Econometric Theory, **17**, 2001, 156–187. Printed in the United States of America.

TESTING FOR DISTRIBUTIONAL CHANGE IN TIME SERIES

ATSUSHI INOUE
North Carolina State University

This paper proposes nonparametric tests of change in the distribution function of a time series. The limiting null distributions of the test statistics depend on a nuisance parameter, and critical values cannot be tabulated a priori. To circumvent this problem, a new simulation-based statistical method is developed. The validity of our simulation procedure is established in terms of size, local power, and test consistency. The finite-sample properties of the proposed tests are evaluated in a set of Monte Carlo experiments, and the distributional stability in financial markets is examined.

1. INTRODUCTION

The purpose of this paper is to develop tests for distributional stability in a time series. Although financial markets have experienced significant episodes of instability such as the Great Depression, the end of the Bretton Woods System, the start of the European Monetary System, and the policy regime shifts by the Federal Reserve System, econometric theory typically assumes structural stability. Indeed, stability of distribution, moments, or parameters is essential to the proofs of asymptotic properties of the maximum likelihood method, generalized method of moments, and nonparametric method. Consequently, instability can affect estimation and inference.

Tests for structural change have already been developed in the econometric literature. Recent parametric and semiparametric tests include Andrews (1993) and Sowell (1996) for generalized method of moments, Bai (1994) for autoregressive moving average (ARMA) models, Bai (1996) and Chu, Stinchcombe, and White (1996) for linear regression models, and Chu (1995) for generalized autoregressive conditional heteroskedasticity (GARCH) models. Recent nonparametric tests include Picard (1985) for the spectral distribution of stationary Gaussian processes and Hidalgo (1995) for conditional moments.

I thank the co-editor, two anonymous referees, V. Corradi, F.X. Diebold, J. Hahn, O. Linton, R.S. Mariano, and the seminar participants at Finance Forum (Osaka, Japan), Kyoto University, Society for Computational Economics Third International Conference on Computing in Economics and Finance, Tokyo Metropolitan University, University of Pennsylvania, and University of Tokyo for helpful comments. Of course, any remaining errors are my own responsibility. Address correspondence to: Atsushi Inoue, Department of Agricultural and Resource Economics, Box 8109, North Carolina State University, Raleigh, NC 27695-8109, USA; e-mail: atsushi_inoue@ncsu.edu. The programs used in this paper and a technical appendix (Appendix B) are available at my homepage, http://www4.ncsu.edu/~atsushi.

 © 2001 Cambridge University Press 0266-4666/01 $9.50

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

In this paper, we propose nonparametric tests for change in the distribution function (d.f.) of a time series. The location of the change or changes is not specified a priori, as was done in the change point literature. There are three key features that distinguish our tests from the extant tests. First, our tests are nonparametric. The parametric and semiparametric tests should perform well under the correct specifications, but inference based on misspecified models is not well studied. On the other hand, nonparametric tests are robust against misspecification. Whereas it is possible to construct nonparametric tests based on density estimators, the convergence rate of density estimators is always slower than $\sqrt{n}$ and suffers from the curse of dimensionality. Thus, such tests may not have power against $\sqrt{n}$ local alternatives. In contrast, our tests are based on empirical d.f.'s. The convergence rate of empirical d.f.'s is always $\sqrt{n}$ and does not suffer from the curse of dimensionality. Our tests have nontrivial power against $\sqrt{n}$ local alternatives. The existing change point tests based on empirical d.f.'s include Bai (1994, 1996) for residual empirical d.f.'s of ARMA and linear regression models with independent disturbance and Szyszkowicz (1994a, 1994b) for independent observations. Carlstein (1988) develops change point estimators based on empirical d.f.'s for independent observations.

Second, our tests allow for dependence in the data. Thus, our tests are useful in time series environments. As a result of allowing for dependence, the limiting null distributions depend on a nuisance parameter. Thus, we cannot tabulate the critical values a priori. As the bootstrap has proven powerful when the limiting distributions depend on nuisance parameters (see, e.g., Andrews, 1997; Hansen, 1996; Linton and Gozalo, 1996), we bootstrap the Kolmogorov–Smirnov statistic and the Cramér–Von Mises statistic.

Third, our tests are robust against the heavy-tailed distributions observed in financial markets. The unconditional fourth moment of exchange rate returns and stock returns may not exist (see Jansen and de Vries, 1991; Phillips, McFarland, and McMahon, 1996), and the limiting distributions of the existing tests may depend on whether the unconditional fourth moment exists (see Loretan and Phillips, 1994). On the contrary, because our test statistics are constructed from empirical d.f.'s, the same asymptotic theory always applies in our framework. To the best of our knowledge, no extant test satisfies all three of our distinguishing key features.

The rest of the paper is organized as follows. Section 2 proposes nonparametric tests for change in the distribution function, develops asymptotic theory, and establishes the validity of the simulation procedure in terms of size, test consistency, and local power. Section 3 implements a set of Monte Carlo experiments to evaluate the finite-sample performance of the proposed tests. Section 4 applies the proposed testing procedure to analyzing the instability of financial markets. Section 5 concludes. Appendix A provides proofs of the main results given in the paper. A supplementary appendix, Appendix B, is available on my homepage, www4.ncsu.edu/~atsushi. In what follows, $I(\cdot)$ denotes the indicator function, $\Rightarrow$ denotes weak convergence in the space of

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

$D([0,1] \times \Re^p)$ under the Skorohod topology,[1] and $\xrightarrow{d}$ denotes convergence in distribution. All limits are taken as n goes to infinity, where n is the sample size.

2. ASYMPTOTIC THEORY

In this section, we propose a new simulation-based approach to testing for change in the d.f. First, we derive the limit process of a sequential empirical process and the limiting distributions of test statistics. Next, we derive the limit process of a simulated sequential empirical process and the limiting distributions of simulated test statistics. Our simulation approximates the limiting null distribution, and our simulation-based tests have nontrivial power against local alternatives. Last, we show the test consistency against the multiple-change alternative hypothesis.

Let $\{x_{ni} : i \leq n,\ n = 1,2,\ldots.\}$ be a triangular array of p-dimensional strong mixing random vectors defined on a complete probability space $(\Omega, \mathcal{A}, P)$.[2] In this paper, we are interested in testing the null hypothesis of no change in the d.f. of $\{x_{ni}\}$. That is, there is a d.f. F such that

$$P(x_{ni} \leq t) = F(t)$$

for all $t \in \Re^p$ and $i = 1,2,\ldots,n$ and $n \geq 1$. A test of change in the d.f. is based on a certain distance between the prechange d.f. and the postchange d.f. of $\{x_{ni}\}$.

Because we do not assume any knowledge of the distributional form and the change point, we measure the difference between the two d.f.'s by the difference between their sample analog:

$$\left| \frac{1}{m} \sum_{i=1}^{m} I(x_{ni} \leq t) - \frac{1}{n-m} \sum_{i=m+1}^{n} I(x_{ni} \leq t) \right|,$$

where m denotes a candidate change date. We consider the following two test statistics based on the preceding distance. One is a weighted Kolmogorov–Smirnov statistic,

$$T_1 = \sup_{1 \leq m < n} \sup_{t \in \Re^p} \left| \frac{m}{n} \left(1 - \frac{m}{n}\right) n^{1/2} \times \left(\frac{1}{m} \sum_{i=1}^{m} I(x_{ni} \leq t) - \frac{1}{n-m} \sum_{i=m+1}^{n} I(x_{ni} \leq t) \right) \right|,$$

and another is a weighted Cramér–Von Mises statistic,

$$T_2 = \frac{1}{n(n-1)} \sum_{m=1}^{n-1} \sum_{j=1}^{n} \left| \frac{m}{n} \left(1 - \frac{m}{n}\right) n^{1/2} \times \left(\frac{1}{m} \sum_{i=1}^{m} I(x_{ni} \leq x_{nj}) - \frac{1}{n-m} \sum_{i=m+1}^{n} I(x_{ni} \leq x_{nj}) \right) \right|^2.$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

The weighting function $(m/n)(1 - m/n)$ guarantees that the limiting distributions are well defined and has been used in the literature (see Deshayes and Picard, 1986; Carlstein, 1988; Bai, 1996).

To study the limiting distributions of T_1 and T_2, we introduce the following assumptions. First, we need to restrict the dependence of $\{x_{ni}\}$.

Assumption A. $\{x_{ni}\}$ is a strong mixing triangular array with mixing coefficients that satisfy

$$\sum_{j=1}^{\infty} j^2 \alpha(j)^{\gamma/(4+\gamma)} < \infty \quad \text{for some } \gamma \in (0,2).$$

Next, we need to define a sequence of local alternatives to analyze the local power. The following specification seems natural.

$$P(x_{ni} \le t) = (1 - \delta n^{-1/2}) P(\xi_i \le t) + \delta n^{-1/2} P\left(\eta_i\left(\frac{i}{n}\right) \le t\right), \tag{1}$$

where $\{\xi_i\}$ is a sequence of identically distributed random variables and $\{\eta_i(\cdot)\}$ is a sequence of random functions such that $\eta_i : \Omega \times [0,1] \to \Re^p$. Thus, $\{\eta_i(i/n)\}$ is a triangular array of p-dimensional random vectors. For instance, Bai (1996) uses a special case of (1). However, it turns out that the derivation of well-defined limiting distributions requires not only (1) but also the specification of $P(x_{ni} \le t, x_{n(i+d)} \le t') - P(x_{ni} \le t)P(x_{n(i+d)} \le t')$ in our context. That is, the covariance kernel of the well-defined limiting null distribution depends on the joint d.f.'s $\{P(x_{ni} \le t, x_{n(i+d)} \le t')\}$ and also on the marginals $\{P(x_{ni} \le t)\}$. More specifically, we require

$$\begin{aligned} &P(x_{ni} \le t, x_{n(i+d)} \le t') - P(x_{ni} \le t)P(x_{n(i+d)} \le t') \\ &\quad = (1 - \pi_n)(P(\xi_i \le t, \xi_{i+d} \le t') - P(\xi_i \le t)P(\xi_{i+d} \le t')) \\ &\qquad + \pi_n \left(P\left(\eta_i\left(\frac{i}{n}\right) \le t, \eta_{i+d}\left(\frac{i+d}{n}\right) \le t'\right) \right. \\ &\qquad\quad \left. - P\left(\eta_i\left(\frac{i}{n}\right) \le t\right) P\left(\eta_{i+d}\left(\frac{i+d}{n}\right) \le t'\right)\right) \end{aligned} \tag{2}$$

for some deterministic sequence $\{\pi_n\}_{n=1}^{\infty} \in [0,1]$ such that $\pi_n = o(1)$.

The next assumption accommodates both (1) and (2).

Assumption B. $P(x_{ni} \le t, x_{n(i+d)} \le t')$ is

$$\begin{aligned} H_{ni,d}(t,t') &= (1 - \delta n^{-1/2})^2 F_d(t,t') + \delta^2 n^{-1} G_d\left(\frac{i}{n}, \frac{i+d}{n}, t, t'\right) \\ &\quad + \delta n^{-1/2}(1 - \delta n^{-1/2}) \left(F(t) G\left(\frac{i+d}{n}, t'\right) + F(t') G\left(\frac{i}{n}, t\right)\right), \end{aligned}$$

$\delta \in [0,1]$,

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

where

(i) $F_d(\cdot,\cdot)$ and $F(\cdot)$ are the d.f.'s of some p-dimensional random vector ξ_i, i.e., $F_d(t,t') = P(\xi_i \le t, \xi_{i+d} \le t')$, $F(t) = P(\xi_i \le t)$, and $F(t)$ is continuous;
(ii) the marginal d.f. of the qth element of ξ_i is strictly increasing on $\Re$ for $q = 1,2,\ldots,p$;
(iii) $G_d(\cdot,\cdot,\cdot,\cdot)$ and $G(\cdot,\cdot)$ are the d.f.'s of some p-dimensional strong mixing random function $\eta_i(\cdot)$, i.e.,

$$G_d(r,r',t,t') = P(\eta_i(r) \le t, \eta_{i+d}(r') \le t'), \qquad G(r,t) = P(\eta_i(r) \le t).$$

The term $G(r,t)$ is continuous in t for all r, and there is (r,r',t) such that $G(r,t) \ne G(r',t)$.

