CONDITIONS ON THE MICROSTRUCTURE NOISE FOR FRACTIONAL AND STANDARD COINTEGRATION

In this section we consider three types of cointegration: weak fractional, strong fractional, and standard cointegration. We describe the conditions assumed for each of these three cases separately.

The weak fractional cointegration case corresponds to $d_\eta \in (-\frac{1}{2}, 0)$. In this case, we require the following condition, stated for a generic process $\{\eta_k\}$:

Condition A. For $d_\eta \in (-\frac{1}{2}, 0)$, $\{\eta_k\}$ is a weakly stationary mean-0 process with memory parameter d_η in the sense that the spectral density $f(\lambda)$ satisfies

$$f(\lambda) = \tilde{\sigma}^2 C^* \lambda^{-2d_\eta}(1 + O(\lambda^\beta)) \quad \text{as } \lambda \to 0^+$$

for some β with $0 < \beta \le 2$, where $\tilde{\sigma}^2 > 0$ and $C^* = (d_\eta + \frac{1}{2})\Gamma(2d_\eta + 1)\sin((d_\eta + \frac{1}{2})\pi)/\pi > 0$.

Condition A, which was originally used in a semiparametric context by Robinson (1995), is very general, specifying only the behavior of the spectral density in a neighborhood of zero frequency. The condition is satisfied by all parametric long-memory models that we have seen in the literature, including the ARFIMA(p, d_η, q) model with $p \ge 0$, $q \ge 0$, and $d_\eta \in (-\frac{1}{2}, 0)$. In the ARFIMA case, $\beta = 2$. Condition A also allows the possibility for seasonal long memory, that is, poles or 0s of $f(\lambda)$ at nonzero frequencies.

The strong fractional cointegration case corresponds to $d_\eta \in (-1, -\frac{1}{2})$. For this case, we assume the following:

Condition B. For $d_\eta \in (-1, -\frac{1}{2})$, $\eta_k = \varphi_k - \varphi_{k-1}$, $k = 1, 2, \ldots,$ where $\varphi_0 = 0$ and $\{\varphi_k\}_{k=1}^\infty$ is a mean-0, weakly stationary long-memory process with memory parameter $d_\varphi = d_\eta + 1 \in (0, \frac{1}{2})$ in the sense that its autocovariances satisfy

$$\operatorname{cov}(\varphi_k, \varphi_{k+j}) = Kj^{2d_\varphi - 1} + O(j^{2d_\varphi - 3}), \qquad j \ge 1, \tag{5}$$

where $K > 0$.

By theorem 1 of Lieberman and Phillips (2006), any stationary, invertible ARFIMA(p, d_φ, q) process with $d_\varphi \in (0, \frac{1}{2})$ has autocovariances satisfying (5), with $K = 2f^*(0)\Gamma(1 - 2d_\varphi)\sin(\pi d_\varphi)$, where $f^*(0)$ is the spectral density of the ARMA component of the model at zero frequency.

The standard cointegration case corresponds to $d_\eta = -1$. In this case we assume the following:

Condition C. If $d_\eta = -1$, then $\{\eta_k\}_{k=1}^\infty$ is given by $\eta_k = \xi_k - \xi_{k-1}$ with $\xi_0 = 0$. The process $\{\xi_k\}_{k=1}^\infty$ is weakly stationary with mean 0 and autocovariance sequence $\{c_{\xi,r}\}_{r=0}^\infty$, where $c_{\xi,r} = E(\xi_{k+r}\xi_k)$ with exponential decay and $|c_{\xi,r}| \le A_\xi e^{-K_\xi r}$ for all $r \ge 0$, where A_ξ and K_ξ are positive constants.

The assumptions on $\{\xi_k\}$ in Condition C are satisfied by all stationary invertible ARMA models.

4. LONG–TERM MARTINGALE–TYPE PROPERTIES OF THE LOG PRICES

In this section we present the properties of the log-price series generated by model (1). Define $\lambda_i = 1/E^0(\tau_{i,k})$, where E^0 denotes expectation under the Palm distribution (see Deo et al. 2009 for information on the Palm probability measure), that is, the distribution under which the $\{\tau_{i,k}\}$ $(i = 1, 2)$ are stationary. The following two theorems show that the log-price series in model (1) have asymptotic variances that scale like t as $t \to \infty$, as would happen for a martingale, and that their discretized differences are asymptotically uncorrelated as the sampling interval increases, as would occur for a martingale difference series.

Theorem 1. For the log-price series in model (1),

$$\operatorname{var}(\log P_{i,t}) \sim C_i t, \qquad i = 1, 2,$$

as $t \to \infty$, where $C_1 = (\sigma_{1,e}^2 \lambda_1 + \theta^2 \sigma_{2,e}^2 \lambda_2)$ and $C_2 = (\sigma_{2,e}^2 \lambda_2 + \frac{1}{\theta^2}\sigma_{1,e}^2 \lambda_1)$.

For a given sampling interval (equally spaced clock-time period) Δt, the returns (changes in log price) for assets 1 and 2 corresponding to model (1) are

$$\begin{aligned} r_{1,j} &= \sum_{k=N_1((j-1)\Delta t)+1}^{N_1(j\Delta t)} (e_{1,k} + \eta_{1,k}) + \sum_{k=N_2(t_{1,N_1((j-1)\Delta t)})+1}^{N_2(t_{1,N_1(j\Delta t)})} (\theta e_{2,k} + g_{21}\eta_{2,k}), \\ r_{2,j} &= \sum_{k=N_2((j-1)\Delta t)+1}^{N_2(j\Delta t)} (e_{2,k} + \eta_{2,k}) + \sum_{k=N_1(t_{2,N_2((j-1)\Delta t)})+1}^{N_1(t_{2,N_2(j\Delta t)})} \left(\frac{1}{\theta} e_{1,k} + g_{12}\eta_{1,k}\right). \end{aligned} \tag{6}$$

Theorem 2. For any fixed integer $k > 0$, the lag-k autocorrelation of $\{r_{i,j}\}_{j=1}^{\infty}$, $i = 1, 2$, tends to 0 as $\Delta t \to \infty$.

5. PROPERTIES OF THE COINTEGRATING ERROR

Here we show that model (1) implies a cointegrating relationship between the two series, treating the weak and strong fractional as well as standard cointegration cases separately.

Theorem 3. Under model (1) with $d_\eta \in (-\frac{1}{2}, 0)$, the memory parameter of the linear combination $(\log P_{1,t} - \theta \log P_{2,t})$ is $(1 + d_\eta) \in (\frac{1}{2}, 1)$, that is,

$$\mathrm{var}(\log P_{1,t} - \theta \log P_{2,t}) \sim Ct^{2d_\eta+1}$$

as $t \to \infty$, where $C > 0$. In this sense, $\log P_{1,t}$ and $\log P_{2,t}$ are weakly fractionally cointegrated.

We next investigate the standard cointegration case. It is important to note that, unlike in Theorem 3, where we measure the strength of cointegration using the asymptotic behavior of the variance of the cointegrating errors $\mathrm{var}(\log P_{1,t} - \theta \log P_{2,t})$, we need a different measure here, because $\log P_{1,t} - \theta \log P_{2,t}$ is stationary and its variance is constant for all t. Instead, we consider the asymptotic covariance of the cointegrating errors,

$$\mathrm{cov}(\log P_{1,t} - \theta \log P_{2,t}, \log P_{1,t+j} - \theta \log P_{2,t+j})$$

as $j \to \infty$. Here we take t and j to be positive integers; that is, we sample the log-price series using $\Delta t = 1$, without loss of generality.

Theorem 4. Under model (1) with $d_\eta \in (-1, -\frac{1}{2})$, the memory parameter of the cointegrating error $(\log P_{1,t} - \theta \log P_{2,t})$ is $(1 + d_\eta) \in (0, \frac{1}{2})$; that is, for any fixed $t > 0$,

$$\mathrm{cov}(\log P_{1,t} - \theta \log P_{2,t}, \log P_{1,t+j} - \theta \log P_{2,t+j}) \sim j^{2(1+d_\eta)-1}\left[C_1 \Pr\{N_1(t) > 0\} + C_2 \Pr\{N_2(t) > 0\}\right]$$

as $j \to \infty$, where $C_1 > 0$, $C_2 > 0$. In this sense, $\log P_{1,t}$ and $\log P_{2,t}$ are strongly fractionally cointegrated.

We say that a sequence $\{a_j\}$ has *nearly exponential decay* if $a_j/j^{-\alpha} \to 0$ as $j \to \infty$ for all $\alpha > 0$. We say that a time series has *short memory* if its autocovariances have nearly exponential decay.

Theorem 5. Under model (1), with $d_\eta = -1$, the cointegrating error $(\log P_{1,t} - \theta \log P_{2,t})$ has short memory. In this sense, $\log P_{1,t}$ and $\log P_{2,t}$ are cointegrated.

6. LEAST SQUARES ESTIMATION OF THE COINTEGRATING PARAMETER

Assume that the log-price series are observed at integer multiples of Δt. The proposed model (1) becomes (with a minor abuse of notation)

$$\log P_{1,j} = \sum_{k=1}^{N_1(j\Delta t)} (e_{1,k} + \eta_{1,k}) + \sum_{k=1}^{N_2(t_{1,N_1(j\Delta t)})} (\theta e_{2,k} + g_{21}\eta_{2,k}),$$

$$\log P_{2,j} = \sum_{k=1}^{N_2(j\Delta t)} (e_{2,k} + \eta_{2,k}) + \sum_{k=1}^{N_1(t_{2,N_2(j\Delta t)})} \left(\frac{1}{\theta} e_{1,k} + g_{12}\eta_{1,k}\right). \quad (7)$$

We show that the cointegrating parameter θ can be consistently estimated by OLS regression.

Theorem 6. For the discretely sampled log-price series in (7) with normally distributed value shocks $\{e_{1,k}\}$, $\{e_{2,k}\}$, the cointegrating parameter θ can be consistently estimated by $\hat{\theta}$, the OLS estimator obtained by regressing $\{\log P_{1,j}\}_{j=1}^n$ on $\{\log P_{2,j}\}_{j=1}^n$ without intercept. For all $\delta > 0$, as $n \to \infty$, we have

Case I: $d_\eta \in (-\frac{1}{2}, 0)$

$$n^{-d_\eta-\delta}(\hat{\theta} - \theta) \xrightarrow{p} 0,$$

Case II: $d_\eta \in (-1, -\frac{1}{2})$

$$n^{1/2-\delta}(\hat{\theta} - \theta) \xrightarrow{p} 0,$$

Case III: $d_\eta = -1$

$$n^{1-\delta}(\hat{\theta} - \theta) \xrightarrow{p} 0.$$

In the weak fractional cointegration case, $d_\eta \in (-\frac{1}{2}, 0)$, the rate of convergence of $\hat{\theta}$ improves as d_η decreases. In the standard cointegration case where $d_\eta = -1$, the rate is arbitrarily close to n. Phillips and Durlauf (1986) and Stock (1987) have demonstrated the n-consistency (super-consistency) of the OLS estimator of the cointegrating parameter in the standard cointegration case for time series in discrete clock time that are linear with respect to a strong-mixing or iid sequence. We are currently unable to derive the asymptotic distribution of the OLS estimator of the cointegrating parameter in the standard cointegration case for our model, because we cannot rely on the strong-mixing condition on returns. This condition would not be expected to hold in the case of LMSD durations, because these are not strong mixing in tick time. In the strong fractional cointegration case $d_\eta \in (-1, -\frac{1}{2})$, even though we established a rate of $n^{1/2-\delta}$, simulations in Section 10 indicate that the actual rate is faster, at $n^{-d_\eta-\delta}$, in keeping with the rates obtained in the weak fractional and standard cointegration cases.

7. A TICK–LEVEL COINTEGRATING PARAMETER ESTIMATOR

In this section we propose a transaction-level estimator, $\tilde{\theta}$, for the cointegrating parameter θ. It may be argued that the OLS estimator $\hat{\theta}$ discussed in Section 6 is not optimal, because it is constructed based on discretized log-prices and thus uses only partial information. Here we propose a tick-level estimator, $\tilde{\theta}$, that uses the full tick-level price series, $\log P_{1,t}$, $\log P_{2,t}$ for $t \in [0, T]$.

Specifically, let $N(T) = N_1(T) + N_2(T)$ be the pooled counting process of transactions for both asset 1 and asset 2 in the time interval $(0, T]$, and let $\{t_k^\star\}_{k=1}^{N(T)}$ denote the transaction times for the pooled process. The proposed estimator is

$$\tilde{\theta} = \frac{\int_0^T \log P_{1,t} \cdot \log P_{2,t}\, dt}{\int_0^T \log P_{2,t}^2\, dt}$$

$$= \left(\sum_{j=1}^{N(T)-1} \left[\log P_{1,N_1(t_j^\star)} \cdot \log P_{2,N_2(t_j^\star)}\right] \cdot (t_{j+1}^\star - t_j^\star) \right.$$

$$+\left[\log P_{1,N_1(T)} \cdot \log P_{2,N_2(T)}\right] \cdot \left(T - t^{\star}_{N(T)}\right)\Bigg)$$

$$\Bigg/ \Bigg(\sum_{j=1}^{N_2(T)-1} \log P^2_{2,j} \cdot \tau_{2,j+1}$$

$$+ \log P^2_{2,N_2(T)} \cdot \left(T - t_{2,N_2(T)}\right)\Bigg), \qquad (8)$$

where the numerator is a summation over all transactions, adding up the product of the most recent log prices of assets 1 and 2 weighted by the corresponding duration for the pooled process. The denominator of $\tilde{\theta}$ has the same structure, except that the product is now of asset 2 log prices with themselves.

