The Annals of Statistics
1995, Vol. 23, No. 2, 450–475

AN OMNIBUS TEST FOR INDEPENDENCE OF A SURVIVAL TIME FROM A COVARIATE

BY IAN W. MCKEAGUE,[1] A. M. NIKABADZE AND YANQING SUN

Florida State University, A. Razmadze Mathematical Institute and University of Rochester

It has been over 60 years since Kolmogorov introduced a distribution-free omnibus test for the simple null hypothesis that a distribution function coincides with a given distribution function. Doob subsequently observed that Kolmogorov's approach could be simplified by transforming the empirical process to an empirical process based on uniform random variables. Recent use of more sophisticated transformations has led to the construction of asymptotically distribution-free omnibus tests when unknown parameters are present. The purpose of the present paper is to use the transformation approach to construct an asymptotically distribution-free omnibus test for independence of a survival time from a covariate. The test statistic is obtained from a certain test statistic *process* (indexed by time and covariate), which is shown to converge in distribution to a Brownian sheet. A simulation study is carried out to investigate the finite sample properties of the proposed test and an application to data from the British Medical Research Council's 4th myelomatosis trial is given.

1. Introduction. A standard way of testing for independence of a survival time from a covariate z is to fit Cox's (1972) model for the conditional hazard function, $\lambda(t \mid z) = \lambda_0(t)\exp(\beta_0 z)$, and test whether the regression parameter β_0 is zero. However, this test has limited power because of the restrictive (viz. parametric and multiplicative) modeling of the covariate effect.

In this paper we develop an omnibus test that can detect *arbitrary* forms of dependence of a (possibly censored) survival time on a one-dimensional covariate, and which is asymptotically distribution-free. The latter property will be achieved via the transformation method of Doob (1949) and Khmaladze (1981, 1993).

We begin by giving some background to the general problem of constructing omnibus tests (i.e., tests consistent against all alternatives) which have the distribution-free property. First consider the simple hypothesis $F = F_0$, where F_0 is specified and the lifetimes $T_1, \dots, T_n$ are completely observed iid

Received December 1992; revised October 1994.
[1]Research partially supported by the Air Force Office of Scientific Research under Grant AFOSR-91-0048.
AMS 1991 *subject classifications*. Primary 62G10; secondary 62G20.
Key words and phrases. Distribution-free, goodness-of-fit, innovation Brownian sheet, counting processes, martingale methods.

random variables having distribution function F. Let

$$\hat{F}(t) = \frac{1}{n} \sum_{i=1}^{n} I(T_i \le t)$$

be the empirical distribution function of the T_i's and $\nu_n(t) = \sqrt{n}(\hat{F}(t) - F_0(t))$ the empirical process. Assume that F_0 is continuous. Doob (1949) transformed $\nu_n(t)$ to the uniform empirical process $u_n(x) = \nu_n(F_0^{-1}(x))$, which is an empirical process based on the iid uniform random variables $F_0(T_i)$, $i = 1, \dots, n$. The distribution of u_n does not depend on F_0 (and it converges weakly to a Brownian bridge), so the distribution of any test statistic that is a functional of u_n is free from F_0. In particular, the Kolmogorov–Smirnov statistic $\sup_x |u_n(x)|$ and the Cramér–von Mises statistic $\int u_n^2(x)\,dx$ are distribution-free.

