

TESTING THE MARKOV PROPERTY WITH ULTRA HIGH  
FREQUENCY FINANCIAL DATA\*

JOÃO AMARO DE MATOS  
Faculdade de Economia  
Universidade Nova de Lisboa  
Rua Marquês de Fronteira, 20  
1099-038 Lisbon, Portugal  
Tel: +351.21.3826100  
Fax: +351.21.3873973  
amatos@fe.unl.pt

MARCELO FERNANDES  
Graduate School of Economics  
Fundação Getulio Vargas  
Praia de Botafogo, 190  
22953-900 Rio de Janeiro, Brazil  
Tel: +55.21.5595827  
Fax: +55.21.5538821  
mfernand@fgv.br

March 2001

\* The second author gratefully acknowledges the hospitality of the Universidade Nova de Lisboa, where part of this paper was written, and a Jean Monnet fellowship at the European University Institute.

TESTING THE MARKOV PROPERTY WITH ULTRA HIGH  
FREQUENCY FINANCIAL DATA

**ABSTRACT:** This paper develops a framework to test whether discrete-valued irregularly-spaced financial transactions data follow a subordinated Markov process. For that purpose, we consider a specific optional sampling in which a continuous-time Markov process is observed only when it crosses some discrete level. This framework is convenient for it accommodates not only the irregular spacing of transactions data, but also price discreteness. Further, it turns out that, under such an observation rule, the current price duration is independent of previous price durations given the current price realization. A simple nonparametric test then follows by examining whether this conditional independence property holds. Finally, we investigate whether or not bid-ask spreads follow Markov processes using transactions data from the New York Stock Exchange. The motivation lies on the fact that asymmetric information models of market microstructures predict that the Markov property does not hold for the bid-ask spread. The results are mixed in the sense that the Markov assumption is rejected for three out of the five stocks we have analyzed.

**JEL CLASSIFICATION:** C14, C52, G10, G19.

**KEYWORDS:** Bid-ask spread, nonparametric tests, price durations, subordinated Markov process, ultra-high frequency data.

## 1. INTRODUCTION

Despite the innumerable studies in financial economics rooted in the Markov property, there are only two tests available in the literature to check such an assumption: Aït-Sahalia (1997) and Fernandes and Flôres (1999). To build a nonparametric testing procedure, the first uses the fact that the Chapman-Kolmogorov equation must hold in order for a Markov process compatible with the data to exist. If, on the one hand, the Chapman-Kolmogorov representation involves a quite complicated nonlinear functional relationship among transition probabilities of the process, on the other hand, it brings about several advantages. First, estimating transition distributions is straightforward and does not require any prior parameterization of conditional moments. Second, a test based on the whole transition density is obviously preferable to tests based on specific conditional moments. Third, the Chapman-Kolmogorov representation is well defined, even within a multivariate context.

Fernandes and Flôres (1999) develop alternative ways of testing whether discretely recorded observations are consistent with an underlying Markov process. Instead of using the highly nonlinear functional characterization provided by the Chapman-Kolmogorov equation, they rely on a simple characterization out of a set of necessary conditions for Markov models. As in Aït-Sahalia (1997), the testing strategy boils down to measuring the closeness of density functionals which are nonparametrically estimated by kernel-based methods.

Both testing procedures assume, however that the data are equally spaced in time. Financial transactions data do not satisfy such an assumption and hence these tests are not appropriate. To design a consistent test for the Markov property that is suitable to ultra-high frequency data, we build on the theory of subordinated Markov processes. We assume that there is an underlying continuous-time Markov process that is observed only when it crosses some discrete level. Accordingly, we accommodate not only the irregular spacing of transaction data, but also price discreteness. Further, such an optional sampling scheme implies that consecutive spells between price changes are conditionally independent given the current price realization. This paper then develops a simple nonparametric test for the Markov property by testing whether this conditional independence property holds.