Assumption B defines a sequence of local alternatives to the null hypothesis with $\delta = 0$ corresponding to the null hypothesis. Passing each element of t' to positive infinity, we obtain $P(x_{ni} \le t)$:

$$H_{ni}(t) = (1 - \delta n^{-1/2})F(t) + \delta n^{-1/2} G\left(\frac{i}{n}, t\right),$$

which is a mixture of two d.f.'s. The d.f. converges to the null as the sample size goes to infinity: $\lim_{n\to\infty} H_{ni}(t) = F(t)$. Assumption B includes a sequence of multiple-change local alternatives and a sequence of smooth-transition local alternatives because $G_d(\cdot,\cdot,\cdot,\cdot)$ can depend on the set of change points. To analyze the asymptotic behavior of our test statistics, we first consider a *sequential empirical process*:

$$K_n(r,t) = n^{-1/2} \sum_{i=1}^{[nr]} (I(x_{ni} \le t) - H_{ni}(t)),$$

defined on $[0,1] \times \Re^p$. The next theorem shows that the limiting distribution of $K_n(r,t)$ depends on a nuisance parameter:

$$\sigma(t,t') = \sum_{d=-\infty}^{\infty} (F_d(t,t') - F(t)F(t')).$$

THEOREM 2.1. *Under Assumptions* A *and* B,

$$K_n(\cdot,\cdot) \Rightarrow K(\cdot,\cdot),$$

where $K(\cdot,\cdot)$ *is a mean-zero Gaussian process with covariance kernel*

$$E(K(r,t)K(r',t')) = \min(r,r')\sigma(t,t')$$

and $P(K(\cdot,\cdot) \in C([0,1] \times \Re^p)) = 1$.

The limit process $K(\cdot,\cdot)$ is sometimes referred to in the literature as a Kiefer process. Yoshihara (1975) proves $K_n(\cdot,\cdot) \Rightarrow K(\cdot,\cdot)$ under the strict stationarity

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

assumption of $\{x_{ni}\}$ with $\delta = 0$. In contrast, Assumptions A and B require only marginal strict stationarity and are independent of the dimension p of x_{ni}.

We can rewrite the test statistics as

$$T_1 = \sup_{1\le m\le n-1} \sup_{t\in\Re^p} \left| d_n\left(\frac{m}{n}, t\right)\right|,$$

$$T_2 = \frac{1}{n(n-1)} \sum_{m=1}^{n-1} \sum_{j=1}^{n} d_n\left(\frac{m}{n}, x_{nj}\right)^2,$$

where

$$d_n\left(\frac{m}{n}, t\right) = n^{-1/2}\left(\sum_{i=1}^{m} I(x_{ni} \le t) - \frac{m}{n}\sum_{i=1}^{n} I(x_{ni} \le t)\right).$$

By applying the continuous mapping theorem (CMT) to Theorem 2.1, we obtain the limiting distributions.

COROLLARY 2.2.

(a) *Under the null hypothesis, i.e., Assumptions* A *and* B(i) *with* $\delta = 0$,

$$T_1 \xrightarrow{d} \sup_{0<r<1} \sup_{t\in\Re^p} |K^\circ(r,t)|,$$

$$T_2 \xrightarrow{d} \int_0^1 \int_{\Re^p} |K^\circ(r,t)|^2 dF(t)dr,$$

where $K^\circ(r,t)$ *is a mean-zero Gaussian process with covariance kernel*

$$E(K^\circ(r,t)K^\circ(r',t')) = (\min(r,r') - rr')\sigma(t,t').$$

(b) *Under the sequence of local alternatives, i.e., Assumptions* A *and* B,

$$T_1 \xrightarrow{d} \sup_{0<r<1} \sup_{t\in\Re^p} \left| K^\circ(r,t) + \delta\left(\int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds\right)\right|,$$

$$T_2 \xrightarrow{d} \int_0^1 \int_{\Re^p} \left| K^\circ(r,t) + \delta\left(\int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds\right)\right|^2 dF(t)dr,$$

where the set $\{(r,t): \int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds \ne 0\}$ *is of positive Lebesgue measure.*

When $\{x_{ni}\}$ is independent and identically distributed (i.i.d.), the limiting null distributions are free of nuisance parameters (see Deshayes and Picard, 1986), and the covariance kernel of $K^\circ(r,t)$ simplifies to $(\min(r,r') - rr')(F_0(t,t') - F(t)F(t'))$. When $\{x_{ni}\}$ is strong mixing, however, the limiting null distributions depend on a nuisance parameter, $\sigma(\cdot,\cdot)$. Thus, we cannot tabulate the critical values for the test statistics a priori. To circumvent the nuisance parameter problem, we develop a new simulation procedure. In what follows, we first simulate the sequential empirical process, and then we simulate the test statis-

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

tics (for the weak convergence of the bootstrap empirical process for i.i.d. random variables and for strictly stationary strong mixing random variables, respectively, see Bickel and Freedman, 1981; Bühlmann, 1994).

Given the sample ω, let $x_{ni}(\omega)$ denote a particular realization of x_{ni}. We define a simulation version of the sequential empirical process by

$$K_n^*(r,t;\omega) = n^{-1/2} \sum_{j=1}^{[nr]-\ell+1} z_j \sum_{i=j}^{j+\ell-1} \{I(x_{ni}(\omega) \le t) - \hat{F}_n(t;\omega)\},$$

where

$$\hat{F}_n(t;\omega) = \frac{1}{n} \sum_{i=1}^{n} I(x_{ni}(\omega) \le t).$$

Whereas the bootstrap of Efron (1979), Künsch (1989), and Bühlmann (1994) uses multinomial random variables as the random weighting $\{z_j\}$, our simulation method uses possibly continuous random variables as the random weighting $\{z_j\}$. Our simulation method can be considered a block version of some weighted bootstrap (see van der Vaart and Wellner, 1996, Example 3.6.12, p. 354). The block length ℓ plays the same role as that in the blockwise bootstrap of Künsch (1989) and Bühlmann (1994). Whereas Bühlmann (1994) shows the weak convergence of the block bootstrap empirical process for strictly stationary mixing random variables, his results cover neither the weak convergence of the sequential empirical process nor the weak convergence under a sequence of local alternatives. We develop a necessary asymptotic theory as follows.

Assumption C. $\{x_{ni}\}$ is a strong mixing triangular array that mixing coefficients satisfy

$$\sum_{j=1}^{\infty} j^{Q/2} \alpha(j)^{\gamma/(\gamma+Q)} < \infty$$

for some $\gamma > 0$ and even integer $Q \ge 16$.

Assumption D. $\{z_i\}_{i=1}^{n-\ell+1}$ are independent random variables independent of $\{x_{ni}\}$ with mean zero, variance $1/\ell$, and $E(z_i^4) = O(1/\ell^2)$, where $\ell \to \infty$ as $n \to \infty$ and $\ell = o(n^{1/2})$. In practice, we recommend the choice $z_i \sim NID(0,1/\ell)$. The reason is that the finite-sample distribution of $K^*(\cdot,\cdot;\omega)$ is Gaussian when $z_i \sim NID(0,1/\ell)$.

THEOREM 2.3. *Under Assumptions* B–D,

$$K_n^*(\cdot,\cdot;\omega) \Rightarrow K(\cdot,\cdot) \quad \omega\text{-a.s.}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

Let

$$d_n^*\left(\frac{m+\ell-1}{n}, t;\omega\right) = n^{-1/2}\left[\sum_{i=1}^{m} z_i \sum_{j=i}^{i+\ell-1}\{I(x_{nj}(\omega)\le t) - \hat{F}_n(t;\omega)\} - \frac{m}{n-\ell+1}\sum_{i=1}^{n-\ell+1} z_i \times \sum_{j=i}^{i+\ell-1}\{I(x_{nj}(\omega)\le t) - \hat{F}_n(t;\omega)\}\right]$$

for $m = 1,\ldots,n-\ell$. We define a simulation version of the Kolmogorov–Smirnov statistic by

$$T_1^*(\omega) = \sup_{1\le m\le n-\ell}\sup_{t\in\Re^p}\left|d_n^*\left(\frac{m+\ell-1}{n}, t;\omega\right)\right|$$

and a simulation version of the Cramér–Von Mises statistic by

$$T_2^*(\omega) = \frac{1}{n(n-\ell)}\sum_{m=1}^{n-\ell}\sum_{j=1}^{n}\left|d_n^*\left(\frac{m+\ell-1}{n}, x_{nj}(\omega);\omega\right)\right|^2.$$

By applying the CMT to Theorem 2.3, we obtain the limiting distributions of T_1^* and T_2^*.

COROLLARY 2.4. *Under Assumptions* B–D,

$$T_1^*(\omega) \xrightarrow{d} \sup_{0<r<1}\sup_{t\in\Re^p}|K^\circ(r,t)| \quad \omega\text{-a.s.},$$

$$T_2^*(\omega) \xrightarrow{d} \int_0^1\int_{\Re^p}|K^\circ(r,t)|^2 dF(t)dr \quad \omega\text{-a.s.}$$

Contrary to Corollary 2.2, the limiting distributions of the simulated test statistics do not depend on δ. Because of the additional randomness due to $\{z_i\}$ and the assumption $\ell = o(n^{1/2})$, the simulation version is less sensitive than the original test statistics. Thus, the $n^{-1/2}$ term in local alternatives does not affect the limiting distribution of the simulated test statistics. Corollaries 2.2 and 2.4 suggest the following simulation procedure.

(1) Compute T_1 and T_2.
(2) Generate $\{T_1^{*(j)}\}_{j=1}^J$ and $\{T_2^{*(j)}\}_{j=1}^J$, where $T_1^{*(j)}$ and $T_2^{*(j)}$ are based on $\{z_i^{(j)}\}_{i=1}^{n-\ell+1}$ for $j = 1,2,\ldots,J$.
(3) Estimate the level-α critical value $\hat{c}_{1\alpha n}^J$ and $\hat{c}_{2\alpha n}^J$ from $\{T_1^{*(j)}\}_{j=1}^J$ and $\{T_2^{*(j)}\}_{j=1}^J$, respectively.

By the Glivenko–Cantelli lemma, $\lim_{J\to\infty}\hat{c}_{1\alpha n}^J = c_{1\alpha n}$ ω-a.s. and $\lim_{J\to\infty}\hat{c}_{2\alpha n}^J = c_{2\alpha n}$ ω-a.s., where $c_{1\alpha n}$ and $c_{2\alpha n}$ are the level-α critical values of T_1^* and T_2^* conditional on ω and n, respectively. By the continuity of $K^\circ(\cdot,\cdot)$ (see Theo-

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

rem 2.1) and Corollary 2.4, $\lim_{n\to\infty} c_{1\alpha n} = c_{1\alpha}$ ω-a.s. and $\lim_{n\to\infty} c_{2\alpha n} = c_{2\alpha}$ ω-a.s., where $c_{1\alpha n}$ and $c_{2\alpha n}$ are the critical values of the limiting distributions of T_1 and T_2, respectively. Thus, Corollaries 2.2(a) and 2.4 imply

$$\lim_{n\to\infty} \lim_{J\to\infty} P(T_1 > c_{1\alpha n}^J) = \alpha, \qquad \lim_{n\to\infty} \lim_{J\to\infty} P(T_2 > c_{2\alpha n}^J) = \alpha.$$

Because $K^\circ(\cdot,\cdot)$ is a zero-mean Gaussian process, the quantiles of $|K^\circ(r,t) + c|$ are always larger than the corresponding quantiles of $|K^\circ(r,t)|$ for every $c \neq 0$ and $(r,t) \in (0,1) \times \Re^p$. Thus, Corollaries 2.2(b) and 2.4 imply that, when $\delta > 0$,

$$\lim_{n\to\infty} \lim_{J\to\infty} P(T_1 > c_{1\alpha n}^J) \geq \alpha, \qquad \lim_{n\to\infty} \lim_{J\to\infty} P(T_2 > c_{2\alpha n}^J) > \alpha.$$

Therefore, the proposed tests have nontrivial power against local alternatives.

From this point on, we focus on the test consistency against a sequence of multiple-change global alternatives, which is defined as follows.

Assumption E. The d.f. of x_{ni} is

$$H_{ni}(t) = F_k(t), \quad [n\pi_{k-1}] < i \leq [n\pi_k]$$

for $k = 1,\ldots,K+1$, where $0 = \pi_0 < \pi_1 < \cdots < \pi_K < \pi_{K+1} = 1$.