We do not derive asymptotic properties of the estimator $\tilde{\theta}$. Nevertheless, the simulation study presented in Section 10 indicates that the tick-level estimator $\tilde{\theta}$ may outperform the OLS estimator $\hat{\theta}$, having smaller bias, variance, and root mean squared error (RMSE), particularly if the sampling interval Δt for $\hat{\theta}$ is large.

8. METHOD–OF–MOMENTS PARAMETER ESTIMATION

In earlier work (Hurvich and Wang 2009), we proposed a transaction-level parameter estimation procedure for model (1) using the method of moments, based on $\log P_{1,t}$, $\log P_{2,t}$ for $t \in [0, T]$. To conserve space, we omit the complete details on how we constructed the method-of-moments estimators here. We make specific assumptions for the sake of definiteness, although most of these assumptions could be relaxed. Specifically, in constructing our estimates of Θ, we assume Gaussian white noise for the value shocks and a Gaussian ARFIMA$(1, d_\eta, 0)$ process for the microstructure noise when $d_\eta \in (-\frac{1}{2}, 0)$, and also assume that the microstructure noise is the difference of a Gaussian ARFIMA$(1, d_\eta + 1, 0)$ process with the initial value set to 0 when $d_\eta \in (-1, -\frac{1}{2})$. In the standard cointegration case $d_\eta = -1$, we assume that the microstructure noise is the difference of a Gaussian AR(1) process with initial value set to 0. We denote the autoregressive parameters (lag-1 autocorrelations) of the two microstructure noise series by α_1 and α_2. The method-of-moments estimator, $\hat{\Theta} = (\hat{\sigma}^2_{1,e}, \hat{\sigma}^2_{2,e}, \hat{\sigma}^2_{1,\eta}, \hat{\sigma}^2_{2,\eta}, \hat{g}_{21}, \hat{g}_{12}, \hat{d}_{\eta_1}, \hat{d}_{\eta_2}, \hat{\alpha}_1, \hat{\alpha}_2)$, is obtained as the solution to the following system of equations, based on certain specific observed sequences of assets 1 and 2 transactions:

$$\widehat{\text{var}}(\text{second transaction of Sequence 1 1}) = \hat{\sigma}^2_{1,e} + \hat{\sigma}^2_{1,\eta},$$

$$\widehat{\text{var}}(\text{second transaction of Sequence 2 2}) = \hat{\sigma}^2_{2,e} + \hat{\sigma}^2_{2,\eta},$$

$$\widehat{\text{cov}}(\text{first and second transactions of Sequence 1 1}) = \hat{\sigma}^2_{1,\eta}\hat{\rho}_{1,1},$$

$$\widehat{\text{cov}}(\text{first and second transactions of Sequence 2 2}) = \hat{\sigma}^2_{2,\eta}\hat{\rho}_{2,1},$$

$$\widehat{\text{cov}}(\text{first and third transactions of Sequence 1 1 1}) = \hat{\sigma}^2_{1,\eta}\hat{\rho}_{1,2},$$

$$\widehat{\text{cov}}(\text{first and third transactions of Sequence 2 2 2}) = \hat{\sigma}^2_{2,\eta}\hat{\rho}_{2,2}, \qquad (9)$$

$$\widehat{\text{var}}(\text{third transaction of Sequence 1 2 1}) = \hat{\sigma}^2_{1,e} + \hat{\sigma}^2_{1,\eta} + \tilde{\theta}^2\hat{\sigma}^2_{2,e} + \hat{g}^2_{21}\hat{\sigma}^2_{2,\eta},$$

$$\widehat{\text{var}}(\text{third transaction of Sequence 2 1 2}) = \hat{\sigma}^2_{2,e} + \hat{\sigma}^2_{2,\eta} + \frac{1}{\tilde{\theta}^2}\hat{\sigma}^2_{1,e} + \hat{g}^2_{12}\hat{\sigma}^2_{1,\eta},$$

$$\widehat{\text{cov}}(g_{21} \text{ pairs in Sequence 1 2 1 2 2}) = \hat{g}_{21}\hat{\sigma}^2_{2,\eta}\hat{\rho}_{2,2},$$

$$\widehat{\text{cov}}(g_{12} \text{ pairs in Sequence 2 1 2 1 1}) = \hat{g}_{12}\hat{\sigma}^2_{1,\eta}\hat{\rho}_{1,2},$$

where $\widehat{\text{var}}$ and $\widehat{\text{cov}}$ are the sample variance and covariance, $\rho_{i,j}$ is the lag-j autocorrelation of the microstructure disturbances $\{\eta_{i,k}\}$ for asset $i = 1, 2$, and $\hat{\rho}_{i,j}$ is the resulting estimate of $\rho_{i,j}$. In ealrier work (Hurvich and Wang 2009) we established the following theorem on consistency of the estimator.

Theorem 7. The method-of-moments estimator $\hat{\Theta}$ is consistent, that is,

$$\hat{\Theta} \xrightarrow{p} \Theta \quad \text{as } T \to \infty,$$

where $\Theta = (\sigma^2_{1,e}, \sigma^2_{2,e}, \sigma^2_{1,\eta}, \sigma^2_{2,\eta}, g_{21}, g_{12}, d_{\eta_1}, d_{\eta_2}, \alpha_1, \alpha_2)$.

Motivated by computational constraints that limit the size of the data set we are analyzing, we propose an alternative ad hoc estimator $\tilde{\Theta}$, which performed reasonably well in simulations reported earlier (Hurvich and Wang 2009). We start by taking the ratio of the third and fifth equations in (9), giving us a numerical estimate of $\rho_{1,1}/\rho_{1,2}$, which we denote by $\tilde{\rho}_{1,1}/\tilde{\rho}_{1,2}$. Then, on a grid of values of (d, α), we compute the corresponding ratio $\rho_{1,1}/\rho_{1,2}$ for the ARFIMA$(1, d, 0)$ model with parameters (d, α). We use the algorithm of Bertelli and Caporin (2002) to compute $\rho_{1,1}$ and $\rho_{1,2}$, since there is no attractive closed form for the autocovariances of an ARFIMA$(1, d, 0)$ process. The supports of (d, α) are $(-1, -\frac{1}{2}) \cup (-\frac{1}{2}, 0)$ and $(-1, 1)$, respectively. Next, we construct $\tilde{d}_{\eta_1}$ and $\tilde{\alpha}_1$ such that $|\frac{\tilde{\rho}_{1,1}}{\tilde{\rho}_{1,2}} - \frac{\rho_{1,1}}{\rho_{1,2}}|$ is minimized, that is,

$$(\tilde{d}_{\eta_1}, \tilde{\alpha}_1) = \min_{d,\alpha}\left|\frac{\tilde{\rho}_{1,1}}{\tilde{\rho}_{1,2}} - \frac{\rho_{1,1}}{\rho_{1,2}}\right|.$$

In addition, $\tilde{\rho}_{1,1}$ and $\tilde{\rho}_{1,2}$ are obtained. Similarly, we obtain $(\tilde{d}_{\eta_2}, \tilde{\alpha}_2)$, $\tilde{\rho}_{2,1}$, and $\tilde{\rho}_{2,2}$. We then obtain the remaining parameter estimates in $\tilde{\Theta}$ from (9). Using $\tilde{\rho}_{1,1}$ in the third equation of (9), we get $\tilde{\sigma}^2_{1,\eta}$, which we then use in the first equation to get $\tilde{\sigma}^2_{1,e}$. Similarly, we obtain $\tilde{\sigma}^2_{2,\eta}$ and $\tilde{\sigma}^2_{2,e}$. Next, we obtain $\tilde{g}^2_{21}$ and $\tilde{g}^2_{12}$ based on the seventh and eighth equations in (9). By this point, we have obtained $\tilde{g}^2_{21}$ and $\tilde{g}^2_{12}$, as well as all entries of $\tilde{\Theta}$ except for $\tilde{g}_{12}$ and $\tilde{g}_{21}$ using only the first eight equations of (9). Finally, we use the last two equations in (9) (which are inherently less accurate than the others, because they are based on five-trade sequences) to determine the signs of $\tilde{g}_{21}$ and $\tilde{g}_{12}$.

9. MODEL SPECIFICATION TEST

In this section we propose a specification test for our model (1), based on Theorem 6. The idea is that, according to Theorem 6, if the model (1) is correctly specified, then the OLS estimator is consistent for any particular sampling interval Δt. Suppose that we choose two sampling intervals, Δt_1 and Δt_2, and denote the corresponding OLS estimators by $\hat{\theta}^{\Delta t_1}$ and $\hat{\theta}^{\Delta t_2}$. Because both estimators are consistent, their difference must converge in probability to 0. Thus we propose a specification test to test whether this difference is significantly different from 0. The test is semiparametric, in that model (1) makes no parametric assumptions on either the duration or the microstructure noise.

To implement the test, we divide the entire time span, say 1 year, into K nonoverlapping subperiods, for example, into months. Within subperiod k ($k = 1, \dots, K$), we sample every Δt_1 to obtain a bivariate log-price series $\{\log P^{\Delta t_1}_{1,j,k}, \log P^{\Delta t_1}_{2,j,k}\}$, where $\log P^{\Delta t_1}_{1,j,k}$ is the jth sampled asset 1 log price in subperiod k using sampling interval Δt_1 and similarly for $\log P^{\Delta t_1}_{2,j,k}$. Based on these results, we obtain an OLS cointegrating parameter estimate, $\hat{\theta}^{\Delta t_1}_k$, and similarly we sample every Δt_2 to obtain $\{\log P^{\Delta t_2}_{1,j,k}, \log P^{\Delta t_2}_{2,j,k}\}$, and then $\hat{\theta}^{\Delta t_2}_k$. Repeating the procedure through all K subperiods, we obtain sequences $\{\hat{\theta}^{\Delta t_1}_k\}^K_{k=1}$ and $\{\hat{\theta}^{\Delta t_2}_k\}^K_{k=1}$. The proposed test statistic is

$$\hat{\delta}_{12} = \frac{\text{sample mean of } \{\hat{\delta}_{12,k}\}}{\sqrt{\frac{1}{K} \cdot \text{sample variance of } \{\hat{\delta}_{12,k}\}}},$$

where $\hat{\delta}_{12,k} = \hat{\theta}^{\Delta t_2}_k - \hat{\theta}^{\Delta t_1}_k$ ($k = 1, \dots, K$). The distribution of the test statistic under the null hypothesis that all model assumptions are correctly specified is unknown; however, the critical value for the test, as well as the corresponding distribution of the test statistic, can be simulated under the null hypothesis, based on the estimated parameter values.

The power of the proposed specification test is unknown, because the precise alternative hypothesis is not specified. As discussed by Hausman (1978), a sufficient requirement for the specification test to be consistent is that the two estimators, $\hat{\theta}^{\Delta t_1}$ and $\hat{\theta}^{\Delta t_2}$, have different probability limits under the alternative.

In Section 10, we first investigate the simulation-based distributions of the test statistics for empirically relevant parameter values, We then compute critical values for the specification test on the empirical example, Tiffany (TIF), which we use in the data analysis in Section 11.

10. SIMULATIONS

10.1 Estimation of the Cointegrating Parameter: $\hat{\theta}$ and $\tilde{\theta}$

We study the performance of $\hat{\theta}$ and $\tilde{\theta}$ in a simulation study carried out as follows. First, we simulate two mutually independent duration process $\{\tau_{i,k}\}$ for asset $i = 1, 2$. Note that for simplicity, we assume that the two duration processes are mutually independent, although this is not required by our theoretical results. Each duration process follows the LMSD model,

$$\tau_{i,k} = e^{h_{i,k}} \epsilon_{i,k},$$

where the $\{\epsilon_{i,k}\}$ are iid positive random variables with all moments finite and the $\{h_{i,k}\}$ are a Gaussian long-memory series with mean 0 and common memory parameter d_τ. Based on empirical work of Deo, Hsieh, and Hurvich (2006), we choose $d_{\tau_1} = d_{\tau_2} = 0.45$. Here we assume that the $\{\epsilon_{i,k}\}$ follow an exponential distribution with unit mean. We simulate the $\{h_{i,k}\}$ from a Gaussian ARFIMA$(0, d_\tau, 0)$ model, with innovation variances chosen such that the mean of the log durations matches those observed in the Tiffany series used in Section 11. Using the simulated durations $\{\tau_{i,k}\}$, $i = 1, 2$, we obtain the corresponding counting processes $\{N_i(t)\}$, using $t_{i,1} =$ Uniform$[0, \tau_{i,1}]$. This ensures that the counting processes are stationary.

Next, we generate mutually independent disturbance series $\{e_{1,k}\}$, $\{e_{2,k}\}$, $\{\eta_{1,k}\}$, and $\{\eta_{2,k}\}$. Here $\{e_{i,k}\}, i = 1, 2$, are iid Gaussian with mean 0. For simplicity, the memory parameters of the microstructure noise series are assumed to be the same: $d_{\eta_1} = d_{\eta_2} = d_\eta$. When $d_\eta \in (-\frac{1}{2}, 0)$, the $\{\eta_{i,k}\}$ are given by ARFIMA$(1, d_\eta, 0)$. When $d_\eta \in (-1, -\frac{1}{2})$, $\{\eta_{i,k}\}$ are simulated as the differences of ARFIMA$(1, d_\eta + 1, 0)$, and when $d_\eta = -1$, $\{\eta_{i,k}\}$ are simulated as the differences of two independent mean-0 Gaussian AR(1) series, $\{\xi_{i,k}\}$. The disturbance variances are $\text{var}(e_{i,k}) = 4 \times 10^{-6}$ and $\text{var}(\eta_{i,k}) = 1 \times 10^{-6}$ for $i = 1, 2$. We set $g_{21} = g_{12} = 1$. We select these particular values because they are close to the corresponding parameter estimates based on several stocks that we have analyzed empirically.