An empirical application is performed using data from five stocks actively traded on the New York Stock Exchange (NYSE), namely Boeing, Coca-Cola, Disney, Exxon, and IBM. Unfortunately, all bid and ask prices seem integrated of order one and hence nonstationary. Notwithstanding, there is no evidence of unit roots in the bid-ask spreads and so they serve as input. The results indicate that the Markov assumption is consistent with the Disney and Exxon bid-ask spreads, whereas the converse is true for Boeing, Coca-Cola and IBM. A possible explanation for the non-Markovian character of the bid-ask spreads relies on sufficiently high adverse selection costs. Asymmetric information models of market microstructure predict that the bid-ask spread depends on the whole trading history, so that the Markov property does not hold (Easley and O'Hara, 1992).

The remainder of this paper is organized as follows. Section 2 discusses how to design a nonparametric test for Markovian dynamics that is suitable to high frequency data. The asymptotic normality of the test statistic is then derived both under the null hypothesis that the Markov property holds and under a sequence of local alternatives. Section 3 applies the above ideas to test whether the bid-ask spreads of five actively traded stocks in the NYSE follow a subordinated Markov process. Section 4 summarizes the results and offers some concluding remarks. For ease of exposition, we collect all proofs and technical lemmas in the appendix.

## 2. TESTING SUBORDINATED MARKOV PROCESSES

Let  $t_i$  ( $i = 1, 2, \dots$ ) denote the observation times of the continuous-time price process  $\{X_t, t > 0\}$  and assume that  $t_0 = 0$ . Suppose further that the shadow price  $\{X_t, t > 0\}$  follows a strong Markov process. To account for price discreteness, we assume that prices are observed only when the cumulative change in the shadow price is at least  $c$ , say a basic tick. The price duration then reads

$$(1) \quad d_{i+1} \equiv t_{i+1} - t_i = \inf_{\tau > 0} \{|X_{t_i+\tau} - X_{t_i}| \geq c\}$$

for  $i = 0, \dots, n - 1$ . The data available for statistical inference are the price durations  $(d_1, \dots, d_n)$  and the corresponding realizations  $(X_1, \dots, X_n)$ , where  $X_i = X_{t_i}$ .

The observation times  $\{t_i, i = 1, 2, \dots\}$  form a sequence of increasing stopping times of the continuous-time Markov process  $\{X_t, t > 0\}$ , hence the

discrete-time price process  $\{X_i, i = 1, 2, \dots\}$  satisfies the Markov property as well. Further, the price duration  $d_{i+1}$  is a measurable function of the path of  $\{X_t, 0 < t_i \leq t \leq t_{i+1}\}$ , and thus depends on the information available at time  $t_i$  only through  $X_i$  (Burgayran and Darolles, 1997). In other words, the sequence of price durations are conditionally independent given the observed price (Dawid, 1979). Therefore, one can test the Markov assumption by checking the property of conditional independence between consecutive durations given the current price realization.

The null hypothesis of conditional independence implied by the Markov character of the price process then reads

$$(2) \quad H_0 : f_{iX_j}(a_1, x, a_2) = f_{i|X}(a_1)f_{X_j}(x, a_2) \text{ a.s.},$$

where  $f_{iX_j}(\cdot, \cdot, \cdot)$ ,  $f_{i|X}(\cdot)$  and  $f_{X_j}(\cdot, \cdot)$  denote the joint density of  $(d_i, X_i, d_j)$ , the conditional density of  $d_i$  given  $X_i$  and the joint density of  $(X_i, d_j)$ , respectively. To keep the nonparametric nature of the testing procedure, we employ kernel smoothing to estimate both the right- and left-hand sides of (2). Next, it suffices to gauge how well the density restriction in (2) fits the data by the means of some discrepancy measure.

For the sake of simplicity, we consider the mean squared difference, yielding the following test statistic

$$(3) \quad \Lambda_f = E[f_{iX_j}(d_i, X_i, d_j) - f_{i|X}(d_i|X_i)f_{X_j}(X_i, d_j)]^2.$$

The sample analog is then

$$\Lambda_{\hat{f}} = \frac{1}{n - i + j} \sum_{k=1}^{n-i+j} [\hat{f}_{iX_j}(d_{k+i-j}, X_{k+i-j}, d_k) - \hat{g}_{iX_j}(d_{k+i-j}, X_{k+i-j}, d_k)]^2,$$

where  $\hat{g}_{iXj}(d_{k+i-j}, X_{k+i-j}, d_k) = \hat{f}_{i|X}(d_{k+i-j}|X_{k+i-j})\hat{f}_{Xj}(X_{k+i-j}, d_k)$ . Any other evaluation of the integral on the right-hand side of (3) can be used.