(i) $F_k(\cdot)$ is continuous for all k, and there are $k \neq k'$ and t such that $F_k(t) \neq F_{k'}(t)$.
(ii)

$$\frac{1}{n} \sum_{i=[n\pi_{k-1}]+1}^{[n\pi_k]} (I(x_{ni} \leq t) - F_k(t)) = o_{as}(1)$$

uniform in $t \in \Re^p$ for $k = 1,\ldots,K+1$.

Assumption E means that the d.f. of x_{ni} changes K times. Assumption E(ii) states that the Glivenko–Cantelli theorem holds in each subsample. Imposing more primitive assumptions, such as Assumptions B and C, in each subsample will imply Assumption E(ii).

THEOREM 2.5. *Under the sequence of multiple-change global alternatives, i.e., Assumptions* D *and* E,

$$\lim_{n\to\infty} n^{-1/2} T_1 > 0 \quad \omega\text{-a.s.}, \qquad \lim_{n\to\infty} n^{-1} T_2 > 0 \quad \omega\text{-a.s.},$$

$$\ell^{-1/2} T_1^* = O_p(1), \qquad \ell^{-1} T_2^* = O_p(1).$$

Because $n^{-1/2}T_1$ and $n^{-1}T_2$ converge to positive constants with probability one, T_1 and T_2 almost surely diverge to positive infinity at rate $n^{1/2}$ and n, respectively. On the other hand, $\ell^{-1/2}T_1^*$ and $\ell^{-1}T_2$ are bounded in probability, and so the simulated statistics diverge in probability at rate $\ell^{1/2}$ and ℓ, respectively. Thus, the proposed tests are consistent because the test statistics diverge faster than the corresponding simulated test statistics by Assumption D. How-

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

ever, a direct consequence of diverging critical values under the alternative is the loss of power in finite samples. The proposed tests would be less powerful than ideal tests in which the fixed critical values of the limiting null distributions are available.

In concluding this section, we mention another potential approach. This approach would take the following steps. First, estimate the covariance kernel by the heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimation methods. Second, generate multivariate normal random variables based on the Cholesky decomposition of the estimated covariance matrix (for a related bootstrap procedure in the context of spectral density estimation, see Diebold, Ohanian, and Berkowitz, 1998). When the Bartlett kernel is used in the HAC estimation, the present simulation approach and this simulation approach are equivalent.[3] The advantage of our simulation approach over this approach is that our simulation method can directly generate random variables without estimating the $n \times n$ covariance matrix and computing the Cholesky decomposition. To the best of our knowledge, the standard statistical packages cannot handle $n \times n$ matrices when n is huge, say, 5,000.

3. MONTE CARLO SIMULATIONS

In this section, we evaluate the finite-sample performance of the simulation procedure in a set of Monte Carlo experiments. We use the following stochastic volatility model as the data generating process (DGP):

$$x_{ni} = \exp\left(\frac{h_{ni}}{2}\right)\varepsilon_i,$$

$$h_{ni} = -0.5 + \Delta I(i \geq [n\pi]) + \rho h_{n(i-1)} + \sqrt{0.30}\eta_i,$$

$$\begin{pmatrix} \varepsilon_i \\ \eta_i \end{pmatrix} \sim N(0_{2\times 1}, I_2),$$

with $\Delta = -0.4, -0.2, 0.0, 0.2, 0.4$, $\rho = 0.90, 0.95$, $n = 250, 500, 1{,}000, 1{,}500$, $\pi = 0.25, 0.50, 0.75$. The parameter values roughly correspond to the parameter estimates for the weekly financial time series studied in Section 4 (see Jacquier, Polson, and Rossi, 1994, Tables 1 and 2). We draw h_0 from the implied unconditional distribution and discard the first 1,000 observations to minimize the effects of the initial values.

The choice of the block length ℓ is an important question in practice. In finite samples, the test results depend on the block length used because simulated critical values depend on the block length. Although there exist data-dependent methods to select the block length, such as Hall, Horowitz, and Jing (1995) and Bühlmann and Künsch (1996), they may not work in our context. Because the test statistics diverge to positive infinity under the alternative hypothesis, the block length selected by the data-dependent method may not satisfy $\ell = o(n^{1/2})$. Therefore, we simply use $\ell = 10, 20, \ldots, 100$ for the size

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

analysis and $n = 10, 100$ for the power analysis. As a result of the computational requirement, we set the number of Monte Carlo replications at 2,000 for $n = 250$, 1,000 for $n = 500$, 500 for $n = 1{,}000$ and 250 for $n = 1{,}500$, and the number of simulation replications at 199 throughout the experiments. We use $z_i \sim N(0,1/\ell)$. Table 1 gives the implied unconditional variances.

Tables 2 and 3 report the rejection frequencies at the 1%, 5%, and 10% levels of significance under the null hypothesis. The two tables suggest that the rejection frequencies are decreasing in ℓ and increasing in ρ. Thus, stronger dependence requires larger block lengths to obtain good size performance.

Tables 4 and 5 report the rejection frequencies at the 5% significance level under the alternative hypothesis. Whereas the shift does not change the unconditional mean, it changes the unconditional variance.

As expected, the rejection frequencies increase in the sample size, increase in the break size, and depend on break location. These results apply not only to our tests but also to other tests for structural change. The sensitivity to the choice of the block length is decreasing in the sample size in the following sense. For given Δ and π, the ratio

$$\frac{\text{the actual power for } \ell = 10}{\text{the actual power for } \ell = 100}$$

is decreasing in n.

The power is asymmetric with respect to the break locations. That is, the power when $\pi = 0.25$ tends to be higher than the power when $\pi = 0.75$. Although the rejection probability should be close to one regardless of the break location in large samples, the break location may affect the power in small samples for the following reason. Immediately after a shift in the parameter, there is some adjustment period during which the effect of the prechange model remains and the empirical d.f.'s cannot be accurately estimated. When $\pi = 0.75$, the reliable sample size for the second empirical d.f. may be relatively small. The degree of the asymmetry decreases as the sample size increases, however.

4. APPLICATIONS

In this section, we apply our simulation approach to testing for change in financial markets. We use the weekly log-difference of the spot exchange rates (middle rate) for the £/$, DM/$, franc/$, and yen/$ and the weekly log-

TABLE 1. The implied unconditional variances

	$\Delta = -0.2$	$\Delta = -0.1$	$\Delta = 0.0$	$\Delta = 0.1$	$\Delta = 0.2$
$\rho = 0.90$	0.0020	0.0055	0.0148	0.0403	0.1096
$\rho = 0.95$	3.9×10^{-6}	2.9×10^{-5}	0.0002	0.0016	0.0115

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

TABLE 2. Finite-sample rejection frequencies under the null ($\rho = 0.90$)

Sample size n	Block size ℓ	Kolmogorov–Smirnov			Cramér–Von Mises		
		0.01	0.05	0.10	0.01	0.05	0.10
250	10	0.013	0.105	0.194	0.019	0.119	0.234
	20	0.001	0.026	0.087	0.003	0.055	0.140
	30	0.000	0.006	0.041	0.002	0.030	0.109
	40	0.000	0.002	0.025	0.001	0.022	0.091
	50	0.000	0.004	0.024	0.000	0.020	0.098
	60	0.000	0.005	0.024	0.002	0.027	0.112
	70	0.000	0.008	0.032	0.002	0.042	0.141
	80	0.000	0.013	0.053	0.007	0.061	0.192
	90	0.002	0.026	0.089	0.012	0.106	0.252
	100	0.004	0.044	0.128	0.022	0.154	0.324
500	10	0.032	0.126	0.215	0.037	0.130	0.242
	20	0.008	0.058	0.128	0.011	0.081	0.157
	30	0.003	0.031	0.080	0.006	0.051	0.122
	40	0.002	0.019	0.054	0.004	0.034	0.104
	50	0.000	0.010	0.041	0.002	0.026	0.101
	60	0.000	0.010	0.033	0.000	0.021	0.091
	70	0.000	0.005	0.024	0.000	0.021	0.088
	80	0.000	0.005	0.022	0.000	0.023	0.076
	90	0.000	0.003	0.020	0.001	0.016	0.086
	100	0.000	0.004	0.018	0.000	0.016	0.083
1,000	10	0.032	0.130	0.220	0.036	0.140	0.250
	20	0.014	0.062	0.126	0.014	0.094	0.170
	30	0.008	0.046	0.114	0.012	0.058	0.138
	40	0.002	0.028	0.068	0.008	0.054	0.126
	50	0.000	0.018	0.068	0.008	0.044	0.112
	60	0.004	0.010	0.050	0.004	0.036	0.098
	70	0.000	0.012	0.040	0.006	0.034	0.084
	80	0.002	0.010	0.032	0.002	0.028	0.078
	90	0.000	0.004	0.028	0.004	0.032	0.086
	100	0.000	0.006	0.020	0.000	0.030	0.076
1,500	10	0.048	0.164	0.280	0.056	0.172	0.292
	20	0.016	0.092	0.216	0.020	0.116	0.204
	30	0.012	0.088	0.160	0.016	0.072	0.172
	40	0.012	0.060	0.124	0.012	0.060	0.140
	50	0.012	0.040	0.112	0.012	0.048	0.128
	60	0.008	0.044	0.104	0.008	0.060	0.112
	70	0.008	0.032	0.088	0.008	0.048	0.104
	80	0.004	0.024	0.088	0.008	0.036	0.088
	90	0.004	0.020	0.064	0.008	0.032	0.092
	100	0.004	0.020	0.068	0.004	0.028	0.088

Notes: The DGP is $x_{ni} = \exp(h_{ni}/2)\varepsilon_i$, $h_{ni} = -0.5 + 0.90h_{n(i-1)} + \sqrt{0.30}\eta_i$, $(\varepsilon_i, \eta_i)' \sim N(0_{2\times 1}, I_2)$. The number of bootstrap replications is set to 199, and the number of simulation replications is set to 2,000 for $n = 200$, 1,000 for $n = 500$, 500 for $n = 1{,}000$, and 250 for $n = 1{,}500$.

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

TABLE 3. Finite-sample rejection frequencies under the null ($\rho = 0.95$)

Sample size n	Block size ℓ	Kolmogorov–Smirnov			Cramér–Von Mises		
		0.01	0.05	0.10	0.01	0.05	0.10
250	10	0.070	0.238	0.378	0.096	0.290	0.440
	20	0.003	0.068	0.172	0.015	0.148	0.274
	30	0.000	0.014	0.068	0.002	0.062	0.191
	40	0.000	0.005	0.029	0.001	0.033	0.134
	50	0.000	0.003	0.021	0.001	0.027	0.116
	60	0.000	0.003	0.019	0.003	0.030	0.125
	70	0.000	0.008	0.032	0.002	0.043	0.158
	80	0.000	0.015	0.057	0.005	0.071	0.203
	90	0.002	0.032	0.097	0.016	0.110	0.250
	100	0.007	0.053	0.126	0.037	0.155	0.319
500	10	0.144	0.339	0.479	0.173	0.367	0.475
	20	0.035	0.157	0.270	0.052	0.209	0.336
	30	0.007	0.072	0.158	0.021	0.127	0.240
	40	0.003	0.032	0.102	0.008	0.078	0.186
	50	0.000	0.015	0.068	0.001	0.057	0.157
	60	0.000	0.009	0.041	0.000	0.040	0.144
	70	0.000	0.005	0.033	0.000	0.036	0.130
	80	0.000	0.003	0.028	0.000	0.024	0.119
	90	0.000	0.001	0.025	0.001	0.031	0.117
	100	0.000	0.001	0.026	0.002	0.026	0.117
1,000	10	0.182	0.364	0.512	0.184	0.380	0.514
	20	0.054	0.208	0.332	0.078	0.222	0.352
	30	0.022	0.134	0.222	0.052	0.160	0.270
	40	0.008	0.066	0.156	0.016	0.112	0.202
	50	0.002	0.044	0.120	0.014	0.094	0.184
	60	0.002	0.028	0.080	0.010	0.074	0.152
	70	0.000	0.024	0.068	0.012	0.058	0.138
	80	0.002	0.014	0.050	0.006	0.050	0.120
	90	0.000	0.016	0.044	0.010	0.046	0.120
	100	0.000	0.006	0.036	0.002	0.040	0.112
1,500	10	0.248	0.460	0.588	0.272	0.452	0.576
	20	0.120	0.288	0.412	0.104	0.292	0.432
	30	0.064	0.176	0.276	0.052	0.168	0.324
	40	0.028	0.120	0.212	0.032	0.124	0.276
	50	0.020	0.100	0.168	0.020	0.120	0.204
	60	0.016	0.076	0.144	0.020	0.092	0.148
	70	0.012	0.068	0.120	0.012	0.068	0.144
	80	0.008	0.052	0.108	0.012	0.068	0.136
	90	0.004	0.052	0.096	0.012	0.060	0.136
	100	0.004	0.040	0.092	0.012	0.048	0.120

Notes: The DGP is $x_{ni} = \exp(h_{ni}/2)\varepsilon_i$, $h_{ni} = -0.5 + 0.95h_{n(i-1)} + \sqrt{0.30}\eta_i$, $(\varepsilon_i, \eta_i)' \sim N(0_{2\times 1}, I_2)$. The number of bootstrap replications is set to 199, and the number of simulation replications is set to 2,000 for $n = 200$, 1,000 for $n = 500$, 500 for $n = 1{,}000$, and 250 for $n = 1{,}500$.