We then construct the log-price series $\{\log P_{i,j}\}^n_{j=1}$, $i = 1, 2$, from (1), using a fixed sampling interval Δt. We calculate the estimated cointegrating parameter $\hat{\theta}$ by regressing $\{\log P_{1,j}\}^n_{j=1}$ on $\{\log P_{2,j}\}^n_{j=1}$, using OLS without intercept. We construct the tick-level estimator, $\tilde{\theta}$, according to (8), using the entire tick-level price series.

In the study, we fixed the cointegrating parameter at $\theta = 1$. We considered various values of the parameters Δt and the sample size n. We consider time as being measured in seconds, so that $\Delta t = 300$ corresponds to observing the price series every 5 minutes; in this case, $n = 390$ would correspond to 1 week of data. (There are 6.5 trading hours each day, so sampling every 5 minutes yields 78 observations per day.) For each parameter configuration, we generated 1000 realizations of the log-price series. The results are summarized in Table 1.

As the sample size n increases, the bias, the standard deviation, and the RMSE of $\hat{\theta}$ decrease, as seen in block A. This is consistent with Theorem 6. We report results only for $d_\eta = -0.75$; however, we found similar patterns for $d_\eta = -0.25, -1$.

In A2 together with block B, we fixed the total time span $T = n\Delta t$, while varying the sampling interval Δt and n. For this specific set of empirically relevant parameter values, the impact of increasing Δt was not obvious until $\Delta t = 9{,}000$, which corresponds to the commonly used sampling frequency of 1 day. Both the standard deviation and the RMSE deteriorated as Δt increased. We found the same pattern for $d_\eta = -0.25, -0.75, -1$, although we report only results for $d_\eta = -0.75$ here. In addition, the bias of $\hat{\theta}$ decreased as the sampling interval Δt increased, possibly because the end effect is

Table 1. Simulation results for estimating the cointegrating parameter

		Simulation parameters				$\hat{\theta}$			$\tilde{\theta}$		
Block	Case	$n\Delta t$	Δt (sec)	d_η	n	Mean	SD	RMSE	Mean	SD	RMSE
A	A1	39,000	300	−0.75	130	0.9510	0.1192	0.1288	0.9511	0.1192	0.1288
	A2	117,000	300	−0.75	390	0.9769	0.0576	0.0620	0.9768	0.0575	0.0620
	A3	351,000	300	−0.75	1170	0.9875	0.0350	0.0371	0.9876	0.0349	0.0370
	A4	1,053,000	300	−0.75	3,510	0.9957	0.0142	0.0148	0.9957	0.0142	0.0148
B	B1	117,000	10	−0.75	11,700	0.9768	0.0575	0.0620	0.9768	0.0575	0.0620
	B2	117,000	60	−0.75	1,950	0.9768	0.0575	0.0620	0.9768	0.0575	0.0620
	B3	117,000	1800	−0.75	65	0.9770	0.0592	0.0635	0.9768	0.0575	0.0620
	B4	117,000	9,000	−0.75	13	0.9780	0.0761	0.0792	0.9768	0.0575	0.0620
	B5	117,000	23,400	−0.75	5	0.9876	0.1148	0.1154	0.9768	0.0575	0.0620

not as important when Δt is large. Finally, in terms of RMSE, $\tilde{\theta}$ performed no worse than $\hat{\theta}$, and performed much better than $\hat{\theta}$ when Δt was large.

We also performed simulations related to the convergence rate of $\hat{\theta}$. In Theorem 6, when $d_\eta \in (-1, -\frac{1}{2})$, the convergence rate is arbitrarily close to $\sqrt{n}$ and does not depend on d_η. But simulations indicate a faster rate in this strong fractional cointegration case. For example, when $d_\eta = -0.75$, we simulated the log-price series in discrete clock-time using sample sizes n ranging from 1000 to 20,000 with an equally spaced increment of 800. The variance of $\hat{\theta}$ for each value of n was obtained based on 1000 realizations. The estimated convergence rate of $\hat{\theta}$ was $n^{0.78}$, obtained from the estimated slope in a log-log plot of these simulated variances versus the corresponding sample sizes. We ran similar simulations for other values of d_η. Based on these, we conjecture that the actual rate of convergence for $\hat{\theta}$ was $n^{-d_\eta-\delta}$, in keeping with the rates obtained in the weak fractional and standard cointegration cases.

10.2 Specification Test

We performed a simulation study for the specification test proposed in Section 9. We used two sets of empirically relevant parameter values to investigate the simulation-based distribution of the test statistic $\hat{\delta}$.

We chose empirically relevant parameter values to investigate the simulation-based distribution of the test statistic $\hat{\delta}$. Specifically, we selected four sampling intervals, $\Delta t_1 = 60$, $\Delta t_2 = 300$, $\Delta t_3 = 600$, and $\Delta t_4 = 1800$ seconds. We set the entire time span at 100 trading days, divided into 25 subperiods of 4 trading days each. Other model parameter values included $d_\eta = d_{\eta_1} = d_{\eta_2} = -0.25, -0.75$, and $d_{\tau_1} = d_{\tau_2} = 0.45$. Results are based on 1000 realizations.

We generated six test-statistic distributions for each pair of sampling intervals; for example, for the pair Δt_1, Δt_2, we obtained the test statistic

$$\hat{\delta}_{12,m} = \frac{\text{sample mean of } \{\hat{\theta}_{k,m}^{\Delta t_2} - \hat{\theta}_{k,m}^{\Delta t_1}\}}{\sqrt{\frac{1}{25} \cdot \text{sample variance of } \{\hat{\theta}_{k,m}^{\Delta t_2} - \hat{\theta}_{k,m}^{\Delta t_1}\}}}$$

for realization m based on $\{\hat{\theta}_{k,m}^{\Delta t_1}\}_{k=1}^{25}$ and $\{\hat{\theta}_{k,m}^{\Delta t_2}\}_{k=1}^{25}$. Overall, we had $\{\hat{\delta}_{12,m}\}_{m=1}^{1000}$, forming the simulation-based empirical distribution of the test statistic $\hat{\delta}_{12}$. This distribution can be used to generate critical values or compute empirical p-values. Table 2 summarizes the quantiles of these empirical distributions, where Q_q represents the qth quantile. For each distribution, the null hypothesis of normality is rejected at a nominal size of 1% based on the Kolmogorov–Smirnov goodness-of-fit test.

11. DATA ANALYSIS

In this section we focus on analyzing the buy prices, $\{P_{1,t}\}$, and sell prices, $\{P_{2,t}\}$, of a single stock, Tiffany Company

Table 2. Summary statistics of the simulation-based empirical distributions

Case	Test-stat	$Q_{0.005}$	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	$Q_{0.975}$	$Q_{0.995}$	Skewness	Excess kurtosis
$d_\eta = -0.25$	$\hat{\delta}_{12}$	−2.32	−1.80	−1.55	−0.04	1.47	1.64	2.12	0.05	−0.50
	$\hat{\delta}_{13}$	−2.26	−1.83	−1.54	0.02	1.48	1.70	1.96	0.11	−0.52
	$\hat{\delta}_{14}$	−2.12	−1.80	−1.51	0.00	1.39	1.66	2.16	0.02	−0.45
	$\hat{\delta}_{23}$	−2.43	−1.82	−1.60	0.03	1.57	1.80	2.25	0.08	−0.54
	$\hat{\delta}_{24}$	−2.09	−1.80	−1.50	0.03	1.39	1.65	2.22	0.03	−0.51
	$\hat{\delta}_{34}$	−2.17	−1.81	−1.54	−0.02	1.41	1.74	2.38	0.04	−0.48
$d_\eta = -0.75$	$\hat{\delta}_{12}$	−2.05	−1.66	−1.41	0.00	1.42	1.66	2.10	0.01	−0.72
	$\hat{\delta}_{13}$	−2.45	−1.69	−1.55	−0.02	1.41	1.59	2.02	0.10	−0.51
	$\hat{\delta}_{14}$	−2.14	−1.67	−1.34	−0.05	1.36	1.64	2.06	0.01	−0.17
	$\hat{\delta}_{23}$	−2.47	−1.77	−1.59	−0.02	1.40	1.62	2.17	0.07	−0.63
	$\hat{\delta}_{24}$	−2.15	−1.64	−1.37	−0.03	1.39	1.72	2.06	0.01	−0.24
	$\hat{\delta}_{34}$	−2.17	−1.73	−1.44	−0.01	1.41	1.66	2.11	0.02	−0.35
Standard normal	Z	−2.58	−1.96	−1.64	0.00	1.64	1.96	2.58	0.00	0.00

Table 3. Buy and sell prices of TIF

Δt (sec)	n	Estimate of θ	$\hat{d}_{\text{buy-price}}$ [SE]	$\hat{d}_{\text{sell-price}}$ [SE]	$\hat{d}_{\text{coint-error}}$ [SE]
1800	1612	$\hat{\theta} = 0.998040$	1.0124 [0.0484]	1.0105 [0.0484]	0.2328 [0.0484]
600	4,836	$\hat{\theta} = 0.998046$	1.0223 [0.0330]	1.0208 [0.0330]	0.1312 [0.0330]
300	9,672	$\hat{\theta} = 0.998042$	0.9906 [0.0259]	0.9914 [0.0259]	0.1068 [0.0259]
–	–	$\tilde{\theta} = 1.001678$	–	–	–

(ticker: TIF). The data were obtained from the TAQ database of WRDS. We considered daily transactions between 9:30 a.m. and 4:00 p.m. We ignore overnight durations and returns, as was also done by, for example, Hasbrouck (1995). The data span the period from January 25, 2000 to July 20, 2000, comprising 124 trading days.

We followed Lee and Ready (1991) in classifying individual trades. If the transaction price was higher than the prior bid-ask midpoint, then the current trade was labeled a buy order; if the transaction price was lower, then it was labeled a sell order. If the transaction price was exactly the same as the prior bid-ask midpoint, then we used the tick test (described in Lee and Ready 1991) to determine whether it should be classified as a buy order or a sell order. Lee and Ready (1991) found that the accuracy of their method was at least 85%. Using this method, we found 26,103 buy trades and 32,812 sell trades during the study period.

We first verified that a strong cointegrating relationship exists between buy and sell prices of TIF. The results are given in Table 3. We estimated the memory parameters of the log-buy prices and log-sell prices as 1 plus the GPH estimator (see Geweke and Porter-Hudak 1983) of the memory parameter of the differences. We estimated the memory parameter of the cointegrating error using a GPH estimator based on the levels of the residuals from an OLS regression of $\{\log P_{1,j}\}$ on $\{\log P_{2,j}\}$ for various choices of Δt. The memory parameter of the cointegrating error was $1 + \max(d_{\eta_1}, d_{\eta_2})$. The number of frequencies used in the log periodogram regressions was $n^{0.5}$. As expected, the estimated cointegrating parameter was close to 1. Evidence of strong cointegration was observed, along with some evidence indicating that the cointegration is fractional, not standard.

Next, using the ad hoc estimator $\tilde{\Theta}$, we estimated the model parameters for three clock-time subperiods, as well as the entire period. During period 1 (day 1 to day 25), the stock price declined by roughly 25%. During period 2 (day 41 to day 70), the price remained relatively stable. In period 3 (day 90 to day 124), the stock price increased by approximately 25%. The results are given in Table 4. The tick-time stock prices are plotted in Figure 2.

Based on the results in Table 4, we report the following findings:

(1) The microstructure noise variance estimates, $\tilde{\sigma}^2_{i,\eta}$, were smaller for period 2 (during which the stock prices varied substantially but showed no clear trend) than for periods 1 and 3 (during which the price showed a decreasing trend and an increasing trend, respectively).
(2) The value-shock variance estimates ($\tilde{\sigma}^2_{i,e}$) showed an opposite pattern, that is, larger in period 2 but smaller in periods 1 and 3.
(3) Comparing $\tilde{\sigma}^2_{i,e}$ and $\tilde{\sigma}^2_{i,\eta}$ shows that the variability of the value shocks usually exceeded that of the microstructure shocks. Indeed, $\tilde{\sigma}^2_{i,e}$ was greater than $\tilde{\sigma}^2_{i,\eta}$ for both buy and sell trades in the entire period.

As for the microstructure noise feedback coefficient estimates, $\tilde{g}_{21}$ and $\tilde{g}_{12}$, their magnitudes were generally around 1, but the signs varied in different periods. In some periods, the estimates of $\tilde{g}^2_{21}$ or $\tilde{g}^2_{12}$ were negative; thus we set the corresponding $\tilde{g}_{21}$ or $\tilde{g}_{12}$ at 0. In general, we found no systematic pattern for $\tilde{g}_{21}$ and $\tilde{g}_{12}$, and their values are not reported in Table 4. We stress, however, that the simulation study of Hurvich and Wang (2009) showed that g_{21} and g_{12} were harder to estimate than the other parameters.