At first glance, deriving the limiting distribution of  $\Lambda_{\hat{f}}$  seems to involve a number of complex steps since one must deal with the cross-correlation among  $\hat{f}_{iXj}$ ,  $\hat{f}_{i|X}$  and  $\hat{f}_{Xj}$ . Happily, the fact that the rates of convergence of the three estimators are different simplifies things substantially. In particular,  $\hat{f}_{iXj}$  converges slower than  $\hat{f}_{i|X}$  and  $\hat{f}_{Xj}$  due to its higher dimensionality. As such, estimating the conditional density  $f_{i|X}$  and the joint density  $f_{Xj}$  does not play a role in the asymptotic behavior of the test statistic.

To derive the necessary asymptotic theory, we impose the following regularity conditions as in Ait-Sahalia (1994).

**A1** The sequence  $\{d_i, X_i, d_j\}$  is strictly stationary and  $\beta$ -mixing with  $\beta_r = O(r^{-\delta})$  as  $r \rightarrow \infty$ , where  $\delta > 1$ . Further,  $E\|(d_i, X_i, d_j)\|^k < \infty$  for some constant  $k > 2\delta/(\delta - 1)$ .

**A2** The density function  $f_{iXj}$  is continuously differentiable up to order  $s + 1$  and its derivatives are bounded and square integrable.

**A3** The kernel  $K$  is of order  $s$  (even integer) and is continuously differentiable up to order  $s$  on  $\mathbb{R}^3$  with derivatives in  $L^2(\mathbb{R}^3)$ . Let  $e_K \equiv \int |K(u)|^2 du$  and  $v_K \equiv \int [\int K(u)K(u+v) du]^2 dv$ .

**A4** The bandwidths  $b_{d,n}$  and  $b_{x,n}$  are of order  $o(n^{-2/(4s+3)})$  as the sample size  $n$  grows.

Assumption A1 restricts the amount of dependence allowed in the observed data sequence to ensure that the central limit theorem holds. As usual, there is a trade-off between the degree of dependence and the number of finite moments. Assumption A2 requires that the joint density function  $f_{iX_j}$  is smooth enough to admit a functional Taylor expansion. Although assumption A3 provides enough room for higher order kernels, hereinafter, we implicitly assume that the kernel is of second order ( $s = 2$ ). Assumption A4 restricts the rate at which the bandwidth must converge to zero.

The following proposition documents the asymptotic normality of the test statistic.

**PROPOSITION 1:** *Under the null and assumptions A1 to A4, the statistic*

$$\hat{\lambda}_n = \frac{n b_n^{1/2} \Lambda_{\hat{f}} - b_n^{-1/2} \hat{\delta}_\Lambda}{\hat{\sigma}_\Lambda} \xrightarrow{d} N(0, 1),$$

where  $b_n = b_{d,n}^2 b_{x,n}$  is the bandwidth for the kernel estimation of the joint density  $f_{iX_j}$ , and  $\hat{\delta}_\Lambda$  and  $\hat{\sigma}_\Lambda^2$  are consistent estimates of  $\delta_\Lambda = e_K E(f_{iX_j})$  and  $\sigma_\Lambda^2 = v_K E(f_{iX_j}^3)$ , respectively.

Ergo, a test that rejects the null hypothesis at level  $\alpha$  when  $\hat{\lambda}_n$  is greater or equal to the  $(1 - \alpha)$ -quantile  $z_{1-\alpha}$  of a standard normal distribution is locally strictly unbiased.

Consider now the following sequence of local alternatives

$$(4) \quad H_1^{[n]} : \sup \left| f_{iX_j}^{[n]}(a_1, x, a_2) - g_{iX_j}^{[n]}(a_1, x, a_2) - \epsilon_n \ell(a_1, x, a_2) \right| = o(\epsilon_n),$$

where  $\left\| f_{iX_j}^{[n]} - f_{iX_j} \right\| = o\left(n^{-1} b_n^{-1/2}\right)$ ,  $\epsilon_n = n^{-1/2} b_n^{-1/4}$  and  $\ell(\cdot, \cdot, \cdot)$  is such that  $E[\ell(a_1, x, a_2)] = 0$  and  $\ell_2 \equiv E[\ell^2(a_1, x, a_2)] < \infty$ . The next result

illustrates the fact that the testing procedure entails nontrivial power under local alternatives that shrink to the null at rate  $\epsilon_n$ .