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

TABLE 4. Finite-sample rejection frequencies under the alternative ($\rho = 0.90$)

Sample size n	Break ratio π	Break size Δ	Kolmogorov–Smirnov		Cramér–Von Mises	
			$\ell = 10$	$\ell = 100$	$\ell = 10$	$\ell = 100$
250	0.25	+0.2	0.431	0.078	0.573	0.299
		+0.1	0.173	0.060	0.228	0.199
		−0.1	0.185	0.056	0.237	0.184
		−0.2	0.449	0.089	0.546	0.292
	0.50	+0.2	0.628	0.016	0.739	0.075
		+0.1	0.234	0.036	0.303	0.127
		−0.1	0.254	0.032	0.332	0.120
		−0.2	0.670	0.013	0.780	0.068
	0.75	+0.2	0.264	0.112	0.364	0.357
		+0.1	0.135	0.065	0.178	0.217
		−0.1	0.138	0.059	0.196	0.184
		−0.2	0.289	0.097	0.416	0.298
500	0.25	+0.2	0.787	0.002	0.868	0.053
		+0.1	0.295	0.005	0.398	0.039
		−0.1	0.352	0.006	0.393	0.045
		−0.2	0.787	0.005	0.830	0.058
	0.50	+0.2	0.939	0.011	0.956	0.155
		+0.1	0.486	0.004	0.543	0.061
		−0.1	0.510	0.016	0.564	0.072
		−0.2	0.941	0.012	0.961	0.157
	0.75	+0.2	0.706	0.008	0.741	0.044
		+0.1	0.288	0.006	0.350	0.035
		−0.1	0.292	0.004	0.363	0.034
		−0.2	0.726	0.003	0.800	0.058
1,000	0.25	+0.2	0.974	0.058	0.980	0.338
		+0.1	0.526	0.034	0.608	0.132
		−0.1	0.572	0.018	0.610	0.118
		−0.2	0.988	0.064	0.986	0.366
	0.50	+0.2	1.000	0.630	1.000	0.930
		+0.1	0.756	0.130	0.804	0.340
		−0.1	0.768	0.126	0.798	0.420
		−0.2	0.998	0.630	0.998	0.944
	0.75	+0.2	0.956	0.056	0.970	0.256
		+0.1	0.564	0.028	0.548	0.112
		−0.1	0.536	0.010	0.608	0.106
		−0.2	0.970	0.022	0.976	0.274

(*continued*)

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

TABLE 4. *Continued*

Sample size n	Break ratio π	Break size Δ	Kolmogorov–Smirnov $\ell = 10$	Kolmogorov–Smirnov $\ell = 100$	Cramér–Von Mises $\ell = 10$	Cramér–Von Mises $\ell = 100$
1,500	0.25	+0.2	0.996	0.516	0.996	0.800
		+0.1	0.712	0.172	0.772	0.356
		−0.1	0.708	0.132	0.720	0.320
		−0.2	1.000	0.432	1.000	0.744
	0.50	+0.2	1.000	0.988	1.000	1.000
		+0.1	0.900	0.508	0.924	0.708
		−0.1	0.884	0.424	0.900	0.668
		−0.2	1.000	0.976	1.000	0.996
	0.75	+0.2	1.000	0.368	1.000	0.732
		+0.1	0.752	0.124	0.744	0.304
		−0.1	0.664	0.124	0.732	0.308
		−0.2	0.996	0.368	1.000	0.708

Notes: The DGP is $x_{ni} = \exp(h_{ni}/2)\varepsilon_i$, $h_{ni} = -0.5 + \Delta I(i \geq [n\pi]) + 0.90h_{n(i-1)} + \sqrt{0.30}\eta_i$, $(\varepsilon_i, \eta_i)' \sim N(0_{2\times 1}, I_2)$. The number of bootstrap replications is set to 199, and the number of simulation replications is set to 2,000 for $n = 200$, 1,000 for $n = 500$, 500 for $n = 1{,}000$, and 250 for $n = 1{,}500$.

difference of the Dow Jones industrial average, Standard & Poor's 500 composite index, and NYSE composite index. To minimize the number of missing values, we use the Wednesday returns. We retrieve the data from the DATASTREAM database. Table 6 summarizes the data.

The choice of the block length is important in practice. We select the block lengths by response surface as follows. Let ℓ_n^* be the block length that minimizes the sum of the absolute deviations from the nominal 1%, 5%, and 10% significance levels in Tables 2 and 3. We regress $\log \ell_n^*$ on one and $\log n$ with the constraint that the slope coefficient is less than or equal to 0.4999. The constraint least square estimates for the Kolmogorov–Smirnov statistic are

$$\log \ell_1^* = 0.134 + 0.499 \log n \quad \textbf{(1)}$$

for $\rho = 0.90$ and

$$\log \ell_2^* = 0.623 + 0.499 \log n \quad \textbf{(2)}$$

for $\rho = 0.95$. The constraint least square estimates for the Cramér–Von Mises statistic are

$$\log \ell_3^* = 0.916 + 0.446 \log n \quad \textbf{(3)}$$

for $\rho = 0.90$ and

$$\log \ell_4^* = 2.444 + 0.292 \log n \quad \textbf{(4)}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

TABLE 5. Finite-sample rejection frequencies under the alternative ($\rho = 0.95$)

Sample size n	Break ratio π	Break size Δ	Kolmogorov–Smirnov		Cramér–Von Mises	
			$\ell = 10$	$\ell = 100$	$\ell = 10$	$\ell = 100$
250	0.25	+0.2	0.817	0.060	0.900	0.250
		+0.1	0.463	0.062	0.538	0.201
		−0.1	0.468	0.060	0.544	0.219
		−0.2	0.836	0.067	0.890	0.252
	0.50	+0.2	0.902	0.016	0.942	0.069
		+0.1	0.525	0.036	0.587	0.132
		−0.1	0.543	0.041	0.623	0.126
		−0.2	0.911	0.011	0.946	0.065
	0.75	+0.2	0.489	0.106	0.577	0.367
		+0.1	0.312	0.063	0.376	0.220
		−0.1	0.315	0.069	0.401	0.225
		−0.2	0.523	0.094	0.642	0.354
500	0.25	+0.2	0.977	0.003	0.983	0.060
		+0.1	0.674	0.009	0.743	0.047
		−0.1	0.695	0.004	0.719	0.066
		−0.2	0.978	0.003	0.981	0.064
	0.50	+0.2	0.990	0.008	0.991	0.189
		+0.1	0.785	0.010	0.808	0.097
		−0.1	0.799	0.012	0.832	0.094
		−0.2	0.992	0.004	0.997	0.165
	0.75	+0.2	0.921	0.005	0.930	0.063
		+0.1	0.584	0.011	0.615	0.057
		−0.1	0.593	0.003	0.645	0.049
		−0.2	0.928	0.004	0.942	0.065
1,000	0.25	0.2	0.998	0.094	1.000	0.462
		0.1	0.880	0.054	0.904	0.228
		−0.1	0.902	0.066	0.916	0.258
		−0.2	1.000	0.124	1.000	0.502
	0.50	0.2	1.000	0.798	1.000	0.980
		0.1	0.960	0.202	0.956	0.502
		−0.1	0.948	0.262	0.960	0.578
		−0.2	1.000	0.814	1.000	0.982
	0.75	0.2	0.998	0.038	0.998	0.284
		0.1	0.846	0.038	0.844	0.168
		−0.1	0.836	0.040	0.856	0.168
		−0.2	0.998	0.034	0.998	0.350

(*continued*)

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

TABLE 5. *Continued*

Sample size n	Break ratio π	Break size Δ	Kolmogorov–Smirnov $\ell = 10$	Kolmogorov–Smirnov $\ell = 100$	Cramér–Von Mises $\ell = 10$	Cramér–Von Mises $\ell = 100$
1,500	0.25	0.2	1.000	0.748	1.000	0.920
		0.1	0.944	0.276	0.952	0.480
		−0.1	0.936	0.244	0.940	0.476
		−0.2	1.000	0.656	1.000	0.868
	0.50	0.2	1.000	1.000	1.000	1.000
		0.1	0.996	0.632	0.996	0.848
		−0.1	0.984	0.604	0.980	0.796
		−0.2	1.000	0.996	1.000	1.000
	0.75	0.2	1.000	0.508	1.000	0.792
		0.1	0.932	0.164	0.940	0.424
		−0.1	0.908	0.176	0.912	0.364
		−0.2	1.000	0.476	1.000	0.772

Notes: The DGP is $x_{ni} = \exp(h_{ni}/2)\varepsilon_i$, $h_{ni} = -0.5 + \Delta I(i \geq [n\pi]) + 0.95h_{n(i-1)} + \sqrt{0.30}\eta_i$, $(\varepsilon_i, \eta_i)' \sim N(0_{2\times 1}, I_2)$. The number of bootstrap replications is set to 199, and the number of simulation replications is set to 2,000 for $n = 200$, 1,000 for $n = 500$, 500 for $n = 1{,}000$, and 250 for $n = 1{,}500$.

for $\rho = 0.95$. The constraint is binding for the Kolmogorov–Smirnov statistic, and the unconstrained least square estimates for the slope coefficient are larger than 0.5.

We use $z_i \sim N(0,1/\ell)$ and set the number of simulation replications at 999. Table 7 reports the simulated p-values for the marginal d.f. of x_{ni} $(p = 1)$ and for the joint d.f. of $(x_{ni}, x_{n(i-1)})$ $(p = 2)$. The first seven rows show the results

TABLE 6. Data

	Datastream code	Sample period	Sample size
British pound	USDOLLR	1973–1996	1,252
Deutsche mark	DMARKER, USDOLLR	1973–1996	1,252
French franc	FRENFRA, USDOLLR	1973–1996	1,252
Japanese yen	JAPAYEN, USDOLLR	1973–1996	1,252
Dow Jones industrial average	DJINDUS	1969–1996	1,461
NYSE composite	NYSEALL	1976–1996	1,095
S&P 500 composite	S&PCOMP	1969–1996	1,461

Notes: The weekly (Wednesday) log-difference is used. The exchange rates against the British pound are transformed into those against the U.S. dollar.

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

TABLE 7. Applications

	$p = 1$					$p = 2$	
	Kolmogorov–Smirnov			Cramér–Von Mises			
	ℓ_1^*	ℓ_2^*		ℓ_3^*	ℓ_4^*	ℓ_3^*	ℓ_4^*
British pound	0.060	0.122	(1985)	0.020	0.061	0.068	0.102
Deutsche mark	0.041	0.090	(1980)	0.101	0.157	0.154	0.180
French franc	0.053	0.135	(1980)	0.070	0.122	0.140	0.138
Japanese yen	0.042	0.100	(1978)	0.091	0.156	0.111	0.163
Dow Jones	0.001	0.015	(1984)	0.011	0.018	0.008	0.015
NYSE	0.120	0.227	(1984)	0.086	0.119	0.075	0.107
S&P500	0.006	0.040	(1984)	0.028	0.058	0.013	0.028
British pound^2	0.006	0.026	(1980)	0.021	0.043	0.023	0.059
Deutsche mark^2	0.009	0.036	(1980)	0.043	0.102	0.041	0.083
French franc^2	0.010	0.034	(1980)	0.055	0.094	0.044	0.107
Japanese yen^2	0.007	0.038	(1977)	0.035	0.088	0.036	0.094
Dow Jones^2	0.014	0.078	(1988)	0.054	0.092	0.037	0.095
NYSE^2	0.024	0.083	(1988)	0.062	0.106	0.044	0.091
S&P500^2	0.054	0.117	(1988)	0.057	0.117	0.053	0.101

Notes: $p = 1$ corresponds to the marginal distribution function of x_{ni}, and $p = 2$ corresponds to the joint distribution function of $(x_{ni}, x_{n(i-1)})$, where x_{ni} is a random variable in question. The block lengths used are $\ell_1^* = \exp(0.134 + 0.499 \log n)$, $\ell_2^* = \exp(0.623 + 0.499 \log n)$, $\ell_3^* = \exp(0.916 + 0.446 \log n)$, and $\ell_4^* = \exp(2.444 + 0.292 \log n)$. Superscript 2 indicates squared returns. The number of bootstrap replications is set to 999. In parentheses are the maximizing years of the Kolmogorov–Smirnov statistics.

for returns, and the last seven rows show the results for squared returns. Change in the d.f. of returns implies change in the d.f. of squared returns; however, we investigate both d.f.'s because the second moment of returns is often of interest. The joint d.f. contains information about the dynamics, which the marginal d.f. may not. Whereas we do not find very strong evidence for structural change in the spot exchange rates, we find strong evidence of structural change in the stock returns. In contrast, we find some evidence for change in the d.f.'s of squared returns of the spot exchange rates. The results show no qualitative difference between the marginal d.f. and joint d.f.