Next consider the composite null hypothesis $F = F_0(\cdot, \theta)$, where θ is an unknown parameter. The natural extension of the above transformation, $\hat{u}_n(x) = \hat{\nu}_n(F_0^{-1}(x, \hat{\theta}))$, where $\hat{\nu}_n(t) = \sqrt{n}(\hat{F}(t) - F_0(t, \hat{\theta}))$ is the parametric empirical process and $\hat{\theta}$ is an estimator of θ, is unfortunately no longer distribution-free or even asymptotically distribution-free [Durbin (1973)]. As a consequence, classical statistics such as $\sup_x |\hat{u}_n(x)|$ or $\int \hat{u}_n^2(x)\,dx$ have limit distributions which depend on F_0. Thus, in order to generalize what the uniform empirical process does in the case of simple hypotheses, it is necessary to construct a more sophisticated transformation of $\hat{\nu}_n$. Khmaladze (1981, 1993) introduced martingale methods to address this problem; see also Nikabadze (1987). The parametric empirical process $\hat{\nu}_n$ converges weakly to some zero-mean Gaussian process ν [Durbin (1973)], so Khmaladze first transformed the process ν to an innovation martingale, which is a Gaussian process with independent increments and covariance function $F_0(s \wedge t, \theta)$ and which preserves the information in ν. Then he transformed the innovation martingale to a standard Brownian motion w. Applying the transformation $\nu \mapsto w$ to $\hat{\nu}_n$ results in a test process that converges weakly to Brownian motion. This leads to an asymptotically distribution-free omnibus test based on the supremum norm (say) of the test process. Note that there is some loss of information in reducing to a single test statistic, via supremum norm, but this does not affect the omnibus property. Also, the transformation approach is not designed to reveal the *nature* of a departure from the null hypothesis—a graphical inspection of the (untransformed) parametric empirical process might be useful for that.

In survival analysis, one is rarely able to observe complete life histories. Important examples occur with right censoring and left truncation [Keiding and Gill (1990)]. These examples fit into the general setting of Aalen's (1978) multiplicative intensity model for counting processes. In that setting it is natural to formulate hypotheses in terms of the hazard function $\lambda(t)$ or the cumulative hazard function $\Lambda(t) = \int_0^t \lambda(s)\,ds$, rather than the distribution function F. Andersen, Borgan, Gill and Keiding (1982, 1993) studied tests of the simple hypothesis $\lambda = \lambda_0$ in terms of functionals of $\sqrt{n}(\hat{\Lambda} - \Lambda_0)$, where $\hat{\Lambda}$

is the Nelson–Aalen estimator. Hjort (1990) considered the composite hypothesis $\lambda = \lambda_0(\cdot, \theta)$, with statistics based on functionals of the process $\sqrt{n}\,(\hat{\Lambda}(t) - \Lambda_0(t, \hat{\theta}))$, where $\hat{\theta}$ is the maximum likelihood estimator of θ. This process converges weakly to a zero-mean Gaussian process under the null hypothesis. An innovation martingale can be found for the limit process and used to construct an asymptotically distribution-free omnibus test; see Andersen, Borgan, Gill and Keiding [(1993), Section VI.3.3.4].

In many applications of survival analysis it is important to consider whether a covariate has some effect upon survival, say through the conditional hazard function $\lambda(t \mid z) = \lambda(t, z)$. That is, one would like to test the null hypothesis

$$H_0\colon \lambda(t, z) \text{ does not depend on the covariate } z$$

against the general alternative that $\lambda(t, z)$ depends on z. For simplicity, we shall restrict the domain of (t, z) to be the unit square. An omnibus test of H_0 is feasible when the covariate is one dimensional, such as age at diagnosis, disease duration and so forth. Indeed, McKeague and Utikal [(1990), subsequently MU] proposed such a test based on the process $X(t, z) = \sqrt{n}\,(\hat{\mathscr{A}} - \bar{\mathscr{A}})$, where $\hat{\mathscr{A}}$ is an estimate of the doubly cumulative hazard function $\mathscr{A}(t, z) = \int_0^t\int_0^z \lambda(s, x)\,dx\,ds$ and $\bar{\mathscr{A}}(t, z) = z\hat{\Lambda}(t)$ is the natural estimate of $\mathscr{A}$ under H_0. They showed that X converges weakly under H_0 to a Gaussian random field of the form

$$m(t, z) = \int_0^t\int_0^z \sqrt{h}\, dW - b(z)\int_0^t\int_0^1 g\, dW, \tag{1.1}$$

where W is a Brownian sheet, $b(z) = z$ and h, g are certain nonrandom functions (see Section 3.1). The above stochastic integrals are defined in the L^2-sense; see Wong and Zakai (1974). MU's test was based on the Kolmogorov–Smirnov statistic computed directly from X. However, while asymptotically omnibus, such a test is not asymptotically distribution-free and would require simulation of the process m to find critical values.