PROPOSITION 2: *Under the sequence of local alternatives  $H_1^{[n]}$  and assumptions A1 to A4,  $\hat{\lambda}_n \xrightarrow{d} N(\ell_2/\sigma_\Lambda, 1)$ .*

Other testing procedures could well be developed relying on the restrictions imposed by the conditional independence property on the cumulative probability functions. For instance, Linton and Gozalo (1999) propose two nonparametric tests for conditional independence restrictions rooted in a generalization of the empirical distribution function. The motivation rests on the fact that, in contrast to smoothing-based tests, empirical measure-based tests usually have power against all alternatives at distance  $n^{-1/2}$ . Linton and Gozalo (1999) show that the asymptotic null distribution of the test statistic is a quite complicated functional of a Gaussian process.

This alternative approach entails two serious drawbacks, however. First, the asymptotic properties are derived in an iid setup, which is obviously not suitable for ultra-high frequency financial data. Second, the complex nature of the limiting null distribution calls for the use of bootstrap critical values. Design a bootstrap algorithm that imposes the null of conditional independence and deals with the time dependence feature is however a daunting task. In effect, Linton and Gozalo (1999) recognize that considerable additional work is necessary to extend their results to a time series context, while the bootstrap technology is still in process of development.

### 3. EMPIRICAL EXERCISE

We focus on New York Stock Exchange (NYSE) data ranging from September to November 1996. In particular, we look at five actively traded stocks from the Dow Jones index: Boeing, Coca-Cola, Disney, Exxon, and IBM.<sup>1</sup> Trading at the NYSE is organized as a combined market maker/order book system. A designated specialist composes the market for each stock by managing the trading and quoting processes and providing liquidity. Apart from an opening auction, trading is continuous from 9:30 to 16:00. Spread durations are defined as the time interval needed to observe a change either in the bid or in the ask price.

For all stocks, durations between events recorded outside the regular opening hours of the NYSE, as well as overnight spells, are removed. As documented by Giot (2000), durations feature a strong time-of-day effect related to predetermined market characteristics, such as trade opening and closing times and lunch time for traders. To account for this feature, we also consider seasonally adjusted spread durations  $d_i^* = d_i/\phi(t_i)$ , where  $d_i$  is the original spread duration in seconds and  $\phi(\cdot)$  denotes a time-of-day factor determined by averaging durations over thirty-minutes intervals for each day of the week and fitting a cubic spline with nodes at each half hour. Such a transformation may however destroy the Markovian character of the process, hence the results based on adjusted durations must be viewed with caution.

The reason why it is interesting to test whether the bid-ask spread follows

---

<sup>1</sup> Data were kindly provided by Luc Bauwens and Pierre Giot and refer to the NYSE's Trade and Quote (TAQ) database. Giot (2000) describes the data more thoroughly.

a Markov process is simple. Information-based models of market microstructure, such as Glosten and Milgrom (1985) and Easley and O’Hara (1987, 1992), predict that the quote-setting process depends on the whole trading history rather than exclusively on the most recent quote, and thus both bid and ask prices are non-Markovian. Therefore, one can test indirectly for the presence of asymmetric information by checking whether bid and ask prices satisfy the Markov property. Table I reports however that the bid and ask quotes are both integrated of order one, and hence nonstationary. In contrast, there is no evidence of unit roots in the bid-ask spread processes. As kernel density estimation relies on the assumption of stationarity, spread data are therefore more convenient to serve as input for the subsequent analysis.

All density estimations are carried out using a (product) Gaussian kernel, namely

$$(5) \quad K(u) = (2\pi)^{-3/2} \exp\left(-\frac{u_1^2 + u_2^2 + u_3^2}{2}\right),$$

which implies that  $e_K = (4\pi)^{-3/2}$  and  $v_K = (8\pi)^{-3/2}$ . Bandwidths are chosen according to Silverman’s (1986) rule of thumb adjusted so as to conform to the degree of undersmoothing required by Assumption A4.