In parentheses are the maximizing years of the Kolmogorov–Smirnov distances. The break years for stock returns are concentrated in the middle of the 1980's. A possible interpretation is that the trading strategies and technologies in stock markets changed during the 1980's. For example, the change in programming trading after the October 1987 crash may explain our finding of structural change in stock markets.[4] Our finding also raises questions about the traditional financial modeling that assumes the stability of distributions.

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

5. CONCLUDING REMARKS

Our tests do not substitute for the existing parametric tests but rather complement them. Under the correct specification, parametric tests should be more powerful and informative than nonparametric tests. When the correct specification is unknown, however, our nonparametric tests will serve as a useful pretest.

The choice of the block length is important in practice. Although the data-dependent methods of Hall et al. (1995) and of Bühlmann and Künsch (1996) seem promising, the asymptotic behavior of the selected block length under the alternative hypothesis must be analyzed. We leave it for future research.

We conjecture that it is possible to extend our results for the residual empirical d.f.'s along the lines of Bai (1994, 1996). For example, one can consider tests for structural change based on the empirical d.f.'s of the residuals when the disturbances are strong mixing.

NOTES

1. $D([0,1] \times \Re^p)$ is the space of functions on $[0,1] \times \Re^p$ that are right continuous and have left limits. The Skorohod topology for $D([0,1] \times \Re^p)$ is defined as in Billingsley (1968, pp. 111–114) except that Λ is replaced by the class of strictly increasing, continuous mappings of $[0,1] \times \Re^p$ onto itself.

2. Under the null hypothesis of no change, $\{x_{ni}\}$ is a single array. However, we use triangular arrays to analyze the asymptotic behavior under local alternatives.

3. See Appendix B.5 for a brief discussion.

4. We owe this interpretation to J. Qian.

REFERENCES

Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61, 821–856.

Andrews, D.W.K. (1997) A conditional Kolmogorov test. *Econometrica* 64, 1097–1128.

Andrews, D.W.K. & D. Pollard (1994) An introduction to functional central limit theorem for dependent stochastic processes. *International Statistical Review* 62, 119–132.

Bai, J. (1994) Weak convergence of the sequential empirical processes of residuals in ARMA models. *Annals of Statistics* 22, 2051–2061.

Bai, J. (1996) Testing for parameter constancy in linear regressions: An empirical distribution function approach. *Econometrica* 64, 597–622.

Bickel, P.J. & D.A. Friedman (1981) Some asymptotic theory for the bootstrap. *Annals of Statistics* 9, 1196–1217.

Billingsley, P. (1968) *Convergence of Probability Measures*. New York: Wiley.

Bühlmann, P. (1994) Blockwise bootstrapped empirical process for stationary sequences. *Annals of Statistics* 22, 995–1012.

Bühlmann, P. & H.R. Künsch (1996) Block Length Selection in the Bootstrap for Time Series. Manuscript.

Carlstein, E. (1988) Nonparametric change-point estimation. *Annals of Statistics* 16, 188–197.

Chu, C.-S. J. (1995) Detecting parameter shift in GARCH models. *Econometric Reviews* 14, 241–266.

Chu, C.-S.J., M. Stinchcombe, & H. White (1996) Monitoring structural change. *Econometrica* 64, 1045–1065.

Davidson, J. (1992) A central limit theorem for globally nonstationary near-epoch dependent functions of mixing processes. *Econometric Theory* 8, 313–329.

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

Deshayes, J. & D. Picard (1986) Off-line statistical analysis of change point models using non parametric and likelihood methods. In M. Basseville & A. Benveniste (eds.), *Detection of Abrupt Changes in Signals and Dynamical Systems*, Lecture Notes in Control and Information Sciences 77, pp. 103–168. Berlin: Springer-Verlag.

Diebold, F.X., L.E. Ohanian, & J. Berkowitz (1998) Dynamic equilibrium economies: A framework for comparing models and data. *Review of Economic Studies* 65, 433–452.

Efron, B. (1979) Bootstrap methods: Another look at the jackknife. *Annals of Statistics* 7, 1–26.

Hall, P. & P. Heyde (1980) *Martingale Limit Theory and Its Application*. New York: Academic Press.

Hall, P., J.L. Horowitz, & B.-Y. Jing (1995) On blocking rules for the bootstrap with dependent data. *Biometrika* 82, 561–574.

Hansen, B.E. (1996) Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* 64, 413–430.

Hidalgo, J. (1995) A nonparametric conditional moment test for structural stability. *Econometric Theory* 11, 671–698.

Jacquier, E., N.G. Polson, & P.E. Rossi (1994) Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* 12, 371–389.

Jansen, D.W. & C.G. de Vries (1991) On the frequency of large stock returns: Putting booms and busts into perspective. *Review of Economics and Statistics* 73, 18–24.

Künsch, H.R. (1989) The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* 17, 1217–1241.

Linton, O. & P. Gozalo (1996) Conditional Independence Restrictions: Testing and Estimation. Manuscript.

Loretan, M. & P.C.B. Phillips (1994) Testing the covariance stationarity of heavy-tailed time series. *Journal of Empirical Finance* 1, 211–248.

Phillips, P.C.B., J.W. McFarland, & P.C. McMahon (1996) Robust tests of forward exchange market efficiency with empirical evidence from the 1920s. *Journal of Applied Econometrics* 11, 1–22.

Picard, D. (1985) Testing and estimating change-points in time series. *Advances in Applied Probability* 17, 841–867.

Pollard, D. (1990) *Empirical Processes: Theory and Applications*. Hayward: Institute of Mathematical Statistics.

Sen, P.K. (1974) Weak convergence of multidimensional empirical processes for stationary ϕ-mixing processes. *Annals of Probability* 2, 147–154.

Sowell, F. (1996) Optimal tests for parameter instability in the generalized method of moments framework. *Econometrica* 64, 1085–1107.

Szyszkowicz, B. (1994a) Empirical-type processes in weighted metrics and change-point problems. *Comptes Rendus Mathématiques de l'Académie des Sciences* 16, 221–226.

Szyszkowicz, B. (1994b) Weak convergence of weighted empirical type processes under contiguous and change point alternatives. *Stochastic Processes and Their Applications* 50, 281–313.

Van der Vaart, A.W. & J.A. Wellner (1996) *Weak Convergence and Empirical Processes*. New York: Springer-Verlag.

Yoshihara, K. (1975) Weak convergence of multidimensional empirical processes for strong mixing sequences of stochastic vectors. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 33, 133–137.

APPENDIX A

We will use the following lemma in the proofs.

LEMMA A.1. *Let* $\{\xi_{ni}\}_{i=1}^{n}$ *be strong mixing random variables with mixing coefficients* $\{\alpha(j)\}_{j=1}^{\infty}$ *and let* $\alpha(0) = 1$.

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

(a) *Suppose that*
(i) $|\xi_{ni}| \le 1$, $E(\xi_{ni}) = 0$, *and* $\sup_{i \le n} E(\xi_{ni}^2) \le \tau_n^{2+\gamma}$ *for* $\forall i \le n$ *and* $\forall n \ge 1$,
(ii) $\sum_{j=1}^{\infty} j^{Q/2} \alpha(j)^{\gamma/Q+\gamma} < \infty$,
for some $\tau_n > 0$, $\gamma > 0$, *and even integer* $Q > 2$. *Then there is* $C_{\alpha,\gamma,Q} > 0$ *such that*

$$E\left(\sum_{i=1}^{n} \xi_{ni}\right)^Q \le C_{\alpha,\gamma,Q}\{n\tau_n^2 + \ldots + (n\tau_n^2)^{Q/2}\}$$

for every $n \ge 1$.

(b) *Suppose that* $E(\xi_{ni}) = 0$, $\tau_{2n} \ge E(\xi_{ni}^{2+\delta})$, *and* $\tau_{4n} \ge E(\xi_{ni}^{2(2+\delta)})$ *for some* $\delta > 0$, $\tau_{2n} > 0$, $\tau_{4n} > 0$, *and* $\forall i \le n$. *Then*

$$E\left(\sum_{i=1}^{n} \xi_{ni}\right)^4 \le 3{,}072 n \tau_{2n}^{4/(2+\delta)} \left(\sum_{j=0}^{n} \alpha(j)^{\delta/(2+\delta)}\right)^2 + 576 n^2 \tau_{4n}^{2/(2+\delta)} \sum_{j=0}^{n} (j+1)^2 \alpha(j)^{\delta/(2+\delta)}.$$

(c) *Suppose that* $F_{n,d}(t,t') = P(\xi_{ni} \le t, \xi_{n(i+d)} \le t')$, $F_n(t) = P(\xi_{ni} \le t)$, *and* $\sum_{j=1}^{\infty} j\alpha(j) < \infty$. *Then*

$$\frac{1}{n} E\left\{\sum_{i=1}^{[nr]} (I(\xi_{ni} \le t) - F_n(t))\right\}\left\{\sum_{j=1}^{[nr']} (I(\xi_{nj} \le t') - F_n(t'))\right\} \to \min(r,r') \lim_{n\to\infty} \sum_{d=-n}^{n} (F_{n,d}(t,t') - F_n(t)F_n(t')).$$

Part (a) is a modified version of Lemma 3.1 in Andrews and Pollard (1994) with their τ replaced by our τ_n. The proofs of (b) and (c) are straightforward and are provided in Appendix B.1.

Proof of Theorem 2.1. First, we show convergence of the sample covariance kernel to the specified covariance kernel. Second, we establish convergence of the finite dimensional distributions. The proof of tightness is analogous to the proofs of Theorems 2.1 and 2.2 in Sen (1974) and is provided in Appendix B.3.

First, we will show that

$$\lim_{n\to\infty} E(K_n(r,t)K_n(r',t')) = \min(r,r')\sigma(t,t'). \tag{A.1}$$

By Theorem A5 in Hall and Heyde (1980) and Assumption A,

$$\frac{1}{n} \sum_{i=1}^{[nr]} \sum_{j=1}^{[nr']} |E(I(x_{ni} \le t) - H_{ni}(t))(I(x_{nj} \le t') - H_{nj}(t'))| \le \frac{4}{n} \sum_{i=1}^{[nr]} \sum_{j=1}^{[nr']} \alpha(|i-j|) \le 4 \sum_{j=-\infty}^{\infty} \alpha(j) < \infty. \tag{A.2}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

Thus, the left-hand side of (A.1) is absolutely convergent. By Lemma A.1(c),

$$E(K_n(r,t)K_n(r',t')) = \frac{(1-\delta n^{-1/2})^2}{n}\sum_{i=1}^{[nr]}\sum_{j=1}^{[nr']}(F_{|j-i|}(t,t') - F(t)F(t'))$$
$$+ \frac{\delta^2}{n^2}\sum_{i=1}^{[nr]}\sum_{j=1}^{[nr']}\left(G_{|j-i|}\left(\frac{i}{n},\frac{j}{n},t,t'\right) - G\left(\frac{i}{n},t\right)G\left(\frac{j}{n},t'\right)\right)$$
$$\to \min(r,r')\sum_{d=-\infty}^{\infty}(F_d(t,t') - F(t)F(t'))$$
$$= \min(r,r')\sigma(t,t'). \tag{A.3}$$

Thus, (A.1) is proved.