Finding (1) is consistent with results from the study of Amihud and Mendelson (1980, 1982), where a market-maker executes buy and sell orders that arrive randomly, with the arrival rate determined by the quoted bid and ask prices, so as to maximize his expected profit per unit time, under the constraint that his inventory position will not exceed a long position (L) and a short position (S). (The analysis applies to traders who act as market makers, that is, quote buying and selling prices and benefit from trading at these prices, rather taking a long-run

Table 4. Method-of-moments parameter estimates of TIF

Period	Type	# of trades	$\tilde{\sigma}^2_{i,e}$ ($\times 10^{-6}$)	$\tilde{\sigma}^2_{i,\eta}$ ($\times 10^{-6}$)
1: Trading day 1 to 25	Buy	5,852	3.01	3.38
	Sell	6,875	3.05	1.93
2: Trading day 41 to 70	Buy	5,360	6.22	0.72
	Sell	7,688	4.08	0.83
3: Trading day 90 to 124	Buy	6,896	3.50	1.18
	Sell	8,827	2.00	2.66
Entire period	Buy	26,103	4.67	1.26
	Sell	32,812	3.35	1.66

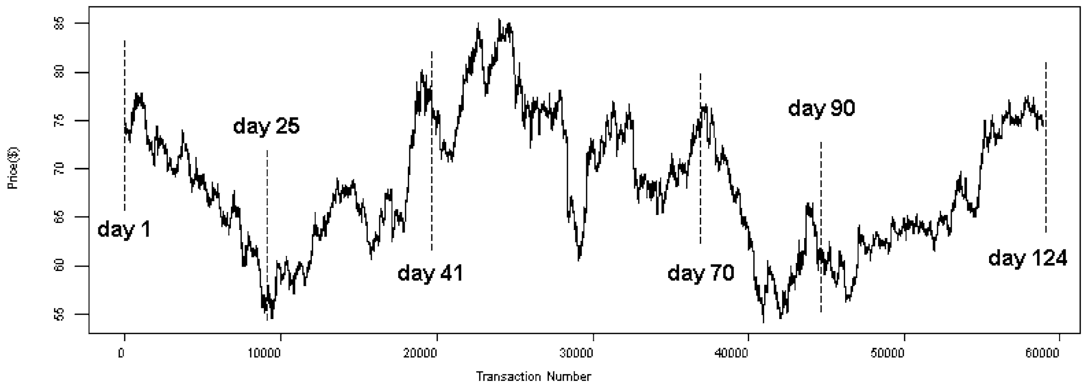

Figure 2. TIF Transaction-level stock price.

position in the stock, based on some information.) The market-maker sets the pair of bid-ask prices to adjust his inventory, and his policy results in having a preferred inventory position to which he reverts. The bid-ask spread is minimized at this preferred position, while it increases as the inventory diverges from the preferred level. This policy applies when there is no change in information about the security's value, in which case prices show no clear trend, hovering within a range. Amihud and Mendelson (1982, pp. 56–58) analyzed a situation involving a change in information about the security's value, unknown to the market-maker. At first, the market-maker maintains the schedule of bid-ask prices that applies to the old valuation, but given the value change, his inventory will deviate from the preferred position, and the bid-ask spread will widen. (For example, if the value is lower, then the market-maker will accumulate a large long position, quoted prices will decline, and the bid-ask spread will widen.) After realizing that the value has changed, the market-maker shifts his schedule of quoted bid-ask prices, and the bid-ask spread reverts to a normal, narrower range. Applying this analysis to the data, periods 1 and 3 demonstrate a major shift in the security value, reflected in the trend in price. Following Amihud and Mendelson (1982), a period of shifting value is associated with wider bid-ask spread. In period 2, when prices vary but do not exhibit a clear trend, the bid-ask spread should be narrower.

A narrower bid-ask spread (i.e., a smaller spread magnitude) indicates a smaller microstructure noise variance (see Amihud and Mendelson 1987, pp. 536, 547), that is, a smaller $\sigma^2_{\eta,i}$ in model (1). Indeed, we found that the estimated microstructure noise variances were smaller when price fluctuated without a clear trend (period 2) and larger otherwise (periods 1 and 3). Unfortunately, it is not possible to test the significance of the change in microstructure noise variances across the three subperiods, because the estimates are not independent.

Another interesting topic is the price discovery process, a popular topic in finance. Specifically, here we focus on the price discovery of a single stock, e.g., TIF, during different market environments. To estimate the information share, estimates for the trading intensities, λ_1 and λ_2, and the value-shock variances, $\sigma^2_{1,e}$ and $\sigma^2_{2,e}$, are required. To estimate λ_i $(i = 1, 2)$, we used the total number of transactions divided by the total period of observation for asset i. We estimated $\sigma^2_{1,e}$ and $\sigma^2_{2,e}$ by the method of moments, as discussed in Section 8. We computed the information share estimates for each of three clock-time periods based on the results given in Table 4. The results are summarized in Table 5.

For period 2, the information shares were approximately equally divided between buys and sells. For period 1, when the stock price declined dramatically, the sell trades had more information than buy trades. In contrast, during period 3, when prices were rising, the buy trades had more information. Unfortunately, we could not test the significance of the change in information share across the three periods, because the estimates were not independent.

As pointed out by Hasbrouck (1995), the information ratios are not related to the microstructure (e.g., spreads) of the markets. This is clear because only the random-walk components of the price series are used in the construction of the information ratios. Thus the results that we have presented so far in this section reflect different aspects of the dynamics of the TIF price series.

Finally, we implemented the specification test described in Section 9 to investigate whether model (1) was misspecified. We divided the 124-day trading period into $K = 62$ subperiods of 2 trading days each. We chose four sampling intervals, $\Delta t_1 = 1800$, $\Delta t_2 = 600$, $\Delta t_3 = 300$, and $\Delta t_4 = 60$ seconds.

First, we simulated the corresponding empirical distributions of $\hat{\delta}$'s, as in Section 10. The results are reported in Table 6.

Based on the values in Table 6, we used the corresponding simulated test statistic distributions to compute the empirical p-values reported in Table 7.

For two-sided hypothesis testing with nominal size of 5%, we found no significant evidence to indicate that model (1) was misspecified for the TIF data set, because the null was not rejected in any of the six cases.

Table 5. Information shares (S) of buy and sell price of TIF

Period	$\tilde{S}_{\text{buy}}$	$\tilde{S}_{\text{sell}}$	$(\tilde{S}_{\text{buy}} - \tilde{S}_{\text{sell}})$
1: Trading day 1 to 25	45.7%	54.3%	−8.6%
2: Trading day 41 to 70	51.5%	48.5%	3.0%
3: Trading day 90 to 124	57.5%	42.3%	15.2%
Entire period	52.6%	47.4%	5.2%

Table 6. Summary statistics of the simulation-based empirical distributions for Tiffany (TIF)

Case	Test-stat	$Q_{0.005}$	$Q_{0.025}$	$Q_{0.05}$	$Q_{0.5}$	$Q_{0.95}$	$Q_{0.975}$	$Q_{0.995}$	Skewness	Kurtosis
Tiffany (TIF)	$\hat{\delta}_{12}$	−2.05	−1.68	−1.41	−0.04	1.39	1.69	2.02	0.03	2.45
	$\hat{\delta}_{13}$	−1.98	−1.61	−1.41	−0.03	1.37	1.53	2.04	0.02	2.32
	$\hat{\delta}_{14}$	−2.07	−1.64	−1.43	−0.09	1.28	1.55	2.11	0.08	2.46
	$\hat{\delta}_{23}$	−2.00	−1.70	−1.47	0.00	1.45	1.73	2.27	0.07	2.38
	$\hat{\delta}_{24}$	−1.94	−1.61	−1.44	−0.10	1.34	1.65	2.12	0.13	2.41
	$\hat{\delta}_{34}$	−1.96	−1.68	−1.46	−0.12	1.38	1.69	2.08	0.13	2.39

12. CONCLUSIONS

Remark 1. So far, we have seen that the model (1) yields cointegration and also captures two stylized facts that have been observed in actual data: volatility clustering and persistence in duration. It is worth mentioning that by modifying the basic model (1) properly, two additional key stylized facts can be captured: the leverage effect (see, e.g., Andersen et al. 2006), and portfolio autocorrelation due to nonsynchronous trading (see, e.g., Fisher 1966; Scholes and Williams 1977; Lo and MacKinlay 1990a, 1990b; Boudoukh, Richardson, and Whitelaw 1994; Kadlec and Patterson 1999). Hurvich and Wang (2009) have verified that modified versions of our model can indeed capture these effects.

Remark 2. There is an important caveat regarding the martingale property in the special case of model (1), in which the microstructure noise components $\{\eta_{1,k}\}$ and $\{\eta_{2,k}\}$ are absent. For each series, as long as the conditioning set involves only returns of the given series up to time t, the log-price series (observed at discrete, equally spaced time intervals) is a martingale. But the martingale property is lost if the conditioning set is augmented to include returns on both assets up to time t. Because of the feedback effect in the model and the nonsynchronous trading, recent information about asset 1 can help predict the asset 2 return, even though the asset 2 return is unpredictable based on its own past. Such a situation can occur in actual markets. For example, to predict the (real) return on the sale of a given home, it helps to know the returns on sales of similar homes that have occurred recently, although knowing the past returns on sales of the given home may not help at all, especially if the home has not been sold for a long time.

We note a few possibilities for future work stemming from the current study. It might be interesting to investigate the interplay between cointegration and option pricing, hedging, asset allocation, pairs trading, and index tracking in the current pure-jump context. So far, work has been done on option pricing based on pure-jump processes (Prigent 2001) and dynamic asset allocation based on jump-diffusion processes (Liu, Longstaff, and Pan 2003), but this work does not allow for cointegration. Another stream of the literature indicates that in a diffusion context, cointegration may have an impact on option pricing (Duan and Pliska 2004) and on index tracking (Alexander and Dimitriu 2005; Dunis and Ho 2005), but this work does not allow for a pure-jump process.

Other estimators of the cointegrating parameter besides OLS could be considered. Although many such estimators have been proposed for both standard and fractional cointegration, none has yet been justified under a transaction-level model such as (1). Semiparametric estimators could be considered, because, by the previous remark, our results do not require a parametric model for durations.

A possible generalization of model (1) to the case of $W \geq 2$ price series $P_{1,t}, \ldots, P_{W,t}$ is

$$\log P_{1,t} = \sum_{k=1}^{N_1(t)} (e_{1,k} + \eta_{1,k}) + \sum_{i=2}^{W} \left\{ \sum_{k=1}^{N_i(t_{1,N_1(t)})} (\theta_{i1} e_{i,k} + g_{i1} \eta_{i,k}) \right\},$$

$$\log P_{2,t} = \sum_{k=1}^{N_2(t)} (e_{2,k} + \eta_{2,k}) + \sum_{i=1,i\neq 2}^{W} \left\{ \sum_{k=1}^{N_i(t_{2,N_2(t)})} (\theta_{i2} e_{i,k} + g_{i2} \eta_{i,k}) \right\}, \quad (10)$$

$$\vdots$$

$$\log P_{W,t} = \sum_{k=1}^{N_W(t)} (e_{W,k} + \eta_{W,k}) + \sum_{i=1}^{W-1} \left\{ \sum_{k=1}^{N_i(t_{W,N_W(t)})} (\theta_{iW} e_{i,k} + g_{iW} \eta_{i,k}) \right\},$$

where for $i = 1, \ldots, W$, the $\{e_{i,k}\}$ are mutually independent mean-0 iid value shock series; the $\{\eta_{i,k}\}$ are mutually independent microstructure shock series satisfying Condition A, B, or C with memory parameters $d_{\eta_i} \in [-1, 0)$; and for $i \neq j$, the parameters θ_{ij} and η_{ij} represent the impact of the value and microstructure shocks from series i on series j.

In the bivariate case, $W = 2$, there are two feedback coefficients for the value shocks, θ_{21} and θ_{12}. When cointegration exists, one coefficient is constrained to be the reciprocal of the

Table 7. Specification test for TIF

Δt pair under testing	Value of the test statistic [empirical p-value]
Δt_1 vs. Δt_2	1.2356 [0.144]
Δt_1 vs. Δt_3	0.0834 [0.920]
Δt_1 vs. Δt_4	−0.0190 [0.936]
Δt_2 vs. Δt_3	−0.7655 [0.474]
Δt_2 vs. Δt_4	−0.3412 [0.806]
Δt_3 vs. Δt_4	−0.0754 [0.958]

other, as in model (1), where $\theta_{21} = \theta$ and $\theta_{12} = 1/\theta$. In the multivariate model (10), there are $W(W-1)$ such feedback coefficients and at most $(W-1)$ cointegrating vectors, although we do not present here the constraints on the coefficients θ_{ij} that would imply a specific cointegrating rank. It also would be of interest to derive a common-components representation for (10), as was obtained for clock-time multivariate models under standard cointegration by Stock and Watson (1988). Such a representation would generalize the representation (2) to the multivariate case, and presumably would facilitate inference on the cointegrating rank (as it did in Stock and Watson 1988). Finally, it would be of interest to derive properties for the OLS and other estimators of the cointegrating vectors in (10), as considered, for example, for OLS in clock-time multivariate models under standard cointegration by Stock (1987).

APPENDIX: PROOFS

A.1 Lemmas

These lemmas, used in the proofs of our theorems and of interest in their own right, were proved in earlier work (Hurvich and Wang 2009). From (2), it can be seen that the microstructure components of the log price are random sums of the microstructure noise. Lemmas 1 and 3 show, for the case of weak and strong fractional cointegration, respectively, that such random sums have memory parameter $1 + d_\eta < 1$, where d_η is the memory parameter of the microstructure noise.

Lemma 1. Suppose that $\{\eta_k\}$ has memory parameter $d_\eta \in (-\frac{1}{2}, 0)$, is independent of $\{\tau_k\}$, and satisfies Condition A. Then

$$\operatorname{var}\left(\sum_{k=1}^{N(t)} \eta_k\right) \sim (\tilde{\sigma}^2 \lambda^{2d_\eta+1}) t^{2d_\eta+1}$$

as $t \to \infty$.

The following lemma is used to prove Lemma 1.