Table II reports mixed results in the sense that the Markov hypothesis seems to suit only some of the bid-ask spreads under consideration. Clear rejection is detected in the Boeing, Coca-Cola and IBM bid-ask spreads, indicating that adverse selection may play a role in the formation of their prices. In contrast, there is no indication of non-Markovian behavior in the Disney and Exxon bid-ask spreads. Interestingly, the results are quite robust in the sense that they do not depend on whether the spread durations are

adjusted or not for the time-of-day effect. This is important because the Markov property is not invariant under such a transformation, so that conflicting results could cast doubts on the usefulness of the analysis. Further, it is also comforting that these results agree to some extent with Fernandes and Grammig's (2000) analysis. Using different techniques, they identify significant asymmetric information effects only in the Boeing and IBM price durations.

#### 4. CONCLUSION

This paper has developed a test for Markovian dynamics that is particularly tailored to ultra-high frequency data. This testing procedure is especially interesting to investigate whether data are consistent with information-based models of market microstructure. For instance, Easley and O'Hara (1987, 1992) predict that the price discovery process is such that the Markov assumption does not hold for the bid-ask spread set by the market maker.

Using data from the New York Stock Exchange, we show that whether the Markov hypothesis is reasonable or not is indeed an empirical issue. The results show that the Markov assumption seems inadequate for the Boeing, Coca-Cola and IBM bid-ask spreads, indicating that the market maker may account for asymmetric information in the quote-setting process. In contrast, a Markovian character suits the Disney and Exxon bid-ask spreads well, suggesting low adverse selection costs. Accordingly, market microstructure models rooted in Markov processes, such as Amaro de Matos and Rosário (2000), may deserve more attention.

APPENDIX: PROOFS

LEMMA 1: Consider the functional

$$I_n = \int \varphi(a_1, x, a_2) \left[ \hat{f}(a_1, x, a_2) - f(a_1, x, a_2) \right]^2 d(a_1, x, a_2).$$

Under assumptions A1 to A4,

$$n b_n^{1/2} I_n - b_n^{-1/2} e_K E [\varphi(a_1, x, a_2)] \xrightarrow{d} N \left( 0, v_K E \left[ \varphi^2(a_1, x, a_2) f(a_1, x, a_2) \right] \right),$$

provided that the above expectations are finite.

PROOF: Let  $z = (a_1, x, a_2)$ ,  $r_n(z, Z) = \varphi(z)^{1/2} K_{b_n}(z - Z)$ , where  $K_{b_n}(z) = b_n^{-1} K(z/b_n)$ , and  $\check{r}_n(z, Z) = r_n(z, Z) - E_Z[r_n(z, Z)]$ . Consider then the following decomposition

$$\begin{aligned} I_n &= \int \varphi(z) [\hat{f}(z) - E\hat{f}(z)]^2 dz + \int \varphi(z) [E\hat{f}(z) - f(z)]^2 dz \\ &\quad + 2 \int \varphi(z) [\hat{f}(z) - E\hat{f}(z)] [E\hat{f}(z) - f(z)] dz, \end{aligned}$$

or equivalently,  $I_n = I_{1n} + I_{2n} + I_{3n} + I_{4n}$ , where

$$I_{1n} = \frac{2}{n^2} \sum_{i < j} \int \check{r}_n(z, Z_i) \check{r}_n(z, Z_j) dz$$

$$I_{2n} = \frac{1}{n^2} \sum_i \int \check{r}_n^2(z, Z_i) dz$$

$$I_{3n} = \int \varphi(z) [E\hat{f}(z) - f(z)]^2 dz$$

$$I_{4n} = 2 \int \varphi(z) [\hat{f}(z) - E\hat{f}(z)] [E\hat{f}(z) - f(z)] dz.$$