Appendix B.2 shows that we can confine our attention to weak convergence in $D([0,1]^{p+1})$ instead of that in $D([0,1]\times\Re^p)$, as Billingsley (1968, p. 197) does. Next, we will establish the convergence of the finite dimensional distributions of $\{K_n\}$ to those of $\{K\}$. By the Cramér–Wold device, it suffices to show that

$$\sum_{j=1}^{k}\omega_j K_n(r_j,t_j) \xrightarrow{d} N\left(0, \lim_{n\to\infty}\operatorname{Var}\left\{\sum_{j=1}^{k}\omega_j K_n(r_j,t_j)\right\}\right) \tag{A.4}$$

for $\forall(\omega_1,\ldots,\omega_k)' \in \Re^k$, $\forall(r_1,\ldots,r_k)' \in [0,1]^k$, $\forall(t_1',\ldots,t_k')' \in [0,1]^{kp}$, and $\forall k \geq 1$. Because the degenerate case is trivial, we assume $\lim_{n\to\infty}\operatorname{Var}(\sum_{j=1}^{k}c_j K_n(r_j,t_j)) > 0$ if the limit exists. Consider $\sum_{i=1}^{n}X_{ni}$, where

$$X_{ni} = c_n\sum_{j=1}^{k}\omega_j I(i \leq [nr_j])(I(x_{ni}\leq t_j) - H_{ni}(t_j)),$$
$$c_n = \left[\operatorname{Var}\left\{\sum_{i=1}^{n}\sum_{j=1}^{k}\omega_j I(i\leq[nr_j])(I(x_{ni}\leq t_j) - H_{ni}(t_j))\right\}\right]^{-1/2}.$$

Because $\sum_{i=1}^{n}X_{ni} = \sum_{j=1}^{k}\omega_j K_n(r_j,t_j)/\{\operatorname{Var}(\sum_{j=1}^{k}\omega_j K_n(r_j,t_j))\}^{1/2}$, it suffices to show that $\sum_{i=1}^{n}X_{ni} \xrightarrow{d} N(0,1)$. We will apply Theorem 3.6 in Davidson (1992). Here $E(X_{ni}) = 0$ and $E(\sum_{i=1}^{n}X_{ni})^2 = 1$ satisfy his Assumption A1, and $|X_{ni}/c_n| \leq \sum_{j=1}^{k}|\omega_j|$ for all $i \leq n$ and $n \geq 1$ satisfy his Assumption A2. Our Assumption A implies his Assumption A3. His Assumption A4 is implied by $\sup_{n\geq1}nc_n^2 = \sup_{n\geq1}\{\operatorname{Var}(\sum_{j=1}^{n}\omega_j K_n(r_j,t_j))\}^{-1/2}$ and absolute convergent $\lim_{n\to\infty}E(K_n(r,t)K_n(r',t'))$. Thus, Theorem 3.6 in Davidson (1992) completes the proof of the finite dimensional convergence.

Finally, $P(K' \in C) = 1$ by Theorem 15.5 in Billingsley (1968). ■

Proof of Corollary 2.2. By Theorem 2.1 and the CMT,

$$d_n(r,t) = \left|n^{-1/2}\sum_{i=1}^{[nr]}(I(x_{ni}\leq t) - H_{ni}(t)) - \frac{[nr]}{n}n^{-1/2}\sum_{i=1}^{n}(I(x_{ni}\leq t) - H_{ni}(t))\right.$$
$$\left. + \delta\left(n^{-1}\sum_{i=1}^{[nr]}G\left(\frac{i}{n},t\right) - \frac{[nr]}{n}n^{-1}\sum_{i=1}^{n}G\left(\frac{i}{n},t\right)\right)\right|$$
$$\Rightarrow \left|K^\circ(r,t) + \delta\left(\int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds\right)\right|. \tag{A.5}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

By applying the CMT again,

$$T_1 \xrightarrow{d} \sup_{0<r<1} \sup_{t\in\Re^p} \left| K^\circ(r,t) + \delta\left(\int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds \right) \right|, \tag{A.6}$$

and

$$T_2 = \frac{1}{n-1}\sum_{m=1}^{n-1} \int_{\Re^p} d_n\left(\frac{m}{n},t\right)^2 dF(t) + \frac{1}{n-1}\sum_{m=1}^{n-1} \int_{\Re^p} d_n\left(\frac{m}{n},t\right)^2 d(\hat{F}_n(t) - F(t))$$

$$\Rightarrow \int_0^1 \int_{\Re^p} \left| K^\circ(r,t) + \delta\left(\int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds \right) \right|^2 dF(t)dr. \tag{A.7}$$

It remains to show that $\{(r,t): \int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds \neq 0\}$ is of positive Lebesgue measure. Suppose that $\int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds = 0$ for all (r,t). Then it implies that

$$G(r,t) = \int_0^1 G(s,t)ds$$

for all (r,t). However, this contradicts Assumption B(iii). Thus, there is (r^*,t^*) such that

$$\int_0^{r^*} G(s,t^*)ds - r^*\int_0^1 G(s,t^*)ds \neq 0.$$

Because $\int_0^r G(s,t)ds - r\int_0^1 G(s,t)$ is continuous in (r,t), there is a neighborhood of (r^*,t^*) in which $\int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds \neq 0$ holds. Therefore, $\{(r,t): \int_0^r G(s,t)ds - r\int_0^1 G(s,t)ds \neq 0\}$ is of positive Lebesgue measure. ■

Proof of Theorem 2.3. Let E_ω and Var_ω denote conditional expectation and variance, respectively, given the sample ω. Let $z_i = 0$ for $i = n-\ell+2,\ldots,n$. Let us write $K_n^*(r,t;\omega) = \sum_{i=1}^n f_{ni}(z_i,r,t;\omega)$, where

$$f_{ni}(z,r,t;\omega) = zI(i \le [nr]-\ell+1)n^{-1/2}\sum_{j=i}^{i+\ell-1} (I(x_{nj}(\omega)\le t) - \hat{F}_n(t;\omega))$$

for $(x,r,t)\in\Re\times[0,1]\times\Re^p$. Then the triangular array $\{f_{ni}(z_i,r,t;\omega)\}$ is independent within rows. Given the sample ω, Theorem 10.6 in Pollard (1990) can deliver the (conditional) weak convergence: $K_n^*(\cdot,\cdot;\omega) \Rightarrow K(\cdot,\cdot)$. Thus, if the following conditions of Pollard hold ω-a.s., then the same theorem of Pollard delivers the (almost surely conditional) weak convergence: $K_n^*(\cdot,\cdot;\omega) \Rightarrow K(\cdot,\cdot)$ ω-a.s.

(i) $\lim_{n\to\infty} E_\omega\{K_n^*(r,t;\omega)K_n^*(r',t';\omega)\}$ exists for every (r,t), (r',t') in $[0,1]\times\Re^p$;
(ii) the pseudo-metric $\rho(r,r',t,t') = \lim_{n\to\infty}\rho_n(r,r',t,t';\omega)$ is well defined, where

$$\rho_n(r,r',t,t';\omega) = \left\{ \sum_{i=1}^n E_\omega[(f_{ni}(\cdot,r',t';\omega) - f_{ni}(\cdot,r,t;\omega))^2] \right\}^{1/2}.$$

if $\rho(r_n,r_n',t_n,t_n') \to 0$, then $\rho_n(r_n,r_n',t_n,t_n';\omega) \to 0$, where $\{(r_n,t_n)\}_{n=1}^\infty$ and $\{(r_n',t_n')\}_{n=1}^\infty$ are deterministic sequences.

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

(iii) $\limsup_{n\to\infty}\sum_{i=1}^{n} E_\omega(F_{ni}^2(z;\omega)) < \infty$ where $F_{ni}(z;\omega)$ is the envelope of f_{ni};
(iv) $\sum_{i=1}^{n} E\{F_{ni}^2(x;\omega)I(F_{ni}(x;\omega) > \varepsilon)\} \to 0$ for each $\varepsilon > 0$;
(v) $\{f_{ni}\}$ are manageable in the sense of Definition 7.9 of Pollard (1990, p. 38).

Our conditions (i), (ii), (iii), (iv), and (v) correspond to Pollard's conditions (ii), (v), (iii), (iv), and (i), respectively.

(i) We will prove the uniform convergence of $E_\omega(K_n^*(r,t;\omega)K_n^*(r',t';\omega))$ to $\min(r,r')\sigma(t,t')$ ω-a.s. Without loss of generality, we assume that $r < r'$. Consider

$$\begin{aligned}
&\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega)) - E[\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))] \\
&\quad = \frac{\ell}{n}\sum_{j=1}^{[nr]-\ell+1} Z_{1j}(\omega) + \frac{\ell}{n}\sum_{j=[nr]-\ell+2}^{[nr']-\ell+1} Z_{2j}(\omega), \qquad \textbf{(A.8)}
\end{aligned}$$

where

$$\begin{aligned}
Z_{1j}(\omega) &= \left\{\frac{1}{\ell}\sum_{i=j}^{j+\ell-1}(I(x_{ni}(\omega)\le t') - \hat{F}_n(t';\omega)) - \frac{1}{\ell}\sum_{i=j}^{j+\ell-1}(I(x_{ni}(\omega)\le t) - \hat{F}_n(t;\omega))\right\}^2 \\
&\quad - E\left[\left\{\frac{1}{\ell}\sum_{i=j}^{j+\ell-1}(I(x_{ni}(\omega)\le t') - \hat{F}_n(t';\omega))\right.\right. \\
&\qquad \left.\left. - \frac{1}{\ell}\sum_{i=j}^{j+\ell-1}(I(x_{ni}(\omega)\le t) - \hat{F}_n(t;\omega))\right\}^2\right] \\
&= V_{1j}(\omega) - E(V_{1j}), \qquad \textbf{(A.9)}
\end{aligned}$$

$$\begin{aligned}
Z_{2j}(\omega) &= \left\{\frac{1}{\ell}\sum_{i=j}^{j+\ell-1}(I(x_{ni}(\omega)\le t') - \hat{F}_n(t';\omega))\right\}^2 \\
&\quad - E\left[\left\{\frac{1}{\ell}\sum_{i=j}^{j+\ell-1}(I(x_{ni}(\omega)\le t') - \hat{F}_n(t';\omega))\right\}^2\right].
\end{aligned}$$

Let $\|\cdot\|_p$ denote $(E|X|^p)^{1/p}$. By the Minkowski inequality and Lemma A.1(a) with $Q = 16$,

$$\begin{aligned}
\|V_{1j}(\omega)\|_8 &\le \left\|\frac{1}{\ell}\sum_{i=j}^{j+\ell-1} I(x_{ni}\le t') - H_{ni}(t') - I(x_{ni}\le t) + H_{ni}(t)\right\|_{16}^2 \\
&\quad + \left\|\frac{1}{\ell}\sum_{i=j}^{j+\ell-1}\left(H_{ni}(t') - H_{ni}(t) - \frac{1}{n}\sum_{k=1}^{n} H_{nk}(t') + \frac{1}{n}\sum_{k=1}^{n} H_{nk}(t)\right)\right\|_{16}^2 \\
&\quad + \left\|\hat{F}_n(t';\omega) - \frac{1}{n}\sum_{k=1}^{n} H_{nk}(t') - \hat{F}_n(t;\omega) + \frac{1}{n}\sum_{k=1}^{n} H_{nk}(t)\right\|_{16}^2 \\
&= O(\ell^{-1}), \qquad \textbf{(A.10)}
\end{aligned}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

and

$$\|E(V_{1j})\|_8 \le \left(\left\| \frac{1}{\ell} \sum_{i=j}^{j+\ell-1} I(x_{ni} \le t') - H_{ni}(t') - I(x_{ni} \le t) - H_{ni}(t) \right\|_2 \right.$$
$$+ \left\| \frac{1}{\ell} \sum_{i=j}^{j+\ell-1} H_{ni}(t') - \frac{1}{n} \sum_{k=1}^{n} H_{nk}(t') H_{ni}(t) - \frac{1}{n} \sum_{k=1}^{n} H_{nk}(t) \right\|_2$$
$$\left. + \left\| \hat{F}_n(t';\omega) - \frac{1}{n} \sum_{k=1}^{n} H_{nk}(t') \hat{F}_n(t;\omega) - \frac{1}{n} \sum_{k=1}^{n} H_{nk}(t) \right\|_2 \right)^{1/2}$$
$$= O(\ell^{-1}). \tag{A.11}$$