Lemma 2. For $d_\eta \in (-\frac{1}{2}, 0)$, suppose that $\{\eta_k\}$ satisfies Condition A. Then there exists a positive constant C such that for all nonnegative integers s, $\operatorname{var}(\sum_{k=1}^{s} \eta_k) = \tilde{\sigma}^2 s^{2d_\eta+1} + R(s)$, where $|R(s)| \le C s^{\max(2d_\eta+1-\beta, 0)}$.

Lemma 3. For $d_\eta \in (-1, -\frac{1}{2})$, suppose that $\{\eta_k\}$ satisfies Condition B and is independent of $N(\cdot)$. Then, for any fixed $t > 0$,

$$\operatorname{cov}\left(\sum_{k=1}^{N(t)} \eta_k, \sum_{k=1}^{N(t+j)} \eta_k\right) \sim C j^{2d_\eta+1} \Pr\{N(t) > 0\} \tag{A.1}$$

as $j \to \infty$, where $C > 0$ is a constant not depending on t.

The following two lemmas are used in the proofs of Theorems 3, 4, and 5.

Lemma 4. If the durations $\{\tau_k\}$ are generated by a LMSD model with memory parameter $d_\tau \in (0, \frac{1}{2})$ and all moments of the durations $\{\tau_k\}$ are finite, then all moments of the backward recurrence time (BRT_t), as defined in (A.2), also are finite.

Lemma 5. For durations $\{\tau_k\}$ satisfying the assumptions in Lemma 4, $E[N(s)^m] \le K_m(s^m + 1)$ for all $s > 0$, where $K_m < \infty$, $m = 1, 2, \ldots$.

A.2 Proof of Theorem 1

We first consider the fractional cointegration case, $d_\eta \in (-\frac{1}{2}, 0)$. We focus on $\log P_{1,t}$; the proof for $\log P_{2,t}$ follows along similar lines.

The log price of asset 1 is

$$\log P_{1,t} = \sum_{k=1}^{N_1(t)} (e_{1,k} + \eta_{1,k}) + \sum_{k=1}^{N_2(t_{1,N_1(t)})} (\theta e_{2,k} + g_{21}\eta_{2,k}).$$

Note that the two terms on the right side are uncorrelated. By Lemma 1, because $d_\eta < 0$, we obtain

$$\begin{aligned}\operatorname{var}\left[\sum_{k=1}^{N_1(t)} (e_{1,k} + \eta_{1,k})\right] &= \sigma_{1,e}^2 E[N_1(t)] + \operatorname{var}\left[\sum_{k=1}^{N_1(t)} \eta_{1,k}\right] \\ &\sim (\sigma_{1,e}^2 \lambda_1) t + (\sigma_{1,\eta}^2 \lambda_1^{2d_{\eta_1}+1}) t^{2d_{\eta_1}+1} \\ &= (\sigma_{1,e}^2 \lambda_1) t + o(t).\end{aligned}$$

Next, consider $E\{N_1(t) - N_1(t_{2,N_2(t)})\}$, which is the expected number of transactions of asset 1 after the most recent transaction of asset 2 up to time t. Define the *backward recurrence time* for asset 2 at time t as

$$BRT_{2,t} = \inf\{s > 0 : N_2(t) - N_2(t-s) > 0\}. \tag{A.2}$$

Clearly, $BRT_{2,t} = t - t_{2,N_2(t)}$. By stationarity of $N_2(\cdot)$, and using (3.1.7) of Daley and Vere-Jones (2003, p. 43), we obtain $E\{N_1(t) - N_1(t_{2,N_2(t)})\} = E[-N_1(-BRT_{2,t})] = E[-N_1(-BRT_{2,0})]$. In the right-side equality, we used the fact that because $N_2(\cdot)$ is a stationary point process, $BRT_{2,t}$ has the same distribution as $BRT_{2,0}$, which does not depend on t. (See Daley and Vere-Jones 2003, pp. 58–59, for a detailed discussion.) Thus

$$E\{N_1(t) - N_1(t_{2,N_2(t)})\} = \tilde{C}_1, \tag{A.3}$$

a finite constant, independent of t. Similarly,

$$E\{N_2(t) - N_2(t_{1,N_1(t)})\} = \tilde{C}_2 \tag{A.4}$$

also is a finite constant, independent of t as well.

It follows from the proof of Lemma 1 that

$$\begin{aligned}&\operatorname{var}\left[\sum_{k=1}^{N_2(t_{1,N_1(t)})} (\theta e_{2,k} + g_{21}\eta_{2,k})\right] \\ &= \theta^2 \sigma_{2,e}^2 \underbrace{E\{N_2(t_{1,N_1(t)})\}}_{T_1} + g_{21}^2 \sigma_{2,\eta}^2 \underbrace{E\{[N_2(t_{1,N_1(t)})]^{2d_\eta+1}\}}_{T_2} \\ &\quad + g_{21}^2 \sigma_{2,\eta}^2 \underbrace{E\{R(N_2(t_{1,N_1(t)}))\}}_{T_3}.\end{aligned}$$

By (A.4), the first term equals

$$T_1 = E\{N_2(t)\} - \tilde{C}_2 = \lambda_2 t - \tilde{C}_2 \sim \lambda_2 t,$$

as $t \to \infty$.

As for the second term, because when $x > 0$ and $0 < p = (2d_\eta + 1) < 1$, the function x^p is concave, we can apply Jensen's inequality to obtain

$$T_2 \le \{E[N_2(t_{1,N_1(t)})]\}^{2d_\eta+1} = (\lambda_2 t - \tilde{C}_2)^{2d_\eta+1} = o(t).$$

It follows from the proof of Lemma 1 that

$$|T_3| \le CE\{[N_2(t_{1,N_1(t)})]^{\max(2d_\eta+1-\beta,0)}\} = o(t).$$

Therefore,

$$\mathrm{var}\left[\sum_{k=1}^{N_2(t_{1,N_1(t)})}(\theta e_{2,k} + g_{21}\eta_{2,k})\right] \sim (\theta^2\sigma_{2,e}^2\lambda_2)t$$

as $t \to \infty$.

Overall,

$$\mathrm{var}[\log P_{1,t}] \sim (\sigma_{1,e}^2\lambda_1)t + (\theta^2\sigma_{2,e}^2\lambda_2)t = C_1 t,$$

where $C_1 = (\sigma_{1,e}^2\lambda_1 + \theta^2\sigma_{2,e}^2\lambda_2)$.

Similarly,

$$\mathrm{var}[\log P_{2,t}] \sim (\sigma_{2,e}^2\lambda_2)t + \left(\frac{1}{\theta^2}\sigma_{1,e}^2\lambda_1\right)t = C_2 t,$$

where $C_2 = (\sigma_{2,e}^2\lambda_2 + \frac{1}{\theta^2}\sigma_{1,e}^2\lambda_1)$.

Next, for both the strong fractional cointegration case [$d_\eta \in (-1, -\frac{1}{2})$] and the standard cointegration case ($d_\eta = -1$), the proof is identical to that for the weak fractional cointegration case, except that here we have $\mathrm{var}(\sum_{k=1}^{N_i(t)}\eta_{i,k})$ $(i = 1, 2)$, equal to some finite constants, which do not increase with t.

A.3 Proof of Theorem 2

We first consider the fractional cointegration case, $d_\eta \in (-\frac{1}{2}, 0)$. We focus on the returns $\{r_{1,j}\}$ of asset 1, which corresponds to the first equation in (6), because the proof for $\{r_{2,j}\}$ follows along similar lines.

Consider the lag-1 autocorrelation of

$$r_{1,j} = \underbrace{\sum_{k=N_1[(j-1)\Delta t]+1}^{N_1[j\Delta t]} e_{1,k}}_{T_1} + \underbrace{\sum_{k=N_1[(j-1)\Delta t]+1}^{N_1[j\Delta t]} \eta_{1,k}}_{T_2} + \underbrace{\sum_{k=N_2(t_{1,N_1((j-1)\Delta t)})+1}^{N_2(t_{1,N_1(j\Delta t)})} \theta e_{2,k}}_{T_3} + \underbrace{\sum_{k=N_2(t_{1,N_1((j-1)\Delta t)})+1}^{N_2(t_{1,N_1(j\Delta t)})} g_{21}\eta_{2,k}}_{T_4}.$$

Denote $\Delta N_{1,j} = N_1(j\Delta t) - N_1((j-1)\Delta t)$ and $\Delta N_{2,j} = N_2(j\Delta t) - N_2((j-1)\Delta t)$. We know that $E(\Delta N_{1,j}) = \lambda_1\Delta t$ and $E(\Delta N_{2,j}) = \lambda_2\Delta t$. Thus

$$\begin{aligned}\mathrm{var}(T_1) &= E\left\{\left[\sum_{k=N_1((j-1)\Delta t)+1}^{N_1(j\Delta t)} e_{1,k}\right]^2\right\} \\ &= E\left[E\left\{\left[\sum_{k=N_1((j-1)\Delta t)+1}^{N_1(j\Delta t)} e_{1,k}\right]^2 \middle| N_1(\cdot)\right\}\right] \\ &= \sigma_{1,e}^2 E\{N_1(j\Delta t) - N_1((j-1)\Delta t)\} \\ &= \sigma_{1,e}^2 E(\Delta N_{1,j}) = \sigma_{1,e}^2\lambda_1\Delta t. \qquad \text{(A.5)}\end{aligned}$$

By Lemma 2, we have

$$\mathrm{var}(T_2) = \sigma_{2,\eta}^2 E[\{\Delta N_{1,j}\}^{2d_\eta+1}] + \sigma_{2,\eta}^2 E\{R(\Delta N_{1,j})\}.$$

Because the function x^p is concave when $x > 0$ and $0 < p < 1$, by Jensen's inequality for $d_\eta \in (-0.5, 0)$, the first part satisfies

$$\begin{aligned}\sigma_{2,\eta}^2 E[\{\Delta N_{1,j}\}^{2d_\eta+1}] &\le \sigma_{2,\eta}^2\{E[\Delta N_{1,j}]\}^{2d_\eta+1} \\ &= \sigma_{2,\eta}^2\{\lambda_1\Delta t\}^{2d_\eta+1} = o(\Delta t), \qquad \text{(A.6)}\end{aligned}$$

as $\Delta t \to \infty$. As for the second part,

$$\begin{aligned}|E\{R(\Delta N_{1,j})\}| &\le CE\{(\Delta N_{1,j})^{\max(2d_\eta+1-\beta,0)}\} \\ &\le CE\{(\Delta N_{1,j})^{2d_\eta+1}\} = o(\Delta t).\end{aligned}$$

Thus $\mathrm{var}(T_2) = o(\Delta t)$ as $\Delta t \to \infty$.

Hurvich and Wang (2009) used Lemma 2 and equations (A.3) and (A.4) to show that $\mathrm{var}(T_3) = \theta^2\sigma_{2,e}^2\lambda_2\Delta t$ and $\mathrm{var}(T_4) = o(\Delta t)$, which, together with the Cauchy–Schwartz inequality and eqs. (A.5) and (A.6), imply that $\mathrm{cov}(T_i, T_j) = o(\Delta t)$ for $i, j = 1, \ldots, 4$ with $i \ne j$. In addition, $\mathrm{cov}(T_1, T_3) = 0$, because $\{e_{1,k}\}$ and $\{e_{2,k}\}$ are mutually independent iid series.

Overall, we obtain $\mathrm{var}(r_{1,j}) \sim (\sigma_{1,e}^2\lambda_1 + \theta^2\sigma_{2,e}^2\lambda_2)\Delta t$, as $\Delta t \to \infty$, that is,

$$\lim_{\Delta t\to\infty}\frac{\mathrm{var}(r_{1,j})}{\Delta t} = (\sigma_{1,e}^2\lambda_1 + \theta^2\sigma_{2,e}^2\lambda_2).$$

Similarly, for

$$(r_{1,j} + r_{1,j+1}) = \sum_{k=N_1((j-1)\Delta t)+1}^{N_1((j+1)\Delta t)}(e_{1,k} + \eta_{1,k}) + \sum_{k=N_2(t_{1,N_1((j-1)\Delta t)})+1}^{N_2(t_{1,N_1((j+1)\Delta t)})}(\theta e_{2,k} + g_{21}\eta_{2,k}),$$

we obtain

$$\mathrm{var}(r_{1,j} + r_{1,j+1}) \sim 2(\sigma_{1,e}^2\lambda_1 + \theta^2\sigma_{2,e}^2\lambda_2)\Delta t,$$

that is,

$$\lim_{\Delta t\to\infty}\frac{\mathrm{var}(r_{1,j} + r_{1,j+1})}{2\Delta t} = (\sigma_{1,e}^2\lambda_1 + \theta^2\sigma_{2,e}^2\lambda_2).$$

Therefore,

$$\begin{aligned}\mathrm{corr}(r_{1,j}, r_{1,j+1}) &= \frac{\mathrm{cov}(r_{1,j}, r_{1,j+1})}{\mathrm{var}(r_{1,j})} \\ &= \frac{(1/2)\,\mathrm{var}(r_{1,j} + r_{1,j+1}) - \mathrm{var}(r_{1,j})}{\mathrm{var}(r_{1,j})} \\ &= \frac{(1/2)\,\mathrm{var}(r_{1,j} + r_{1,j+1})}{\mathrm{var}(r_{1,j})} - 1 \\ &= \frac{\mathrm{var}(r_{1,j} + r_{1,j+1})/(2\Delta t)}{\mathrm{var}(r_{1,j})/\Delta t} - 1 \to 0,\end{aligned}$$

as $\Delta t \to \infty$.