We show in the sequel that the first term is a degenerate U-statistic and contributes with the variance in the limiting distribution, while the second gives the asymptotic bias. In turn, assumption A4 ensures that the third

and fourth terms are negligible. To begin, observe that the first moment of  $r_n(z, Z)$  reads

$$\begin{aligned}
E_Z[r_n(z, Z)] &= \varphi^{1/2}(z) \int K_{b_n}(z - Z) f(Z) dZ \\
&= \varphi^{1/2}(z) \int K(u) f(z + ub_n) du \\
&= \varphi^{1/2}(z) \int K(z) \left[ f(z) + \frac{1}{2} f'(z) ub_n + f''(z^*) u^2 b_n^2 \right] du \\
&= \varphi^{1/2}(z) f(z) + O(b_n^2),
\end{aligned}$$

where  $f^{(i)}(\cdot)$  denotes the  $i$ -th derivative of  $f(\cdot)$  and  $z^* \in [z, z + ub_n]$ . Applying similar algebra to the second moment yields  $E_Z[r_n^2(z, Z)] = b_n^{-1} e_K \varphi(z) f(z) + O(1)$ . This means that

$$\begin{aligned}
E(I_{2n}) &= \frac{1}{n} \int E_Z[r_n^2(z, Z)] dz - \frac{1}{n} \int E_Z^2[r_n(z, Z)] dz \\
&= \frac{1}{n} \int [b_n^{-1} e_K \varphi(z) f(z) + O(1)] dz + O(n^{-1}) \\
&= n^{-1} b_n^{-1} e_K \int \varphi(z) f(z) dz + O(n^{-1}),
\end{aligned}$$

whereas  $\text{Var}(I_{2n}) = O(n^{-3} b_n^{-2})$ . It then follows from Chebyshev's inequality that  $n b_n^{1/2} I_{2n} - b_n^{-1/2} e_K E[\varphi(z)] = o_p(1)$ . In turn, the deterministic term  $I_{3n}$  is proportional to the integrated squared bias of the fixed kernel density estimation, hence it is of order  $O(b_n^4)$ . Assumption A4 then implies that  $n b_n^{1/2} I_{3n} = o(1)$ . Further,

$$E(I_{4n}) = 2 \int \varphi(z) E_Z \left[ \hat{f}(z) - E\hat{f}(z) \right] \left[ E\hat{f}(z) - f(z) \right] dz = 0,$$

whereas  $E(I_{4n}^2) = O(n^{-1} b_n^4)$  as in Hall (1984, Lemma 1). It then suffices to impose assumption A4 to ensure, by Chebyshev's inequality, that  $n b_n^{1/2} I_{4n} =$

$o_p(1)$ . Lastly, recall that  $I_{1n} = \sum_{i < j} H_n(Z_i, Z_j)$ , where

$$H_n(Z_i, Z_j) = 2n^{-2} \int \check{r}_n(z, Z_i) \check{r}_n(z, Z_j) dz.$$

Because  $H_n(Z_i, Z_j)$  is symmetric, centered and such that  $E[H_n(Z_i, Z_j)|Z_j] = 0$  almost surely,  $I_{1n}$  is a degenerate U-statistic. Khashimov's (1992) central limit theorem for degenerate U-statistics implies that, under assumptions A1 to A4,  $n b_n^{1/2} I_{1n} \xrightarrow{d} N(0, \Omega)$ , where

$$\begin{aligned} \Omega &= \frac{n^4 b_n}{2} E_{Z_1, Z_2} [H_n^2(Z_1, Z_2)] \\ &= 2b_n \int_{Z_1, Z_2} \left[ \int \check{r}_n(z, Z_1) \check{r}_n(z, Z_2) dz \right]^2 f(Z_1, Z_2) d(Z_1, Z_2) \\ &= 2b_n \int \left[ \int \check{r}_n(z, Z) \check{r}_n(z', Z) f(Z) dZ \right]^2 d(z, z') \\ &= 2 \int \varphi^2(z) \left[ \int K(u) K(u+v) f(z - ub_n) du \right. \\ &\quad \left. - b_n \int K(u) f(z - ub_n) du \int K(u) f(z + vb_n - ub_n) du \right]^2 d(z, v) \\ &\cong 2 \int \varphi^2(z) \left[ \int K(u) K(u+v) f(z - ub_n) \right]^2 d(z, v) \\ &\cong 2 v_K \int \varphi^2(z) f(z) dF(z), \end{aligned}$$

which completes the proof.