We shall construct a transformation J that maps m to its innovation Brownian sheet. An estimated version $\hat{J}$ of J will be obtained by plugging an estimate of h into J (it turns out that J does not involve g). We then show that $\hat{J}(X)$ converges weakly to a Brownian sheet. In this way we obtain an asymptotically distribution-free omnibus test for H_0, with the Kolmogorov–Smirnov statistic computed from $\hat{J}(X)$. No simulation technique is needed to find critical values. The test statistic converges weakly to $\sup|W(t, z)|$. Although an exact formula for the distribution function of $\sup|W(t, z)|$ is not known [only approximations are available; see Adler (1991)], it is straightforward to carry out a single Monte Carlo experiment to evaluate it quite accurately. Thus, our test avoids difficulties arising from simulating the null distribution for each particular problem.

A competing procedure would be a bootstrap based test, provided it could be justified theoretically. However, although there is some bootstrap theory available in cases where censoring and covariates are present, none of it applies to our specific problem. Another competing procedure would be an appropriately modified version of the Monte Carlo approach of Lin, Wei and Ying (1993), but we expect that our test would be computationally less demanding.

The paper is organized as follows. In Section 2, we construct the transformation J. In Section 3, we introduce the estimate $\hat{J}$ and define the test statistic. Results of a simulation study are reported in Section 4. In Section 5, the test is applied to a set of data from the British Medical Research Council's (1984) 4th myelomatosis trial. Properties of the test are proved in Section 6. Various lemmas needed through the paper are collected in the Appendix.

2. Transformation of *m* to Brownian sheet. In this section we construct our transformation J of the Gaussian random field m in (1.1) to a Brownian sheet. Such a transformation is likely to have further applications in nonparametric statistics beyond our test for independence—in any setting where a test process converges weakly to a process of the form (1.1); for example, in testing whether $\lambda(t, z)$ is independent of t (the roles of t and z are reversed) or in testing whether a pure jump process on a finite state space is a semi-Markov process; see MU (Section 4.2). Of course, it is usually necessary to estimate J and how that is done will depend on the particular application.

We begin with a key proposition showing that the law of a Brownian sheet W is preserved under a shift by a certain functional of W.

PROPOSITION 2.1. *Let $k \in L^2([0,1]^2)$ satisfy $\int_u^1 k^2(s,v)\,dv > 0$ a.e. $[ds]$ for $u < 1$, and let W be a Brownian sheet. Then*

$$B(t,z) = W(t,z) - \int_0^z \left[\int_0^t \int_u^1 \frac{k(s,x)k(s,u)}{\int_u^1 k^2(s,v)\,dv}\,dW(s,x) \right] du \tag{2.1}$$

is a Brownian sheet on $[0,1]^2$.

PROOF. Let

$$a(t,u;s,x) = \frac{k(s,x)k(s,u)I(x \geq u)I(s \leq t)}{\int_u^1 k^2(s,v)\,dv}.$$

Then

$$B(t,z) = W(t,z) - \int_0^z \left[\int_0^1 \int_0^1 a(t,u;s,x)\,dW(s,x) \right] du.$$

Notice that B is a Gaussian random field, so we only need to inspect its covariance function. For $(t', z') \in [0,1]^2$,

$$\begin{aligned}\operatorname{cov}(B(t,z), B(t',z')) &= (t \wedge t')(z \wedge z') \\ &\quad - \int_0^{z'} \left[\int_0^t \int_0^z a(t',u';s,x)\, ds\, dx \right] du' \\ &\quad - \int_0^{z} \left[\int_0^{t'} \int_0^{z'} a(t,u;s,x)\, ds\, dx \right] du \\ &\quad + \int_0^z \int_0^{z'} \left[\int_0^1 \int_0^1 a(t,u;s,x) a(t',u';s,x)\, ds\, dx \right] du\, du' \\ &= (t \wedge t')(z \wedge z') \\ &\quad + \int_0^z \int_0^{z'} \left[\int_0^1 \int_0^1 a(t,u;s,x) a(t',u',s,x)\, ds\, dx \right. \\ &\qquad\qquad - \int_0^1 a(t,u;s,u') I(s \le t')\, ds \\ &\qquad\qquad \left. - \int_0^1 a(t',u';s,u) I(s \le t)\, ds \right] du'\, du.\end{aligned}$$