The fact that the lag-2 autocorrelation also converges to 0 can be shown by recognizing that

$$\mathrm{corr}(r_{1,j}, r_{1,j+2}) = \frac{1}{2}\left[\frac{\mathrm{var}(r_{1,j} + r_{1,j+1} + r_{1,j+2})}{\mathrm{var}(r_{1,j})} - 3 - 4\,\mathrm{corr}(r_{1,j}, r_{1,j+1})\right]$$

and using the lag-1 autocorrelation results proved earlier, along with

$$\lim_{\Delta t\to\infty}\frac{\operatorname{var}(r_{1,j}+r_{1,j+1}+r_{1,j+2})}{3\Delta t}=(\sigma_{1,e}^2\lambda_1+\theta^2\sigma_{2,e}^2\lambda_2).$$

The result follows for any fixed lag k by induction.

Next, for both the strong fractional cointegration case [$d_\eta \in (-1,-\frac{1}{2})$] and the standard cointegration case ($d_\eta=-1$), the proof is identical to that for the weak fractional cointegration case, except that here we have $\operatorname{var}(\sum_{k=N_i((j-1)\Delta t)+1}^{N_i(j\Delta t)}\eta_{i,k})$ ($i=1,2$) (as well as other similar terms) equal to some finite constants, which do not increase with Δt.

A.4 Proof of Theorem 3

Consider a linear combination of $\log P_{1,t}$ and $\log P_{2,t}$ using vector $(1,-\theta)$,

$$\begin{aligned}\log P_{1,t}-\theta\log P_{2,t}&\\ =\underbrace{\sum_{k=N_1(t_{2,N_2(t)})+1}^{N_1(t)} e_{1,k}}_{T_1}-\theta\underbrace{\sum_{k=N_2(t_{1,N_1(t)})+1}^{N_2(t)} e_{2,k}}_{T_2}&\\ +(1-\theta g_{12})\underbrace{\sum_{k=1}^{N_1(t)}\eta_{1,k}}_{T_3}+\theta g_{12}\underbrace{\sum_{k=N_1(t_{2,N_2(t)})+1}^{N_1(t)}\eta_{1,k}}_{T_4}&\\ -(\theta-g_{21})\underbrace{\sum_{k=1}^{N_2(t)}\eta_{2,k}}_{T_5}-g_{21}\underbrace{\sum_{k=N_2(t_{1,N_1(t)})+1}^{N_2(t)}\eta_{2,k}}_{T_6}.&\qquad\text{(A.7)}\end{aligned}$$

Because all shock series are mutually independent and also independent of the counting processes $N_1(t)$ and $N_2(t)$, we obtain

$$\begin{aligned}\operatorname{var}[\log P_{1,t}-\theta\log P_{2,t}]&\\ =\operatorname{var}(T_1)+\theta^2\operatorname{var}(T_2)+(1-\theta g_{12})^2\operatorname{var}(T_3)&\\ +\theta^2 g_{12}^2\operatorname{var}(T_4)+2\theta g_{12}(1-\theta g_{12})\operatorname{cov}(T_3,T_4)&\\ +(\theta-g_{21})^2\operatorname{var}(T_5)&\\ +g_{21}^2\operatorname{var}(T_6)+2g_{21}(\theta-g_{21})\operatorname{cov}(T_5,T_6).&\qquad\text{(A.8)}\end{aligned}$$

First, by Lemma 1,

$$\begin{aligned}\operatorname{var}(T_3)&\sim(\sigma_{1,\eta}^2\lambda_1^{2d_\eta+1})t^{2d_\eta+1},\\ \operatorname{var}(T_5)&\sim(\sigma_{2,\eta}^2\lambda_2^{2d_\eta+1})t^{2d_\eta+1}.\end{aligned}\qquad\text{(A.9)}$$

Using (A.3) and Lemma 2, we obtain

$$\begin{aligned}\operatorname{var}(T_4)&=\sigma_{1,\eta}^2 E\big[\{N_1(t)-N_1(t_{2,N_2(t)})\}^{2d_\eta+1}\big]\\ &\quad+\sigma_{1,\eta}^2 E\{R[N_1(t)-N_1(t_{2,N_2(t)})]\}\\ &\le\sigma_{1,\eta}^2\big[E\{N_1(t)-N_1(t_{2,N_2(t)})\}\big]^{2d_\eta+1}\\ &\quad+\sigma_{1,\eta}^2 C\big[E\{N_1(t)-N_1(t_{2,N_2(t)})\}\big]^{2d_\eta+1}\\ &=(1+C)\sigma_{1,\eta}^2\tilde{C}_1^{2d_\eta+1},\end{aligned}\qquad\text{(A.10)}$$

where we apply Jensen's inequality in the last inequality, noting that for $x>0$ and $0<p=(2d_\eta+1)<1$, the function x^p is concave. Similarly,

$$\operatorname{var}(T_6)\le(1+C)\sigma_{2,\eta}^2\tilde{C}_2^{2d_\eta+1}.\qquad\text{(A.11)}$$

In addition, by (A.3) and (A.4),

$$\begin{aligned}\operatorname{var}(T_1)&=\operatorname{var}(e_{1,k})E\{N_1(t)-N_1(t_{2,N_2(t)})\}\\ &=\sigma_{1,e}^2\tilde{C}_1,\end{aligned}\qquad\text{(A.12)}$$

$$\begin{aligned}\operatorname{var}(T_2)&=\operatorname{var}(e_{2,k})E\{N_2(t)-N_2(t_{1,N_1(t)})\}\\ &=\sigma_{2,e}^2\tilde{C}_2.\end{aligned}\qquad\text{(A.13)}$$

Next, we consider the covariance terms in (A.8) using the Cauchy–Schwartz inequality. By (A.9) and (A.10),

$$\begin{aligned}|\operatorname{cov}(T_3,T_4)|&\le\sqrt{\operatorname{var}(T_3)\operatorname{var}(T_4)}\\ &\le\sqrt{(1+C)\sigma_{1,\eta}^2\tilde{C}_1^{2d_\eta+1}\operatorname{var}(T_3)}\\ &=o(t^{2d_\eta+1}),\end{aligned}\qquad\text{(A.14)}$$

and similarly, by (A.9) and (A.11),

$$\begin{aligned}|\operatorname{cov}(T_5,T_6)|&\le\sqrt{\operatorname{var}(T_5)\operatorname{var}(T_6)}\\ &\le\sqrt{(1+C)\sigma_{2,\eta}^2\tilde{C}_2^{2d_\eta+1}\operatorname{var}(T_5)}\\ &=o(t^{2d_\eta+1}).\end{aligned}\qquad\text{(A.15)}$$

Overall, using (A.9) to (A.15) for (A.8), we obtain

$$\operatorname{var}(\log P_{1,t}-\theta\log P_{2,t})\sim Ct^{2d_\eta+1},\qquad\text{(A.16)}$$

where $C=(1-\theta g_{12})^2(\sigma_{1,\eta}^2\lambda_1^{2d_{\eta_1}+1})+(\theta-g_{21})^2(\sigma_{2,\eta}^2\lambda_2^{2d_{\eta_2}+1})$.

The cointegrating vector is $(1,-\theta)$, and the memory parameter decreases from 1 for both log prices to $1+d_\eta$.

A.5 Proof of Theorem 4

The proof follows along the same lines as the proof of Theorem 3, except that we now use Lemma 3 to obtain the asymptotic behavior of the autocovariances of the partial sums of the microstructure noise.

A.6 Proof of Theorem 5

As in the proof of Theorem 3, we denote

$$\begin{aligned}S_t&=\log P_{1,t}-\theta\log P_{2,t}\\ &=\underbrace{\sum_{k=N_1(t_{2,N_2(t)})+1}^{N_1(t)} e_{1,k}}_{S_{1,t}}-\theta\underbrace{\sum_{k=N_2(t_{1,N_1(t)})+1}^{N_2(t)} e_{2,k}}_{S_{2,t}}\\ &\quad+\underbrace{\sum_{k=1}^{N_1(t)}\eta_{1,k}}_{S_{3,t}}-\theta g_{12}\underbrace{\sum_{k=1}^{N_1(t_{2,N_2(t)})}\eta_{1,k}}_{S_{4,t}}\end{aligned}$$

$$- \theta \underbrace{\sum_{k=1}^{N_2(t)} \eta_{2,k}}_{S_{5,t}} + g_{21} \underbrace{\sum_{k=1}^{N_2(t_{1,N_1(t)})} \eta_{2,k}}_{S_{6,t}}$$

$$= S_{1,t} - \theta S_{2,t} + S_{3,t} - \theta g_{12} S_{4,t} - \theta S_{5,t} + g_{12} S_{6,t},$$

and evaluate the terms in $\operatorname{cov}(S_t, S_{t+j})$.

(1) Consider $\operatorname{cov}(S_{1,t}, S_{1,t+j}) = E(S_{1,t}S_{1,t+j})$. The term $S_{1,t}$ is a sum of shocks occurring in the time interval between the last transaction of asset 2 before time t and time t. Similarly, $S_{1,t+j}$ is a sum of shocks occurring between the last transaction of asset 2 before time $t+j$ and time $t+j$. Clearly, if at least one transaction of asset 2 occurs in $(t, t+j]$, then we must have $t_{2,N_2(t+j)} > t$, so that $E[S_{1,t}S_{1,t+j}|N_1(\cdot), N_2(\cdot)] = 0$, because $\{e_{1,k}\}$ is iid. Otherwise, $t_{2,N_2(t+j)} = t_{2,N_2(t)}$ and $E[S_{1,t}S_{1,t+j}|N_1(\cdot), N_2(\cdot)] = \sigma_{1,e}^2[N_1(t) - N_1(t_{2,N_2(t)})]$. Therefore, by the Cauchy–Schwartz inequality,

$$\begin{aligned}
&\operatorname{cov}(S_{1,t}, S_{1,t+j}) \\
&\quad = E(S_{1,t}S_{1,t+j}) \\
&\quad = E\{E[S_{1,t}S_{1,t+j}|N_1(\cdot), N_2(\cdot)]\} \\
&\quad = E\{\sigma_{1,e}^2[N_1(t) - N_1(t_{2,N_2(t)})] \cdot I\{N_2(t+j) - N_2(t) = 0\}\} \\
&\quad \le \sigma_{1,e}^2\{E[N_1(t) - N_1(t_{2,N_2(t)})]^2\}^{1/2} \\
&\qquad \cdot \{P[N_2(t+j) - N_2(t) = 0]\}^{1/2}.
\end{aligned}$$

By Lemma 5 and the stationarity of $N_1(\cdot)$, we obtain

$$\begin{aligned}
&E\{[N_1(t) - N_1(t_{2,N_2(t)})]^2\} \\
&\quad = E\{[N_1(t - t_{2,N_2(t)})]^2\} \\
&\quad = E\{[N_1(BRT_{2,t})]^2\} \\
&\quad = E(E\{[N_1(BRT_{2,t})]^2|N_1(\cdot), N_2(\cdot)\}) \\
&\quad \le E[K_2(BRT_{2,t}^2 + 1)],
\end{aligned}$$

which is bounded uniformly in t using Lemma 4.

Next, because $N_2(\cdot)$ is stationary, for any positive integer m, we obtain

$$\begin{aligned}
P[N_2(t+j) - N_2(t) = 0] &= P[N_2(j) \le 0] \\
&\le P[|Z_2(j)| \ge \lambda_2 j^{1/2-d_\tau}] \\
&\le \frac{E|Z_2(j)|^m}{\lambda_2^m j^{m(1/2-d_\tau)}} \\
&= O(j^{m(d_\tau - 1/2)}), \qquad \text{(A.17)}
\end{aligned}$$

where $Z_2(j) = \frac{N_2(j) - \lambda_2 j}{j^{1/2+d_\tau}}$. This is true because it follows from the proof of proposition 1 of Deo et al. (2009) that $E|Z_2(j)|^m$ is bounded uniformly in j for all m. Therefore, $P[N_2(t+j) - N_2(t) = 0]$ has nearly exponential decay, because (A.17) holds for all m. Thus $\operatorname{cov}(S_{1,t}, S_{1,t+j})$ has nearly exponential decay. Similarly, $\operatorname{cov}(S_{2,t}, S_{2,t+j})$ has nearly exponential decay.

(2) In earlier work (Hurvich and Wang 2009) we used Lemma 4 and Lemma 5 to show that $\operatorname{cov}(S_{3,t}, S_{3,t+j})$, $\operatorname{cov}(S_{3,t}, S_{4,t+j})$, $\operatorname{cov}(S_{4,t}, S_{3,t+j})$, and $\operatorname{cov}(S_{4,t}, S_{4,t+j})$ all have nearly exponential decay.

(3) So far, we have shown that the following terms have nearly exponential decay as $j \to \infty$: $\operatorname{cov}(S_{1,t}, S_{1,t+j})$, $\operatorname{cov}(S_{2,t}, S_{2,t+j})$, $\operatorname{cov}(S_{3,t}, S_{3,t+j})$, $\operatorname{cov}(S_{3,t}, S_{4,t+j})$, $\operatorname{cov}(S_{4,t}, S_{3,t+j})$, $\operatorname{cov}(S_{4,t}, S_{4,t+j})$, $\operatorname{cov}(S_{5,t}, S_{5,t+j})$, $\operatorname{cov}(S_{5,t}, S_{6,t+j})$, $\operatorname{cov}(S_{6,t}, S_{5,t+j})$ and $\operatorname{cov}(S_{6,t}, S_{6,t+j})$. Because $\{e_{1,k}\}$, $\{e_{2,k}\}$, $\{\eta_{1,k}\}$, and $\{\eta_{2,k}\}$ are mutually independent, the remaining covariances are all 0.

A.7 Proof of Theorem 6

Here we treat the weak fractional cointegration case (case 1), the standard cointegration case (case 2), and the strong fractional cointegration case (case 3) separately.