**PROOF OF PROPOSITION 1:** Consider the second-order functional Taylor expansion

$$\Lambda_{f+h} = \Lambda_f + D\Lambda_f(h) + \frac{1}{2} D^2\Lambda_f(h, h) + O(\|h\|^3),$$

where  $h$  denotes the perturbation  $h_{iXj} = \hat{f}_{iXj} - f_{iXj}$ . Under the null hypothesis that  $f_{iXj} = g_{iXj}$ , both  $\Lambda_f$  and  $D\Lambda_f$  equal zero. To appreciate the

singularity of the latter, it suffices to compute the Gâteaux derivative of

$\Lambda_{f,h}(\lambda) = \Lambda_{f+\lambda h}$  with respect to  $\lambda$  evaluated at  $\lambda \rightarrow^+ 0$ . Let

$$g_{iX_j}(\lambda) = \frac{\int [f_{iX_j} + \lambda h_{iX_j}](a_1, x, a_2) da_2 \int [f_{iX_j} + \lambda h_{iX_j}](a_1, x, a_2) da_1}{\int [f_{iX_j} + \lambda h_{iX_j}](a_1, x, a_2) d(a_1, a_2)}.$$

It then follows that

$$\begin{aligned} \frac{\partial \Lambda_{f,h}(0)}{\partial \lambda} &= 2 \int [f_{iX_j} - g_{iX_j}][h_{iX_j} - Dg_{iX_j}] f_{iX_j}(a_1, x, a_2) d(a_1, x, a_2) \\ &\quad + \int [f_{iX_j} - g_{iX_j}]^2 h_{iX_j}(a_1, x, a_2) d(a_1, x, a_2), \end{aligned}$$

where  $Dg_{iX_j}$  is the functional derivative of  $g_{iX_j}$  with respect to  $f_{iX_j}$ , namely

$$Dg_{iX_j} = \left( \frac{h_{iX}}{f_{iX}} + \frac{h_{X_j}}{f_{X_j}} - \frac{h_X}{f_X} \right) g_{iX_j}.$$

As is apparent, imposing the null hypothesis induces singularity in the first functional derivative  $D\Lambda_f$ . To complete the proof, it then suffices to appreciate that, under the null, the second-order derivative reads

$$D^2\Lambda_f(h, h) = 2 \int [h_{iX_j}(a_1, x, a_2) - Dg_{iX_j}(a_1, x, a_2)]^2 dF_{iX_j}(a_1, x, a_2)$$

given that all other terms will depend on  $f_{iX_j} - g_{iX_j}$ . Observe, however, that

$Dg_{iX_j}$  converges at a faster rate than does  $h_{iX_j}$  due to its lower dimensionality.

The result then follows from a straightforward application of Lemma 1 with

$$\varphi(a_1, x, a_2) = f_{iX_j}(a_1, x, a_2).$$

**PROOF OF PROPOSITION 2:** The conditions imposed are such that the second-order functional Taylor expansion is also valid in the double array case  $(d_{i,n}, X_{i,n}, d_{j,n})$ . Thus, under  $H_1^{[n]}$  and assumptions A1 to A4,

$$\hat{\lambda}_n - \frac{b_n^{1/2}}{\hat{\sigma}_\Lambda} \sum_{k=1}^{n-i+j} [f_{iX_j}(d_{k+i-j,n}, X_{k+i-j,n}, d_{k,n}) - g_{iX_j}(d_{k+i-j,n}, X_{k+i-j,n}, d_{k,n})]^2$$

converges weakly to a standard normal distribution under  $f^{[n]}$ . The result then follows by noting that  $\hat{\sigma}_\Lambda \xrightarrow{p^{[n]}} \sigma_\Lambda$  and

$$\begin{aligned}\Lambda_{f^{[n]}} &= E \left[ f^{[n]}(d_{i,n}, X_{i,n}, d_{j,n}) - g^{[n]}(d_{i,n}, X_{i,n}, d_{j,n}) \right]^2 + O_p(n^{-1/2}) \\ &= n^{-1} b_n^{-1/2} \ell_2 + o_p(n^{-1} b_n^{-1/2}).\end{aligned}$$