Since

$$\begin{aligned}&\int_0^1 a(t,u;s,x) a(t',u';s,x)\, dx \\ &\quad = \frac{k(s,u) I(s \le t) k(s,u') I(s \le t') \int_0^1 k^2(s,x) I(x \ge u \vee u')\, dx}{\int_u^1 k^2(s,v)\, dv \int_{u'}^1 k^2(s,v)\, dv} \\ &\quad = \frac{k(s,u') k(s,u) I(u' \ge u) I(s \le t) I(s \le t')}{\int_u^1 k^2(s,v)\, dv} \\ &\qquad + \frac{k(s,u') k(s,u) I(u \ge u') I(s \le t) I(s \le t')}{\int_{u'}^1 k^2(s,v)\, dv} \\ &\quad = a(t,u;s,u') I(s \le t') + a(t',u',s,u) I(s \le t),\end{aligned}$$

for almost all $(u, u', s) \in [0,1]^3$, we have that B is a Brownian sheet. □

We now give the main result of this section, showing that the process m in (1.1) can be transformed to a Brownian sheet.

THEOREM 2.1. *Suppose that $h\colon [0,1]^2 \to \mathbb{R}$ is a bounded positive measurable function which is bounded away from zero, $b\colon [0,1] \to \mathbb{R}$ is differentiable with square integrable derivative, $\int_z^1 (b'(x))^2\, dx > 0$, $z \in [0,1)$ and $g \in L^2([0,1]^2)$. Then*

$$\begin{aligned}B(t,z) &= \int_0^t \int_0^z h^{-1/2}\, dm \\ &\quad - \int_0^t \int_0^1 \left[\int_0^{u \wedge z} h^{-1/2}(s,u) Q(s,u,x)\, dx \right] dm(s,u)\end{aligned} \tag{2.2}$$

is a Brownian sheet, where

$$Q(s,u,x) = \frac{h^{-1/2}(s,u)b'(u)h^{-1/2}(s,x)b'(x)}{\int_x^1 h^{-1}(s,v)(b'(v))^2\,dv}.$$

PROOF. Notice that

$$\int_0^1\int_0^{u\wedge z} Q(s,u,x)h^{-1/2}(s,u)b'(u)\,dx\,du = \int_0^z h^{-1/2}(s,x)b'(x)\,dx.$$

Let

$$U(t) = \int_0^t\int_0^1 g\,dW.$$

Substituting m into (2.2) we get

$$\begin{aligned}
B(t,z) &= W(t,z) - \int_0^t\left[\int_0^z h^{-1/2}(s,x)b'(x)\,dx\right]U(ds)\\
&\quad - \int_0^t\int_0^1\left[\int_0^{u\wedge z} Q(s,u,x)\,dx\right]dW(s,u)\\
&\quad + \int_0^t\int_0^1\left[\int_0^{u\wedge z} Q(s,u,x)\,dx\right]h^{-1/2}(s,u)b'(u)\,du\,U(ds)\\
&= W(t,z) - \int_0^t\int_0^1 g(s,y)\left[\int_0^z h^{-1/2}(s,x)b'(x)\,dx\right]dW(s,y)\\
&\quad - \int_0^z\left[\int_0^t\int_x^1 Q(s,y,x)\,dW(s,y)\right]dx\\
&\quad + \int_0^t\int_0^1\left[g(s,y)\int_0^1\int_0^{u\wedge z} Q(s,u,x)h^{-1/2}(s,u)b'(u)\,dx\,du\right]\\
&\quad \times dW(s,y)\\
&= W(t,z) - \int_0^z\left[\int_0^t\int_x^1 Q(s,y,x)\,dW(s,y)\right]dx.
\end{aligned} \tag{2.3}$$

This is a Brownian sheet by Proposition 2.1 with $k(s,x) = h^{-1/2}(s,x)b'(x)$. □

We shall use the notation J for the transformation $\xi \mapsto J(\xi)$, where ξ is a random field and $J(\xi)$ is defined by the right side of (2.2) with m replaced by ξ. The domain of J is composed of random fields ξ for which the stochastic integrals in $J(\xi)$ exist in the L^2-sense. Theorem 2.1 shows that $J(m)$ is a Brownian sheet.