The term O's in (A.10) and (A.11) are uniform in j, t, and t' by Lemma A.1(a) with $\tau_n = 1$ and the definition of mixing coefficients. By applying the Minkowski inequality to (A.9)–(A.11), it follows that

$$E(Z_{1j}^8) = O(\ell^{-8}), \tag{A.12}$$

where O is uniform in j, t, and t'. Using Lemma A.1(b) with $\delta = 2$,

$$E\left(\sum_{i=1}^{[nr]-\ell+1} Z_{1i} \right)^4 = O(\ell^2 (n\tau_{2n} + n^2 \tau_{4n}^{1/2})), \tag{A.13}$$

where $\tau_{2n} = \sup_{i \le n} E(Z_{1i}^4)$, $\tau_{4n} = \sup_{i \le n} E(Z_{1i}^8)$, and the term O is uniform in r, t, and t'. Equations (A.12) and (A.17) imply

$$E\left(\sum_{i=1}^{[nr]-\ell+1} Z_{1i}^4 \right) = O(\ell^{-2} n^2). \tag{A.14}$$

A similar argument leads to

$$E\left(\sum_{[nr]-\ell+2}^{[nr']-\ell+1} Z_{2j} \right)^4 = O(\ell^{-2} n^2). \tag{A.15}$$

By the Minkowski inequality,

$$E[\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega)) - E\{\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))\}]^4$$
$$\le \left[\frac{\ell}{n} \left\{ E\left(\sum_{j=1}^{[nr]-\ell+1} Z_{1j} \right)^4 \right\}^{1/4} + \frac{\ell}{n} \left\{ E\left(\sum_{j=[nr]-\ell+2}^{[nr']-\ell+1} Z_{2j} \right)^4 \right\}^{1/4} \right]^4. \tag{A.16}$$

It follows from (A.14)–(A.16) that

$$E\{\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega)) - E[\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))]\}^4$$
$$= O\left(\frac{\ell^2}{n^2} \right), \tag{A.17}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

where the term O is uniform in r, r', t, and t'. A similar argument leads to (A.17) for $r \geq r'$. By the Markov inequality, (A.17) implies that, for every $\eta > 0$,

$$P\{|\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega)) - E[\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))]| > \eta\}$$

$$= O\left(\frac{\ell^2}{n^2}\right). \tag{A.18}$$

Because $\ell = o(n^{1/2})$, it follows that

$$\sum_{n=1}^{\infty} P\{|\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega)) - E[\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))]| > \eta\}$$

$$< \infty$$

for every $\eta > 0$. By the Borel–Cantelli lemma,

$$\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega)) - E[\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))]$$

$$= o(1) \quad \omega\text{-a.s.}, \tag{A.19}$$

where the term o is uniform in r, r', t, and t'. If we prove

$$E[\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))] - \mathrm{Var}(K(r',t') - K(r,t)) = o(1), \tag{A.20}$$

where the term o is uniform in r, r', t, and t', then (A.19) and (A.20) will imply

$$\mathrm{Var}_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega)) - \mathrm{Var}(K(r',t') - K(r,t)) = o(1) \quad \omega\text{-a.s.}, \tag{A.21}$$

where the term o is uniform in r, r', t, and t'. Then we have

$$E_\omega(K_n^*(r,t;\omega)K_n^*(r',t';\omega)) = E(K(r,t)K(r',t')) + o(1) \quad \omega\text{-a.s.}, \tag{A.22}$$

where the term o is uniform in r, r', t, and t', which implies the pointwise convergence (i).

Equation (A.20) remains to be proved. It suffices to show that

$$E[E_\omega(K_n^*(r,t;\omega)K_n^*(r',t';\omega))] = \min(r,r')\sigma(t,t') + o(1), \tag{A.23}$$

where the term o is uniform in r, r', t, and t'. The left-hand side of (A.23) is

$$E[E_\omega(K_n^*(r,t;\omega)K_n^*(r',t';\omega))]$$

$$= \frac{1}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}\frac{1}{\ell}E\left[\sum_{i_1=j}^{j+\ell-1}(I(x_{ni_1} \leq t) - \hat{F}_n(t))\sum_{i_2=j}^{j+\ell-1}(I(x_{ni_2} \leq t') - \hat{F}_n(t'))\right]$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

$$= \frac{1}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}\frac{1}{\ell}E\left[\sum_{i_1=j}^{j+\ell-1}\sum_{i_2=j}^{j+\ell-1}(I(x_{ni_1}\le t)-F(t))(I(x_{ni_2}\le t')-F(t'))\right]$$

$$+ \frac{1}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}\frac{1}{\ell}E\left[\sum_{i=j}^{j+\ell-1}(I(x_{ni}\le t)-F(t))(F(t')-\hat{F}_n(t'))\right]$$

$$+ \frac{1}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}\frac{1}{\ell}E\left[\sum_{i=j}^{j+\ell-1}(I(x_{ni}\le t')-F(t'))(F(t)-\hat{F}_n(t))\right]$$

$$+ \frac{\ell}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}E[(\hat{F}_n(t)-F(t))(\hat{F}_n(t')-F(t'))]$$

$$= A_n + B_n + C_n + D_n. \tag{A.24}$$

First,

$$A_n = \frac{1}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}\frac{1}{\ell}\sum_{i_1=j}^{j+\ell-1}\sum_{i_2=j}^{j+\ell-1}(H_{n,i_1,i_2-i_1}(t,t')-H_{ni_1}(t)H_{ni_2}(t'))$$

$$= \frac{1}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}\frac{1}{\ell}\sum_{i_1=j}^{j+\ell-1}\sum_{i_2=j}^{j+\ell-1}(F_{i_2-i_1}(t,t')-F(t)F(t'))+O(\ell n^{-1/2})$$

$$= \min(r,r')\sigma(t,t') - \min(r,r')\sum_{d=-\infty}^{-\ell}(F_d(t,t')-F(t)F(t'))$$

$$- \min(r,r')\sum_{d=\ell}^{\infty}(F_d(t,t')-F(t)F(t'))$$

$$- \frac{\min(r,r')}{\ell}\sum_{d=-\ell+1}^{\ell-1}|d|(F_d(t,t')-F(t)F(t'))+O(\ell n^{-1/2})$$

$$= \min(r,r')\sigma(t,t')+o(1), \tag{A.25}$$

where the term o is uniform in r, r', t, and t' by the definition of mixing coefficients. Next,

$$B_n = \frac{1}{n}\sum_{j=1}^{\min([nr],[nr'])-\ell+1}\frac{1}{\ell}$$

$$\times\left\{\sum_{i=j}^{j+\ell-1}(H_{ni}(t)-F(t))F(t') + \frac{1}{n}\sum_{i=j}^{j+\ell-1}\sum_{k=1}^{n}H_{n,i,k-i}(t,t') - F(t)\frac{1}{n}\sum_{k=1}^{n}H_{n,k}(t')\right\}$$

$$= O(n^{-1/2}) + O(\ell n^{-1}) = o(1), \tag{A.26}$$

where the term o is uniform in r, r', t, and t' by the definition of $H_{ni,d}(t,t')$ and $H_{ni}(t)$. Similarly,

$$C_n = o(1). \tag{A.27}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

Last,

$$D_n = \frac{\ell}{n} \sum_{j=1}^{\min([nr],[nr'])-\ell+1} \left\{ \frac{1}{n^2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} H_{n,i_1,i_2-i_1}(t,t') - \frac{1}{n} \sum_{i=1}^{n} H_{ni}(t)F(t') - \frac{1}{n} \sum_{i=1}^{n} H_{ni}(t')F(t) + F(t)F(t') \right\}$$

$$= \frac{\ell \min([nr],[nr']) - \ell + 1}{n^3} \sum_{i=1}^{n} \sum_{j=1}^{n} (F_{i-j}(t,t') - F(t)F(t')) + O(\ell n^{-1/2})$$

$$= o(1), \tag{A.28}$$

where the term o is uniform in r, r', t, and t' by the definitions of $H_{ni,d}(t,t')$, $H_{ni}(t)$, and mixing coefficients. Therefore, (A.20) follows from (A.24)–(A.28).

(ii) Because $\rho_n(r,r',t,t';\omega) = [E_\omega(K_n^*(r',t';\omega) - K_n^*(r,t;\omega))^2]^{1/2}$, (i) implies that, for $\forall(r,t), \forall(r',t') \in (0,1) \times \Re^p$,

$$\rho(r,r',t,t') = \lim_{n\to\infty} \rho_n(r,r',t,t';\omega)$$

$$= (r'\sigma(t',t') - 2\min(r,r')\sigma(t',t) + r\sigma(t,t))^{1/2} \quad \omega\text{-a.s.} \tag{A.29}$$

Thus, $\rho(r,r',t,t')$ is well defined for all (r,t), (r',t') in $[0,1] \times \Re^p$ ω-a.s. Take a deterministic sequence $\{r_n, r_n', t_n, t_n'\}$ such that $\lim_{n\to\infty} \rho(r_n, r_n', t_n, t_n') = 0$. Because

$$\rho_n(r_n,r_n',t_n,t_n';\omega)^2 \le \sup_{r,r',t,t'} |\rho_n(r,r',t,t';\omega)^2 - \rho(r,r',t,t')^2| + \rho(r_n,r_n',t_n,t_n')^2, \tag{A.30}$$

(A.22) implies that $\lim_{n\to\infty} \rho_n(r_n,r_n',t_n,t_n';\omega) = 0$ ω-a.s.

(iii) $\{F_{ni}(z;\omega)\}$ are called the envelopes of $\{f_{ni}(z,r,t;\omega)\}$ if $|f_{ni}(z,r,t;\omega)| \le F_{ni}(z;\omega)$ for all x, r, and t. We use the minimum envelopes, i.e.,

$$F_{ni}(z;\omega) = n^{-1/2}|z| \sup_{t\in\Re^p} \left| \sum_{j=i}^{i+\ell-1} I(x_{nj}(\omega) \le t) - \hat{F}_n(t;\omega) \right|. \tag{A.31}$$

Then

$$\limsup_{n\to\infty} \sum_{i=1}^{n} E_\omega(F_{ni}(z_i;\omega)^2)$$

$$= \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sup_{t\in\Re^p} \frac{1}{\ell} \left\{ \sum_{j=i}^{i+\ell-1} (I(x_{nj}(\omega) \le t) - \hat{F}_n(t;\omega)) \right\}^2$$

$$\le 3 \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sup_{t\in\Re^p} \frac{1}{\ell} \left\{ \sum_{j=1}^{i+\ell-1} (I(x_{nj}(\omega) \le t) - H_{nj}(t)) \right\}^2 \tag{A.32}$$

$$+ 3 \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sup_{t\in\Re^p} \frac{1}{\ell} \left\{ \sum_{j=1}^{i+\ell-1} \left(H_{nj}(t) - \frac{1}{n} \sum_{k=1}^{n} H_{nk}(t) \right) \right\}^2 \tag{A.33}$$

$$+ 3 \limsup_{n\to\infty} \ell \sup_{t\in\Re^p} \left(\hat{F}_n(t;\omega) - \frac{1}{n} \sum_{i=1}^{n} H_{ni}(t) \right)^2. \tag{A.34}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

By Lemma A1(b) and the Borel–Cantelli lemma, one can show that (A.32) is $O_{as}(1)$. By the definition of $H_{ni}(t)$, (A.33) is $O(\ell/n)$. By a version of the Glivenko–Cantelli lemma given in Appendix B.4, (A.34) is $O_{as}(1)$. Therefore, (iii) is proved.