## REFERENCES

- AÏT-SAHALIA, Y. (1994): “The delta method for nonparametric kernel functionals,” Graduate School of Business, University of Chicago.
- (1997): “Do interest rates really follow continuous-time Markov diffusions?,” Graduate School of Business, University of Chicago.
- AMARO DE MATOS, J., AND J. S. ROSÁRIO (2000): “The equilibrium dynamics for an endogeneous bid-ask spread in competitive financial markets,” Universidade Nova de Lisboa.
- BURGAYRAN, E., AND S. DAROLLES (1997): “Nonparametric estimation of a diffusion equation from tick observations,” CREST Working Paper 9756, INSEE.
- DAWID, A. P. (1979): “Conditional independence in statistical theory,” *Journal of the Royal Statistical Society B*, 41, 1–31.
- EASLEY, D., AND M. O’HARA (1987): “Price, trade size, and information in security markets,” *Journal of Financial Economics*, 19, 69–90.
- (1992): “Time and the process of security price adjustment,” *Journal of Finance*, 47, 577–605.
- FERNANDES, M., AND R. G. FLÔRES (1999): “Non-parametric tests for the Markov property,” European University Institute and Getúlio Vargas Foundation.
- FERNANDES, M., AND J. GRAMMIG (2000): “Non-parametric specification tests for conditional duration models,” Working Paper ECO 2000/4, European University Institute.

- GIOT, P. (2000): “Time transformations, intraday data and volatility models,” *Journal of Computational Finance*, 4, 31–62.
- GLOSTEN, L. R., AND P. R. MILGROM (1985): “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders,” *Journal of Financial Economics*, 14, 71–100.
- HALL, P. (1984): “Central limit theorem for integrated squared error multivariate nonparametric density estimators,” *Journal of Multivariate Analysis*, 14, 1–16.
- KHASHIMOV, S. A. (1992): “Limiting behaviour of generalized U-statistics of weakly dependent stationary processes,” *Theory of Probability and Its Applications*, 37, 148–150.
- LINTON, O., AND P. GOZALO (1999): “Conditional independence restrictions: Testing and estimation,” London School of Economics and Brown University.
- NEWHEY, W. K., AND K. D. WEST (1987): “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix,” *Econometrica*, 55, 703–708.
- PHILLIPS, P. C. B., AND P. PERRON (1988): “Testing for a unit root in time series regression,” *Biometrika*, 75, 335–346.
- SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.

**TABLE I**  
 PHILLIPS AND PERRON'S (1988) UNIT ROOT TESTS

	stock	sample size	truncation lag	test statistic
Boeing	ask	6,317	10	-1.6402
	bid	6,317	10	-1.6655
	spread	6,317	10	-115.3388
Coca-Cola	ask	3,823	8	-2.1555
	bid	3,823	8	-2.1615
	spread	3,823	8	-110.2846
Disney	ask	5,801	9	-1.2639
	bid	5,801	9	-1.2318
	spread	5,801	9	-112.1909
Exxon	ask	6,009	9	-0.6694
	bid	6,009	9	-0.6405
	spread	6,009	9	-121.8439
IBM	ask	15,124	12	-0.2177
	bid	15,124	12	-0.2124
	spread	15,124	12	-163.0558

Both ask and bid prices are in logs, whereas the spread refers to the difference of the logarithms of the ask and bid prices. The truncation lag  $\ell$  of the Newey and West's (1987) heteroskedasticity and autocorrelation consistent estimate of the spectrum at zero frequency is based on the automatic criterion  $\ell = [4(T/100)^{2/9}]$ , where  $[z]$  denotes the integer part of  $z$ .

**TABLE II**  
NONPARAMETRIC TESTS OF THE MARKOV PROPERTY

stock	duration		adjusted duration	
	$\hat{\lambda}_n$	p-value	$\hat{\lambda}_n$	p-value
Boeing	3.2935	(0.0005)	5.0327	(0.0000)
Coca-Cola	19.4971	(0.0000)	17.9364	(0.0000)
Disney	-2.5783	(0.9950)	-1.6367	(0.9491)
Exxon	-0.5484	(0.7083)	1.3966	(0.0813)
IBM	21.2064	(0.0000)	17.9247	(0.0000)

Adjusted durations refer to the correction for time-of-day effects.