Notice that J does not involve the function g; phenomena like this are typical of the innovation approach; compare goodness-of-fit testing for parametric hazard function models [Andersen, Borgan, Gill and Keiding (1993), formulae (6.3.16) and (6.3.27)].

3. The test procedure. In this section we first describe the counting process framework for our problem and formally define $\hat{\mathscr{A}}$ and $\bar{\mathscr{A}}$. Then we show that the transformation J given above asymptotically transforms $X = \sqrt{n}(\hat{\mathscr{A}} - \bar{\mathscr{A}})$ to a Brownian sheet. This is done via the continuous mapping theorem. Finally, we construct an estimate $\hat{J}$ of J and show that $\hat{J}(X)$ converges weakly to a Brownian sheet. This will complete the construction of our test.

3.1. *The estimators $\hat{\mathscr{A}}$ and $\bar{\mathscr{A}}$.* Let $\mathbf{N}(t) = (N_1(t), \dots, N_n(t))$, $t \in [0,1]$, be a multivariate counting process with respect to a right-continuous filtration $(\mathscr{F}_t)$, that is, $\mathbf{N}$ is adapted to the filtration and has components N_i which are right-continuous step functions, zero at time zero, with jumps of size $+1$ such that no two components jump simultaneously. Assume that N_i has intensity

$$\lambda_i(t) = Y_i(t)\lambda(t, Z_i(t)),$$

where Y_i is a predictable $\{0,1\}$-valued process, indicating that the ith individual is at risk when $Y_i(t) = 1$, and Z_i is a predictable $[0,1]$-valued covariate process. The function $\lambda(t,z)$ represents the failure rate for an individual at time t with covariate $Z_i(t) = z$. We assume throughout that (N_i, Y_i, Z_i), $i = 1, \dots, n$, are iid replicates of an underlying triple (N, Y, Z). Let $F(s,x) = P(Z_s \le x, Y_s = 1)$, and assume that for each $s \in [0,1]$, $F(s, \cdot)$ is absolutely continuous on $[0,1]$ with subdensity $f(s,\cdot)$. The functions b, h, g in (1.1) are given by $b(z) = z$, $h = \lambda/f$ and $g = \sqrt{\lambda \cdot f}$. The transformation J will only be used with these b and h from now on. We assume that f and λ are Lipschitz, of bounded variation and bounded away from zero.

Consider d_n equal width covariate strata $\mathscr{I}_r = [x_{r-1}, x_r)$, $r = 1, \dots, d_n$, where $x_r = rw_n$ and $w_n = 1/d_n$ is the stratum width, and let $\mathscr{I}_z = \mathscr{I}_r$ for $z \in \mathscr{I}_r$. As in MU, we estimate $\mathscr{A}$ by integrating the "covariate stratum-specific" Nelson–Aalen estimator to obtain

$$\hat{\mathscr{A}}(t,z) = \int_0^z \int_0^t \frac{N^{(n)}(ds, x)}{Y^{(n)}(s,x)}\,dx, \tag{3.1}$$

where $N^{(n)}(t,z) = \sum_{i=1}^n \int_0^t I(Z_i(s) \in \mathscr{I}_z)\,dN_i(s)$ is the number of z-specific failures observed up to time t and $Y^{(n)}(t,z) = \sum_{i=1}^n I(Z_i(t) \in \mathscr{I}_z)Y_i(t)$ is the size of the z-specific risk set at time t. The estimator $\bar{\mathscr{A}}$ does not involve stratification of the covariate and can be obtained by setting $w_n = 1$ in $\hat{\mathscr{A}}$. In (3.1) and throughout the paper, we use the convention $1/0 \equiv 0$.

3.2. *A continuous version of J.* We now introduce a version $\bar{J}$ of J that is defined on a suitably large function space and is continuous on a subspace supporting m, so the continuous mapping theorem is applicable.