(iv) By the Cauchy–Schwarz inequality,

$$\begin{aligned}
&\sum_{i=1}^{n} E_\omega\{F_{ni}^2(z_i;\omega) I(F_{ni}(z_i;\omega) > \varepsilon)\} \\
&\quad = \frac{1}{n}\sum_{i=1}^{n} E_\omega\left\{ z_i^2 \sup_{t\in\Re^p} \left| \sum_{j=i}^{i+\ell-1} I(x_{ni}(\omega) \le t) - \hat{F}_n(t;\omega) \right|^2 I(F_{ni}(z_i;\omega) > \varepsilon) \right\} \\
&\quad \le \frac{1}{n}\sum_{i=1}^{n} \sup_{t\in\Re^p} \frac{1}{\ell} \left\{ \sum_{j=i}^{i+\ell-1} (I(x_{nj}(\omega) \le t) - \hat{F}_n(t;\omega)) \right\}^2 \\
&\qquad \times P\left(z_i > \varepsilon n^{1/2} \Big/ \sup_{t\in\Re^p} \left| \frac{1}{\ell} \left\{ \sum_{j=i}^{i+\ell-1} (I(x_{nj}(\omega) \le t) - \hat{F}_n(t;\omega)) \right\} \right| \right)^{1/2}. \qquad \textbf{(A.35)}
\end{aligned}$$

By the Markov inequality,

$$\begin{aligned}
&P\left(z_i > \varepsilon n^{1/2} \Big/ \sup_{t\in\Re^p} \left| \frac{1}{\ell} \left\{ \sum_{j=i}^{i+\ell-1} (I(x_{nj}(\omega) \le t) - \hat{F}_n(t;\omega)) \right\} \right| \right) \\
&\quad \le \frac{1}{\ell} \sup_{t\in\Re^p} \left| \frac{1}{\ell} \left\{ \sum_{j=i}^{i+\ell-1} (I(x_{nj}(\omega) \le t) - \hat{F}_n(t;\omega)) \right\} \right| \Big/ (\varepsilon^2 n) \\
&\quad = O\left(\frac{\ell}{n}\right), \qquad \textbf{(A.36)}
\end{aligned}$$

where O is uniform in i. Thus, (iv) follows from (iii) and (A.35).

(v) We use the covering number instead of the packing number (see Pollard, 1990, Definition 3.3 and the following inequality, p. 10). The covering number is defined as the smallest number of closed balls with radius $(x/2)\sqrt{\sum_{i=1}^n \alpha_i^2 F_{ni}(z_i;\omega)^2}$ whose union covers $\{\alpha_i f_{ni}(z_i, r, t;\omega)\}$ where $\alpha_i \ge 0$ for $i = 1,2,\ldots,n$ (see Pollard, 1990, inequality (10.7), p. 54). Thus, the following inequality must hold within each closed ball:

$$\sum_{i=1}^{n} \alpha_i^2 E_\omega(f_{ni}(z_i, r', t';\omega) - f_{ni}(z_i, r, t;\omega))^2 \le \frac{x^2}{4} \sum_{i=1}^{n} \alpha_i^2 E_\omega F_{ni}(z_i;\omega)^2 \qquad \textbf{(A.37)}$$

for all z_i, $\alpha \ge 0$, and $x > 0$. An argument similar to the proof of (A.21) leads to

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{i=1}^{n} \alpha_i^2 E_\omega(f_{ni}(z_i, r', t';\omega) - f_{ni}(z_i, r, t;\omega))^2 \\
&\quad = \sum_{i=1}^{\infty} \alpha_i^2 (r\sigma(t,t) - 2\min(r,r')\sigma(t,t') + r(t',t')) \\
&\quad = \sum_{i=1}^{\infty} \alpha_i^2 \rho^2(r,r',t,t') \quad \omega\text{-a.s.}
\end{aligned}$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

An argument analogous to the proof of (iii) yields

$$\lim_{n\to\infty} \frac{x^2}{4} \sum_{i=1}^{n} \alpha_i^2 E_\omega F_{ni}(z_i;\omega)^2 = \frac{x^2}{4} C \quad \omega\text{-a.s.}$$

for some $C > 0$. Thus, (A.37) can be approximated by

$$\sum_{i=1}^{\infty} \alpha_i^2 \rho^2(r, r', t, t') \le \frac{x^2}{4} C \tag{A.38}$$

for sufficiently large n. Let $r_i = \tilde{\delta} i$ and choose t_j so that $F(t_j) = \tilde{\delta} j$ for $i, j = 1,2,\ldots, [1/\tilde{\delta}]$ and $\tilde{\delta} > 0$. Then, for every $(r,t) \in [0,1] \times \Re^p$, there is (i,j) such that

$$\begin{aligned}
&\rho(r_i, r, t_j, t) \\
&\quad = r\sigma(t,t) - 2\min(r, r_i)\sigma(t, t_j) + r_i \sigma(t_j, t_j) \\
&\quad = r \lim_{n\to\infty} n^{-1} \sum_{i=-n}^{n} E(I(x_{ni} \le t_j) - I(x_{ni} \le t) - F(t_j) + F(t))^2 + (r_i - r)\sigma(t_j, t_j) \\
&\quad \le rC_{\alpha,\gamma,2}\{(F(t_j) + F(t) - F(\min(t, t_j)))(1 - F(t_j) + F(t))\}^{2/(2+\gamma)} \\
&\qquad + (r_i - r)\sigma(t_j, t_j) \\
&\quad \le C_{\alpha,\gamma,2}(F(t_j) - F(t))^{2/(2+\gamma)} + (r_i - r) \sup_{t' \in \Re^p} \sigma(t', t') \\
&\quad \le \left(C_{\alpha,\gamma,2} + \sup_{t' \in \Re^p} \sigma(t', t')\right) \tilde{\delta}^{2/(2+\gamma)},
\end{aligned} \tag{A.39}$$

where the inequality follows from Lemma A.1(a) and we assume $r < r_i$ without loss of generality. If $\tilde{\delta}$ solves

$$\sum_{i=1}^{\infty} \alpha_i^2 \left(C_{\alpha,\gamma,2} + \sup_{t' \in \Re^p} \sigma(t', t')\right) \tilde{\delta}^{2/(2+\gamma)} = \frac{x^2}{4} \sup_{t \in \Re^p} \rho(1, 0, t, 0)^2, \tag{A.40}$$

then (A.38) holds for all sufficiently large n ω-a.s. within each closed ball. Because the capacity bound is $O(x^{-2(2+\gamma)})$, the integrability condition is also satisfied. ■

Proof of Corollary 2.4. Because

$$d^*(r, t; \omega) = \left(\frac{n}{n - \ell + 1}\right)^{1/2} \left(K_n^*(r, t; \omega) - \frac{[nr] - \ell + 1}{n - \ell + 1} K_n^*(1, t; \omega)\right), \tag{A.41}$$

it follows that $d^*(\cdot,\cdot;\omega) \Rightarrow K^\circ(\cdot,\cdot)$ ω-a.s. Then the CMT delivers the limiting distribution of $T_1^*(\omega)$. Let μ_n denote the probability measure that puts probability $1/(n - \ell)$ on each element of $\{\ell/n, (\ell + 1)/n, \ldots, (n - 1)/n\}$. By Theorem 3.2 of Billingsley (1968),

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

the probability measure $\mu_n \times \hat{F}_n(\cdot)$ converges weakly to $\mu \times F(\cdot)$ where μ is the uniform probability measure on $[0,1]$. Then we have

$$T_2^*(\omega) = \int_0^1 \int_{t \in \Re^p} d_n^*(r,t;\omega)^2 d(\mu_n(r) \times \hat{F}_n(t))$$

$$= \int_0^1 \int_{t \in \Re^p} K^\circ(r,t)^2 d(\mu_n(r) \times \hat{F}_n(t)) + o_{as}(1)$$

$$= \int_0^1 \int_{t \in \Re^p} K^\circ(r,t)^2 d(\mu(r) \times F(t)) + o_{as}(1), \tag{A.42}$$

where the second equality follows from the weak convergence of $d_n^*(\cdot,\cdot)$ *to* $K^\circ(\cdot,\cdot)$ and the last equality follows from the continuity of $K^\circ(\cdot,\cdot)$ and the weak convergence of $\mu_n \times \hat{F}_n(\cdot)$ to $\mu \times F(\cdot)$. ■

Proof of Theorem 2.5. We first consider the normalized distance, and then we consider the normalized bootstrap distance. Let k be a integer and r be a real number such that $\pi_k < r < \pi_{k+1}$ and $0 \le k \le K$. Then the normalized distance is

$$n^{-1/2} d_n(r,t) = I(k>0) \sum_{j=1}^{k} n^{-1} \sum_{i=[n\pi_{j-1}]+1}^{[n\pi_j]} (I(x_{ni} \le t) - F_j(t))$$

$$+ \frac{1}{n} \sum_{i=[n\pi_k]+1}^{[nr]} (I(x_{ni} \le t) - F_{k+1}(t))$$

$$- \frac{[nr]}{n} \sum_{j=1}^{K+1} n^{-1} \sum_{i=[n\pi_{i-1}]+1}^{[n\pi_i]}$$

$$\times (I(x_{ni} \le t) - F_j(t)) + I(k>0) \sum_{j=1}^{k} \frac{[n\pi_j] - [n\pi_{j-1}]}{n} F_j(t)$$

$$+ \frac{[nr] - [n\pi_k]}{n} F_{k+1}(t) - \frac{[nr]}{n} \sum_{j=1}^{K+1} \frac{[n\pi_j] - [n\pi_{j-1}]}{n} F_j(t)$$

$$\xrightarrow{as} I(k>0) \sum_{j=1}^{k} (\pi_j - \pi_{j-1}) F_j(t) + (r - \pi_k) F_{k+1}(t)$$

$$- r \sum_{j=1}^{K+1} (\pi_j - \pi_{j-1}) F_j(t). \tag{A.43}$$

By Assumption E, there are k^* and t^* such that

$$F_{k^*+1}(t^*) - \sum_{j=1}^{K+1} (\pi_j - \pi_{j-1}) F_j(t^*) \neq 0.$$

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press

Let r^* be a real number such that $\pi_{k^*} < r^* < \pi_{k^*+1}$ and

$$r^* \neq \frac{\pi_{k^*} F_{k^*+1}(t^*) - I(k^* > 0) \sum_{j=1}^{k^*} (\pi_j - \pi_{j-1}) F_j(t^*)}{F_{k^*+1}(t^*) - \sum_{j=1}^{K+1} (\pi_j - \pi_{j-1}) F_j(t^*)}.$$

Then it follows from (A.43) that

$$\lim_{n\to\infty} n^{-1/2} d_n(r^*, t^*) \neq 0 \quad \omega\text{-a.s.}$$

Because the right-hand side of (A.43) is continuous in (r, t), $\lim_{n\to\infty} n^{-1/2} d_n(r, t) \neq 0$ ω-a.s. for (r, t) in a neighborhood of (r^*, t^*). Therefore, $\lim_{n\to\infty} n^{-1/2} T_1 > 0$ ω-a.s. and $\lim_{n\to\infty} n^{-1} T_2 > 0$ ω-a.s.

Next, the normalized bootstrap distance is

$$\begin{aligned}
\ell^{-1/2} d_n^*(r, t; \omega) = {} & \frac{1}{\ell^{1/2}(n-\ell+1)^{1/2}} \sum_{i=1}^{[nr]-\ell} z_i \sum_{j=i}^{i+\ell-1} (I(x_{nj} \le t) - H_{nj}(t)) \\
& - \frac{[nr]-\ell+1}{\ell^{1/2}(n-\ell+1)^{3/2}} \sum_{i=1}^{n-\ell+1} z_i \sum_{j=i}^{i+\ell-1} (I(x_{nj} \le t) - H_{nj}(t)) \\
& - \left\{ \left(\frac{\ell^{1/2}}{n-\ell+1} \sum_{i=1}^{[nr]-\ell+1} z_i - \frac{\ell^{1/2}([nr]-\ell+1)}{(n-\ell+1)^2} \sum_{i=1}^{n-\ell+1} z_i \right) \right. \\
& \qquad \left. \times \frac{(n-\ell+1)^{1/2}}{n} \sum_{j=1}^{n} (I(x_{nj} \le t) - H_{nj}(t)) \right\} \\
& + \frac{1}{\ell^{1/2}(n-\ell+1)^{1/2}} \sum_{i=1}^{[nr]-\ell+1} z_i \sum_{j=i}^{i+\ell-1} \left(H_{nj}(t) - \frac{1}{n} \sum_{k=1}^{n} H_{nk}(t) \right) \\
& - \frac{[nr]-\ell+1}{\ell^{1/2}(n-\ell+1)^{3/2}} \sum_{i=1}^{n-\ell+1} z_i \sum_{j=i}^{i+\ell-1} \left(H_{nj}(t) - \frac{1}{n} \sum_{k=1}^{n} H_{nk}(t) \right) \\
= {} & O_p(1).
\end{aligned}$$

Thus, $\ell^{-1/2} T_1^* = O_p(1)$ and $\ell^{-1} T_2^* = O_p(1)$. ■

https://doi.org/10.1017/S0266466601171057 Published online by Cambridge University Press