Case 1: Fractional cointegration, $d_\eta \in (-\frac{1}{2}, 0)$. The log prices given by (7) can be written as

$$\begin{aligned}
A_j &\equiv \log P_{1,j} \\
&= \sum_{k=1}^{N_1(j\Delta t)} (e_{1,k} + \eta_{1,k}) + \sum_{k=1}^{N_2(t_{1,N_1(j\Delta t)})} (\theta e_{2,k} + g_{21}\eta_{2,k}), \\
B_j &\equiv \log P_{2,j} \\
&= \sum_{k=1}^{N_2(j\Delta t)} (e_{2,k} + \eta_{2,k}) + \sum_{k=1}^{N_1(t_{2,N_2(j\Delta t)})} \left(\frac{1}{\theta} e_{1,k} + g_{12}\eta_{1,k}\right) \\
&= \sum_{k=1}^{N_2(j\Delta t)} (e_{2,k} + \eta_{2,k}) + \sum_{k=1}^{N_1(j\Delta t)} \left(\frac{1}{\theta} e_{1,k} + g_{12}\eta_{1,k}\right) \\
&\quad - \sum_{k=N_1(t_{2,N_2(j\Delta t)})+1}^{N_1(j\Delta t)} \left(\frac{1}{\theta} e_{1,k} + g_{12}\eta_{1,k}\right) \\
&= \underbrace{\sum_{k=1}^{N_2(j\Delta t)} e_{2,k} + \frac{1}{\theta}\sum_{k=1}^{N_1(j\Delta t)} e_{1,k}}_{B_{1,j}} + \underbrace{\sum_{k=1}^{N_2(j\Delta t)} \eta_{2,k}}_{B_{2,j}} + g_{12} \underbrace{\sum_{k=1}^{N_1(j\Delta t)} \eta_{1,k}}_{B_{3,j}} \\
&\quad - \frac{1}{\theta} \underbrace{\sum_{k=N_1(t_{2,N_2(j\Delta t)})+1}^{N_1(j\Delta t)} e_{1,k}}_{B_{4,j}} - g_{12} \underbrace{\sum_{k=N_1(t_{2,N_2(j\Delta t)})+1}^{N_1(j\Delta t)} \eta_{1,k}}_{B_{5,j}},
\end{aligned}$$

and

$$\begin{aligned}
T_j &\equiv A_j - \theta B_j \\
&= \underbrace{\sum_{k=N_1(t_{2,N_2(j\Delta t)})+1}^{N_1(j\Delta t)} e_{1,k}}_{T_{1,j} = B_{4,j}} - \theta \underbrace{\sum_{k=N_2(t_{1,N_1(j\Delta t)})+1}^{N_2(j\Delta t)} e_{2,k}}_{T_{2,j}} \\
&\quad + (1 - \theta g_{12}) \underbrace{\sum_{k=1}^{N_1(j\Delta t)} \eta_{1,k}}_{T_{3,j} = B_{3,j}} + \theta g_{12} \underbrace{\sum_{k=N_1(t_{2,N_2(j\Delta t)})+1}^{N_1(j\Delta t)} \eta_{1,k}}_{T_{4,j} = B_{5,j}} \\
&\quad - (\theta - g_{21}) \underbrace{\sum_{k=1}^{N_2(j\Delta t)} \eta_{2,k}}_{T_{5,j} = B_{2,j}} - g_{21} \underbrace{\sum_{k=N_2(t_{1,N_1(j\Delta t)})+1}^{N_2(j\Delta t)} \eta_{2,k}}_{T_{6,j}}.
\end{aligned}$$

The OLS slope estimator, $\hat{\theta}$, obtained from regressing $\{\log P_{1,j}\}_{j=1}^n$ on $\{\log P_{2,j}\}_{j=1}^n$ is

$$\hat{\theta} = \frac{\sum_{j=1}^n A_j B_j}{\sum_{j=1}^n B_j^2} = \frac{\sum_{j=1}^n (\theta B_j + T_j) B_j}{\sum_{j=1}^n B_j^2} = \theta + \frac{\sum_{j=1}^n T_j B_j}{\sum_{j=1}^n B_j^2}. \quad (A.18)$$

First, we show that $n^{-r}\sum_{j=1}^n T_j B_j \xrightarrow{p} 0$, where $r = 2 + d_\eta + \delta$ for $\forall \delta > 0$. By the Cauchy–Schwartz inequality,

$$\frac{1}{n^r}\sum_{j=1}^n T_{i,j} B_{k,j} \le \sqrt{\left(\frac{1}{n^{2r-2}}\sum_{j=1}^n T_{i,j}^2\right)\left(\frac{1}{n^2}\sum_{j=1}^n B_{k,j}^2\right)}. \quad (A.19)$$

Thus it is sufficient to show that the right side of (A.19) converges in probability to 0 for all $i = 1, \ldots, 6$ and $k = 1, \ldots, 5$. We showed this earlier (Hurvich and Wang 2009). In that work, we also showed that $\frac{1}{(1/n^2)\sum_{j=1}^n B_j^2}$ is $O_p(1)$ by bounding it by a random variable that converges in distribution. Here we make use of the assumption that the $\{e_{i,j}\}$ are Gaussian. Thus

$$n^{2-r}(\hat{\theta} - \theta) = \frac{(1/n^r)\sum_{j=1}^n T_j B_j}{(1/n^2)\sum_{j=1}^n B_j^2} \xrightarrow{p} 0.$$

Case 3: standard cointegration, $d_\eta = -1$. When $d_\eta = -1$, $\eta_{1,k} = \xi_{1,k} - \xi_{1,k-1}$ and $\eta_{2,k} = \xi_{2,k} - \xi_{2,k-1}$. Denote

$$B_j \equiv \underbrace{\sum_{k=1}^{N_2(t_{1,N_1(j\Delta t)})} e_{2,k}}_{B_{1,j}^*} + \underbrace{\sum_{N_2(t_{1,N_1(j\Delta t)})+1}^{N_2(j\Delta t)} e_{2,k}}_{B_{2,j}^*} + \frac{1}{\theta}\underbrace{\sum_{k=1}^{N_1(t_{2,N_2(j\Delta t)})} e_{1,k}}_{B_{3,j}^*}$$
$$+ g_{12}\cdot\underbrace{\xi_{1,N_1(t_{2,N_2(j\Delta t)})} I\{N_1(t_{2,N_2(j\Delta t)}) > 0\}}_{B_{4,j}^*}$$
$$+ \underbrace{\xi_{2,N_2(j\Delta t)} I\{N_2(j\Delta t) > 0\}}_{B_{5,j}^*}$$

and

$$T_j \equiv A_j - \theta B_j$$
$$= \underbrace{\sum_{k=N_1(t_{2,N_2(j\Delta t)})+1}^{N_1(j\Delta t)} e_{1,k}}_{T_{1,j}^*} - \theta \underbrace{\sum_{k=N_2(t_{1,N_1(j\Delta t)})+1}^{N_2(j\Delta t)} e_{2,k}}_{T_{2,j}^*=B_{2,j}^*}$$
$$+ \underbrace{\xi_{1,N_1(j\Delta t)} I\{N_1(j\Delta t) > 0\}}_{T_{3,j}^*}$$
$$- \theta g_{12}\cdot\underbrace{\xi_{1,N_1(t_{2,N_2(j\Delta t)})} I\{N_1(t_{2,N_2(j\Delta t)}) > 0\}}_{T_{4,j}^*=B_{4,j}^*}$$
$$- \theta\cdot\underbrace{\xi_{2,N_2(j\Delta t)} I\{N_2(j\Delta t) > 0\}}_{T_{5,j}^*=B_{5,j}^*}$$
$$+ g_{21}\cdot\underbrace{\xi_{2,N_2(t_{1,N_1(j\Delta t)})} I\{N_1(j\Delta t) > 0\}}_{T_{6,j}^*}.$$

(1) First, we consider $\sum_{j=1}^n B_{1,j}^* T_{1,j}^*$. Earlier (Hurvich and Wang 2009), we showed that

$$\operatorname{var}\left(\sum_{j=1}^n B_{1,j}^* T_{1,j}^*\right) \le O(n^2) + Kn\sum_{j=1}^n\sum_{s=j+1}^n (s-j)^{m(d_\tau - 1/2)/4}. \quad (A.20)$$

Consider $\sum_{j=1}^n\sum_{s=j+1}^n (s-j)^{m(d_\tau-1/2)/4}$. For any fixed integer $1 \le j \le n$, we choose $m > \frac{8}{1-2d_\tau}$ so that $\sum_{s=j+1}^n (s-j)^{m(d_\tau-1/2)/4}$ is summable in s, and thus $\sum_{j=1}^n\sum_{s=j+1}^n (s-j)^{m(d_\tau-1/2)/4} = O(n)$. Therefore, $\operatorname{var}(\sum_{j=1}^n B_{1,j}^* T_{1,j}^*) = O(n^2)$, and, by Chebyshev's inequality, we obtain that for any $\delta > 0$,

$$\frac{1}{n^{1+\delta}}\sum_{j=1}^n B_{1,j}^* T_{1,j}^* \xrightarrow{p} 0.$$

(2) Earlier (Hurvich and Wang 2009), we used Isserlis' formula (Isserlis 1918) to show that, similarly as in (1),

$$\frac{1}{n^{1+\delta}}\sum_{j=1}^n B_{1,j}^* T_{2,j}^* \xrightarrow{p} 0 \qquad \forall\delta > 0.$$

(3) Similar to (1), for $\sum_{j=1}^n B_{1,j}^* T_{3,j}^* = \sum_{j=1}^n B_{1,j}^* \xi_{1,N_1(j\Delta t)} \times I\{N_1(j\Delta t) > 0\}$, we have

$$\operatorname{var}\left(\sum_{j=1}^n B_{1,j}^* \xi_{1,N_1(j\Delta t)} I\{N_1(j\Delta t) > 0\}\right)$$
$$\le \sigma_{2,e}^2\sigma_{1,\xi}^2 \sum_{j=1}^n E[N_2(t_{1,N_1(j\Delta t)})]$$
$$+ 2\sigma_{2,e}^2\sigma_{1,\xi}^2 \sum_{j=1}^n\sum_{s=j+1}^n\sum_{r=0}^\infty E[N_2(t_{1,N_1(j\Delta t)})$$
$$\cdot I\{N_1(s\Delta t) - N_1(j\Delta t) = r\}]\cdot|c_{\xi_{1,r}}|$$
$$\le \sigma_{2,e}^2\sigma_{1,\xi}^2 \sum_{j=1}^n E[N_2(j\Delta t)]$$
$$+ 2\sigma_{2,e}^2\sigma_{1,\xi}^2 \sum_{j=1}^n\sum_{s=j+1}^n\sum_{r=0}^\infty E[N_2(j\Delta t)$$
$$\cdot I\{N_1(s\Delta t) - N_1(j\Delta t) = r\}]\cdot|c_{\xi_{1,r}}|$$
$$\le \sigma_{2,e}^2\sigma_{1,\xi}^2\lambda_2\Delta t\frac{n(n+1)}{2}$$
$$+ 2\sigma_{2,e}^2\sigma_{1,\xi}^2 \sum_{j=1}^n\sum_{s=j+1}^n\sum_{r=0}^\infty \sqrt{E\{[N_2(j\Delta t)]^2\}}$$
$$\cdot\sqrt{P[N_1(s\Delta t) - N_1(j\Delta t) = r]}\cdot|c_{\xi_{1,r}}|$$
$$\le \sigma_{2,e}^2\sigma_{1,\xi}^2\lambda_2\Delta t\frac{n(n+1)}{2}$$

$$+2\sigma_{2,e}^2\sigma_{1,\xi}^2 \underbrace{\sqrt{E\{[N_2(n\Delta t)]^2\}}}_{O(n)}$$

$$\cdot \sum_{j=1}^n \sum_{s=j+1}^n \sum_{r=0}^\infty \sqrt{P[N_1(s\Delta t) - N_1(j\Delta t) \le r]} \cdot |c_{\xi_{1,r}}|.$$

Because

$$\begin{aligned}
&\sum_{r=0}^\infty \sqrt{P[N_1(s\Delta t) - N_1(j\Delta t) \le r]} \cdot |c_{\xi_{1,r}}| \\
&\quad \le \sum_{r=0}^\infty |c_{\xi_{1,r}}| \sqrt{\frac{E|Z_1(s-j) - r(s-j)^{-1/2-d_\tau}|^m}{\lambda_1^m (s-j)^{m(1/2-d_\tau)}}} \\
&\quad \le (s-j)^{(m/2)(d_\tau - 1/2)} \\
&\qquad \cdot C_m \sum_{r=0}^\infty e^{-K_{\xi_1} r}\big[1 + r^{m/2}(s-j)^{(m/2)(-1/2-d_\tau)}\big] \\
&\quad = O\big((s-j)^{(m/2)(d_\tau - 1/2)}\big),
\end{aligned}$$

we can choose m sufficiently large so that $\sum_{s=j+1}^n \sum_{r=0}^\infty \sqrt{P[N_1(s\Delta t) - N_1(j\Delta t) \le r]} \cdot |c_{\xi_{1,r}}|$ is summable in s. Thus

$$\underbrace{\sqrt{E\{[N_2(n\Delta t)]^2\}}}_{O(n)}$$

$$\cdot \sum_{j=1}^n \underbrace{\sum_{s=j+1}^n \sum_{r=0}^\infty \sqrt{P[N_1(s\Delta t) - N_1(j\Delta t) \le r]} \cdot |c_{\xi_{1,r}}|}_{\text{summable in } s} = O(n^2).$$

Therefore, $\mathrm{var}(\sum_{j=1}^n B_{1,j}^* \xi_{1,N_1(j\Delta t)} I\{N_1(j\Delta t) > 0\}) = O(n^2)$ and

$$\frac{1}{n^{1+\delta}} \sum_{j=1}^n B_{1,j}^* T_{3,j}^* \xrightarrow{p} 0 \qquad \forall \delta > 0$$

using Chebyshev's inequality.