Let $D_2 = D_2([0,1]^2)$ be the extension of the usual Skorohod space to functions on $[0,1]^2$; see Neuhaus (1971). Let BV_2 denote the subspace of functions $\xi \in D_2$ for which ξ, $\xi(0,\cdot)$, $\xi(\cdot,0)$ have bounded variation and let

C_2 denote the space of continuous functions on $[0,1]^2$. Equip C_2 with the uniform norm.

For $\xi \in C_2 \cup BV_2$ and $(t, z) \in [0,1] \times [0, \rho]$, with $0 < \rho < 1$, define

$$(3.2)\quad \bar{J}(\xi)(t,z) = \int_0^t \int_0^z f_1(s,x)\, d\xi(s,x) - \int_0^t \int_0^1 f_2(s,u,z)\, d\xi(s,u),$$

where the integrals are considered to be weak net integrals [Hildebrandt (1963), Section III.8] for which integration by parts works as expected and

$$f_1(s,x) = h^{-1/2}(s,x),$$
$$f_2(s,u,z) = h^{-1}(s,u) \int_0^{z \wedge u} \frac{h^{-1/2}(s,x)}{\int_x^1 h^{-1}(s,v)\, dv}\, dx.$$

The upper bound ρ on the domain of z is used to keep the denominator in f_2 bounded away from zero. Note that $\bar{J}$ is a well-defined map from $C_2 \cup BV_2$ into $D_2([0,1] \times [0, \rho])$ since Lemma 1 in the Appendix shows that h inherits the properties of f, λ, and Lemma 2 in the Appendix ensures the existence of the weak net integrals when $\xi \in C_2$. We have included BV_2 in the domain of $\bar{J}$ because the paths of X belong to BV_2, but not to C_2.

THEOREM 3.1. *Suppose that* $w_n \to 0$, $nw_n^2 \to 0$ *and* $nw_n^{1+\delta} \to \infty$ *for some* $0 < \delta < 1$. *Then, under* H_0, $\bar{J}(X)$ *converges weakly to a Brownian sheet in* $D_2([0,1] \times [0, \rho])$.

PROOF. Properties of f_1, f_2 obtained via Lemmas 1 and 2 in the Appendix can be used to show that $\bar{J}$ is continuous as a map from C_2 into $D_2([0,1] \times [0, \rho])$. In particular, we use the property that $f_2(\cdot, \cdot, z)$ has bounded variation uniformly in z, $0 < z < \rho$. MU (Theorem 4.1) gives that X converges weakly in D_2 to m, where m is defined by (1.1) with $b(z) = z$. Thus, since the sample paths of m belong to C_2 a.s., the continuous mapping theorem [Billingsley (1968)] gives $\bar{J}(X) \to_{\mathscr{D}} \bar{J}(m)$ in $D_2([0,1] \times [0, \rho])$. The processes $\bar{J}(m)$ and $J(m)$ have continuous sample paths and, by Lemma 3, they agree a.s. at each fixed (t, z), so they are indistinguishable. Theorem 2.1 [with $b(z) = z$] implies that $J(m)$, and hence $\bar{J}(m)$, is a Brownian sheet. □

3.3. *Estimating the transformation.* In order to use the above result to build a test statistic, we need to estimate the unknown function in $\bar{J}$, namely h. First consider the kernel estimator $\hat{h}$ suggested by MU:

$$\hat{h}(t,z) = \frac{1}{b_n^2} \int_0^1 \int_0^1 K\left(\frac{t-s}{b_n}\right) K\left(\frac{z-x}{b_n}\right) d\hat{H}(s,x),$$

where b_n is a bandwidth parameter, K is a Lipschitz nonnegative kernel function with compact support and integral 1, and

$$\hat{H}(t,z) = nw_n \int_0^z \int_0^t \frac{N^{(n)}(ds,x)}{\left(Y^{(n)}(s,x)\right)^2}\,dx$$

is an estimator of $H(t,z) = \int_0^t \int_0^z h(s,x)\,ds\,dx$. Here $\hat{h}$ is a smoothed version of $\hat{H}$.

We will need to apply methods from stochastic calculus to various martingale integrals involving $\hat{h}$, which is possible provided that $\hat{h}(\cdot, z)$ is an $\mathscr{F}_t$-predictable process for each fixed z. Since $\hat{h}(\cdot, z)$ is continuous, it is enough that it be adapted to the filtration $\mathscr{F}_t$. Thus, we shall use a kernel function K having nonnegative (as well as compact) support.

The estimated transformation $\hat{J}$ is defined by inserting a truncated version $\check{h}$ of $\hat{h}$ in place of h in $\bar{J}$, where $\check{h}$ is given by

$$\check{h}(t,z) = \left(c_n^{-1} \wedge \hat{h}(t,z)\right) \vee c_n,$$

$c_n > 0$. The truncation is needed to prevent instability in $\hat{J}$. Note that $\hat{J}(X)$ is well defined since the paths of X belong to BV_2.

3.4. *The test statistic*. If we show that $\hat{J}(X)$ converges weakly to a Brownian sheet, then our test for H_0 can be based on the Kolmogorov–Smirnov statistic

$$S = \sup_{0 \le t \le 1,\, 0 \le z \le \rho} \left| \hat{J}\left(\sqrt{n}\,(\hat{\mathscr{A}} - \mathscr{A})\right)(t,z) \right|$$

with P-values calculated from the distribution of

$$S^* = \sup_{0 \le t \le 1,\, 0 \le z \le \rho} |W(t,z)|.$$

For that purpose we restrict the choice of w_n, b_n, c_n as follows: $w_n \asymp n^{-\alpha}$, $b_n \asymp n^{-\beta}$, $c_n \asymp n^{\gamma}$, where the following condition holds.

CONDITION 3.1.

$$\tfrac{1}{2} < \alpha < 1,$$
$$0 < \beta < \min\{\tfrac{1}{3}\alpha, \tfrac{1}{2}(1-\alpha)\},$$
$$0 < \gamma < \tfrac{1}{5}\min\{\tfrac{2}{3}\beta, 2\alpha - 6\beta, 1 - \alpha - 2\beta\}.$$

This condition is satisfied, for example, by $\alpha = 5/9$, $\beta = 1/6$, $\gamma = 1/46$. The following result implies that S converges weakly to S^*. The distribution of S^* can be found quite accurately by simulation. Bounds on the tail of the distribution of S^* are given in Adler (1991).

THEOREM 3.2. *Under H_0, $\hat{J}(X)$ converges weakly to a Brownian sheet in* $D_2([0,1] \times [0,\rho])$.

The restriction to $[0, \rho]$ is used to avoid instability in the estimate of the denominator in f_2. The test statistic S will only be affected by the choice of ρ

when $|\hat{J}(X)|$ attains its supremum in the small strip $[0, 1] \times (\rho, 1]$, and that is unlikely for ρ close to 1. In applications it is worthwhile to plot the transformed process over the whole of $[0, 1]^2$, as we have done in Figure 2.

Our final result shows that the test based on S is omnibus—consistent against any departure from the null hypothesis H_0.

THEOREM 3.3. *The test based on S is consistent against the general alternative that $\lambda(t, z)$ depends on z, for (t, z) in the domain $[0, 1] \times [0, \rho]$.*

4. A simulation study. We carried out a simulation study to assess the performance of the proposed test. We considered the Kolmogorov–Smirnov statistic S with the supremum taken over $[0, 1] \times [0, 0.9]$, that is, $\rho = 0.9$. The covariate was taken to be uniformly distributed over $[0, 1]$. The censoring was simple right censoring, independent of the failure time, and exponentially distributed, with the parameter adjusted to give a prescribed proportion (moderate: 27%; heavy: 60%) of censored observations (including those lost to followup at time 1). The covariate strata were arranged to contain equal numbers of observations. For sample sizes 500 and 1000, the number of strata d_n was taken to be 10 and 14, resulting in about 50 and 71 covariate values per stratum. The corresponding bandwidths b_n were taken as 0.32 and 0.18, and the kernel function K was taken to be the indicator of $[0, 1]$.