By similar arguments for $\sum_{j=1}^n B_{1,j}^* T_{3,j}^*$, we obtain that $\forall \delta > 0$,

$$\frac{1}{n^{1+\delta}} \sum_{j=1}^n B_{1,j}^* T_{i,j} \xrightarrow{p} 0, \qquad i = 4, 5, 6.$$

(4) The proof for $\sum_{j=1}^n B_{3,j}^* T_{i,j}$ $(i = 1, \ldots, 6)$ follows along similar lines as for $\sum_{j=1}^n B_{1,j}^* T_{i,j}$ $(i = 1, \ldots, 6)$, because $B_{3,j}^*$ and $B_{1,j}^*$ are essentially the same, because one is for asset 1 and the other is for asset 2. Thus $\forall \delta > 0$,

$$\frac{1}{n^{1+\delta}} \sum_{j=1}^n B_{3,j}^* T_{i,j}^* \xrightarrow{p} 0, \qquad i = 1, \ldots, 6.$$

(5) The remaining terms, $\sum_{j=1}^n B_{i,j}^* T_{k,j}^*$ $(i = 2, 4, 5$ and $k = 1, \ldots, 6)$, are all $O_p(n)$, as can be easily shown using the Cauchy–Schwartz inequality and Chebyshev's inequality. Consider the following examples: (5.1) We have

$$\sum_{j=1}^n B_{2,j}^* T_{1,j}^* \le \sqrt{\sum_{j=1}^n B_{2,j}^{*2} \cdot \sum_{j=1}^n T_{1,j}^{*2}} = O_p(n),$$

because, by Chebyshev's inequality, for any $\epsilon > 0$, we can choose $M > \frac{\sigma_{2,e}^2 \tilde{C}_2}{\epsilon}$, so that

$$\begin{aligned}
P\left(\frac{1}{n}\sum_{j=1}^n B_{2,j}^{*2} > M\right) &\le \frac{E(\sum_{j=1}^n B_{2,j}^{*2})}{nM} \\
&= \frac{\sum_{j=1}^n \mathrm{var}(B_{2,j}^*)}{nM} = \frac{\sigma_{2,e}^2 \tilde{C}_2}{M} < \epsilon
\end{aligned}$$

and similarly $\sum_{j=1}^n T_{1,j}^{*2} = O_p(n)$.

(5.2) We have

$$\sum_{j=1}^n B_{2,j}^* T_{2,j}^* = \sum_{j=1}^n B_{2,j}^{*2} = O_p(n);$$

therefore, $\forall \delta > 0$,

$$\frac{1}{n^{1+\delta}} \sum_{j=1}^n B_{i,j}^* T_{k,j}^* \xrightarrow{p} 0, \qquad i = 2, 4, 5 \text{ and } k = 1, \ldots, 6.$$

(6) Overall, when $d_\eta = -1$,

$$\frac{1}{n^{1+\delta}} \sum_{j=1}^n B_j T_j \xrightarrow{p} 0 \tag{A.21}$$

for any $\delta > 0$.

Furthermore, the proof of $\frac{1}{(1/n^2)\sum_{j=1}^n B_j^2} = O_p(1)$ in the standard cointegration case is identical to that for the fractional cointegration case, except that here we have $\mathrm{var}(\sum_{k=1}^{N_i(t)} \eta_{i,k}) \le 2\sigma_{i,\xi}^2$ $(i = 1, 2)$, which does not increase with t. (We still have the telescope sum even if ξ_i is not iid and the variance of the partial sum is still some constant.) This, together with (A.21), gives that

$$n^{1-\delta}(\hat{\theta} - \theta) \xrightarrow{p} 0.$$

Case 2: Strong fractional cointegration, $d_\eta \in (-1, -\frac{1}{2})$. Following along the same lines as the proof of case 1, we can show that the convergence rate of $\hat{\theta}$ is arbitrarily close to $\sqrt{n}$, using the fact that the variance of the partial sums of the microstructure noise is a constant and not increasing with time.

ACKNOWLEDGMENTS

The authors thank Yakov Amihud, Torben Andersen, Jushan Bai, Menachem Brenner, Xiaohong Chen, Rohit Deo, Rob Engle, Steve Figlewski, Joel Hasbrouck, Jim Stock, Mark Watson, and the participants at the 2006 NBER-NSF Time Series conference in Montreal, and at the Econometrics Seminar, Department of Economics, New York University, for helpful comments and suggestions. The authors also thank two anonymous referees and the associate editor for their help in improving the article.

[Received March 2007. Revised November 2008.]

REFERENCES

Alexander, C., and Dimitriu, A. (2005), "Indexing, Cointegration and Equity Market Regimes," *International Journal of Finance and Economics*, 10, 213–231. [551]

Amihud, Y., and Mendelson, H. (1980), "Dealership Markets: Market-Making With Inventory," *Journal of Financial Economics*, 8, 31–53. [549]

——— (1982), "Asset Price Behavior in a Dealershipt Market," *Financial Analysts Journal*, May/June. [549,550]

——— (1987), "Trading Mechanisms and Stock Returns: An Empirical Investigation," *Journal of Finance*, XLII, 533–553. [540,541,550]

Andersen, T., Bollerslev, T., Frederiksen, P., and Nielsen, M. (2006), "Continuous-Time Models, Realized Volatilities, and Testable Distributional Implications for Daily Stock Returns," working paper, University of Aarhus. [551]

Bertelli, S., and Caporin, M. (2002), "A Note on Calculating Autocovariances of Long-Memory Processes," *Journal of Time Series Analysis*, 23, 503–508. [546]

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327. [540]

Boudoukh, J., Richardson, M., and Whitelaw, R. (1994), "A Tale of Three Schools: Insights on Autcorrelations of Short-Horizon Returns," *Review of Financial Studies*, 7, 539–573. [551]

Bowsher, C. G. (2007), "Modelling Security Market Events in Continuous Time: Intensity Based, Multivariate Point Process Models," *Journal of Econometrics*, 141, 876–912. [542]

Breidt, F., Crato, N., and de Lima, P. (1998), "The Detection and Estimation of Long Memory in Stochastic Volatility," *Journal of Econometrics*, 83, 325–348. [540]

Chen, W., and Hurvich, C. (2003a), "Estimating Fractional Cointegration in the Presence of Polynomial Trends," *Journal of Econometrics*, 117, 95–121. [539]

——— (2003b), "Semiparametric Estimation of Multivariate Fractional Cointegration," *Journal of the American Statistical Association*, 98, 629–642. [539]

——— (2006), "Semiparametric Estimation of Fractional Cointegrating Subspaces," *The Annals of Statistics*, 34, 2939–2979. [539]

Comte, F. (1998), "Discrete and Continuous Time Cointegration—Le Cas Multidimensionnel," *Journal of Econometrics*, 88, 207–226. [539]

Comte, F., and Renault, E. (1996), "Long Memory Continuous Time Models," *Journal of Econometrics*, 73, 101–149. [539]

——— (1998), "Long Memory in Continuous-Time Stochastic Volatility Models," *Mathematical Finance*, 8, 291–323. [539]

Daley, D. J., and Vere-Jones, D. (2003), *An Introduction to the Theory of Point Processes* (2nd ed.), New York: Springer. [544,552]

Deo, R., Hsieh, M., and Hurvich, C. (2006), "The Persistence of Memory: From Durations to Realized Volatility," working paper, New York University, Stern School of Business. [540,542,547]

Deo, R., Hurvich, C., Soulier, P., and Wang, Y. (2009), "Conditions for the Propagation of Memory Parameter From Durations to Counts and Realized Volatility," *Econometric Theory*, 25, 764–792. [540,542,544,555]

Duan, J. C., and Pliska, S. (2004), "Option Valuation With Cointegrated Asset Prices," *Journal of Economic Dynamics and Control*, 28, 727–754. [551]

Dueker, M., and Startz, R. (1998), "Maximum-Likelihood Estimation of Fractional Cointegration With an Application to U.S. and Canadian Bond Rates," *The Review of Economics and Statistics*, 80, 420–426. [539]

Dunis, C. L., and Ho, R. (2005), "Cointegration Portfolios of European Equities for Index Tracking and Market Neutral Strateries," *Journal of Asset Management*, 6, 33–52. [551]

Engle, R., and Granger, C. W. J. (1987), "Co-Integration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 66, 1127–1162. [539]

Engle, R., and Russell, J. (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127–1162. [540]

Fisher, L. (1966), "Some New Stock Market Indices," *Journal of Business*, 39, 191–225. [551]

Frijns, B., and Schotman, P. (2006), "Price Discovery in Tick Time," Working Paper, SC-CFE-05-02, New York University. [542]

Geweke, J., and Porter-Hudak, S. (1983), "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, 4, 221–237. [543,549]

Glosten, L. R. (1987), "Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices," *Journal of Finance*, XLII, 1293–1307. [540]

Harvey, A. C. (1998), "Long Memory in Stochastic Volatility," in *Forecasting Volatility in Financial Markets*, eds. J. Knight and S. Satchell, Oxford: Butterworth–Heineman, pp. 307–320. [540]

Hasbrouck, J. (1995), "One Security, Many Markets: Determining the Contributions to Price Discovery," *Journal of Finance*, 50, 1175–1199. [542-544, 549]

Hausman, J. (1978), "Specification Tests in Econometrics," *Econometrica*, 46, 1251–1271. [547]

Hurvich, C. M., and Wang, Y. (2009), "A Pure-Jump Transaction-Level Price Model Yielding Cointegration, Leverage, and Nonsynchronous Trading Effects," working paper, New York University, Stern School of Business. [546,549,551-553,555,556]

Isserlis, L. (1918), "On a Formula for the Product-Moment Coefficient of Any Order of a Normal Frequency Distribution in Any Number of Variables," *Biometrika*, 12, 134–139. [556]

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, 231–254. [539]

——— (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551–1580. [539]

Kadlec, G., and Patterson, D. (1999), "A Transactions Data Analysis of Nonsynchronous Trading," *Review of Financial Studies*, 12, 609–630. [551]

Lee, C., and Ready, M. (1991), "Inferring Trade Direction From Intraday Data," *Journal of Finance*, 46, 733–746. [549]

Lieberman, O., and Phillips, P. C. B. (2006), "A Complete Asymptotic Series for the Autocovariance Function of a Long Memory Process," working paper, Yale University. [544]

Liu, J., Longstaff, F., and Pan, J. (2003), "Dynamic Asset Allocation With Event Risk," *Journal of Finance*," 58, 231–259. [551]

Lo, A., and MacKinlay, A. C. (1990a), "An Econometric Analysis of Nonsynchronous Trading," *Journal of Econometrics*, 45, 181–211. [551]

——— (1990b), "When Are Constrarian Profits Due to Stock Market Overreaction?" *Review of Financial Studies*, 3, 175–205. [551]

Marinucci, D., and Robinson, P. M. (2001), "Semiparametric Fractional Cointegration Analysis. Forecasting and Empirical Methods in Finance and Macroeconomics," *Journal of Econometrics*, 105, 225–247. [539]

Nijman, T. E., Spierdijk, L., and Soest, A. H. O. (2004), "Modeling Co-Movements in Trading Intensities to Distinguish Stock and Sector-Specific News," working paper, Tilburg University. [542]

Oomen, R. C. A. (2006), "Properties of Realized Variance Under Alternative Sampling Schemes," *Journal of Business & Economic Statistics*, 24, 219–237. [540]

Phillips, P. C. B. (1991a), "Optimal Inference in Cointegrated Systems," *Econometrica*, 59, 283–306. [539]

——— (1991b), "Error Correction and Long-Run Equilibrium in Continuous Time," *Econometrica*, 59, 967–980. [539]

Phillips, P. C. B., and Durlauf, S. (1986), "Multiple Time Series Regression With Integrated Processes," *The Review of Economic Studies*, 53, 473–495. [545]

Press, S. J. (1967), "A Compound Events Model for Security Prices," *Journal of Business*, 40, 317–335. [540]

Prigent, J.-L. (2001), "Option Pricing With a General Marked Point Process," *Mathematics of Operations Research*, 26, 50–66. [551]

Robinson, P. M. (1995), "Gaussian Semiparametric Estimation of Long Range Dependence," *The Annals of Statistics*, 23, 1630–1661. [544,550]

Robinson, P. M., and Haulde, J. (2003), "Cointegration in Fractional Systems With Unknown Integration Orders," *Econometrica*, 71, 1727–1766. [539]

Robinson, P. M., and Marinucci, D. (2001), "Narrow-Band Analysis of Nonstationary Processes," *The Annals of Statistics*, 29, 947–986. [539]

Robinson, P. M., and Yajima, Y. (2002), "Determination of Cointegrating Rank in Fractional Systems," *Journal of Econometrics*, 106, 217–241. [539]

Scholes, M., and Williams, J. (1977), "Estimating Betas From Nonsynchronous Data," *Journal of Financial Economics*, 5, 309–327. [551]

Stock, J. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, 55, 1035–1056. [545,552]

Stock, J., and Watson, M. (1988), "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097–1107. [539,541,552]

Velasco, C. (2003), "Gaussian Semiparametric Estimation of Fractional Cointegration," *Journal of Time Series Analysis*, 24, 345–378. [539]

Velasco, C., and Marmol, F. (2004), "Consistent Testing of Cointegration Relationships," *Econometrica*, 72, 1809–1844. [539]