The survival times were generated using the Cox model $\lambda(t, z) = \exp(\beta_0 z)$, for $\beta_0 = 0$ (null hypothesis) and $\beta_0 = 1, 2$ (alternative hypotheses), and using the non-Cox model $\lambda(t, z) = 7.5 \min(z, 1 - z)$. Table 1 gives observed levels and powers of the test at a nominal (asymptotic) level 5%, with each entry based on 1000 samples. The corresponding values for the Cox model based (Wald type) test of $\beta_0 = 0$ are given in parentheses. In order to obtain the asymptotic 5% critical level for our test (i.e., the 95th percentile of $\sup_{0 \le t \le 1, 0 \le z \le 0.9} |W(t, z)|$), we generated 10,000 replicates of the Brownian

TABLE 1
Observed levels and powers of the proposed test. Corresponding values for the Cox model based test are given in parentheses. Nominal level is 5%

$\lambda(t, z)$	Sample Size	Censoring	
		27%	60%
1	500	0.044 (0.060)	0.034 (0.044)
	1000	0.052 (0.056)	0.039 (0.051)
e^z	500	0.207 (1.00)	0.090 (0.983)
	1000	0.471 (1.00)	0.228 (1.00)
e^{2z}	500	0.455 (1.00)	0.212 (1.00)
	1000	0.819 (1.00)	0.484 (1.00)
$7.5 \min(z, 1 - z)$	500	0.393 (0.056)	0.187 (0.031)
	1000	0.997 (0.060)	0.794 (0.043)

sheet evaluated on a grid defined by 300 equally spaced points on each axis. The 5% critical level was found to be 2.2811.

The observed levels of our test are close to their nominal 5% values at sample sizes 500 and 1000 when censoring is moderate. The test appears to be slightly conservative under heavy censoring. In general we recommend that our test only be used for sample size at least 500, and at least 1000 under heavy censoring. Under the Cox model our test is naturally much less powerful (with 82% power for $\beta_0 = 2$, $n = 1000$ and moderate censoring) than the Cox model based test (with almost 100% power, even for $\beta_0 = 1$, $n = 500$ and heavy censoring). However, our test has adequate power away from the Cox model, as an omnibus test should, whereas the Cox model based test can have very poor power; see the last two rows of Table 1.

Some quantile–quantile plots of the observed distribution of S against the distribution of S^* are shown in Figure 1. Each plot refers to a combination of model and censoring level, and contains curves for sample sizes 500 and 1000. Under $\lambda(t, z) = 1$, the curves are close to the diagonal. This indicates that the observed distribution of S is close to its asymptotic null distribution. Under the alternatives e^{2z} and 7.5 $\min(z, 1 - z)$, when the sample size is 1000, the curves lie well above the diagonal, giving some idea of the power of the proposed test.

5. Application to myelomatosis data. We applied our test to a set of data from the British Medical Research Council's (BMRC) (1984) 4th myelomatosis trial. The data set contains records for 495 patients, including censoring indicator, serum β_2 microglobulin (at presentation) and survival time (in days).

Many studies [e.g., Cuzick, Cooper and MacLennan (1985)] have suggested that serum β_2 microglobulin has a strong effect on survival, at least in the first two years of followup. In our analysis of the data we ignore all covariates except for serum β_2 microglobulin (which is taken on a log scale). We standardized this covariate by its sample mean and sample standard deviation, then transformed it by the standard normal distribution function. The resulting covariate values were then more or less uniformly distributed over $[0, 1]$. (As a general rule we recommend transforming the covariate to uniform, as it helps stabilize the various estimators at points where the covariate data are sparse. Also, our simulation study provides strong support for the accuracy of our test when the covariates are uniform.) The end of followup is taken to be 2000 days, before which 3% of the observations are censored; 81 patients were still at risk at the end of followup. Each covariate stratum was arranged to contain 50 covariate values except for the last stratum.

We have plotted the test process $\hat{J}(X)$ over the whole unit square (see Figure 2); for comparison the untransformed process X is plotted underneath. The magnitude of the negative part of $\overline{J}(X)$ suggests strong departure from a Brownian sheet. The statistic S was found to be 2.41, giving a P-value of 0.033. Thus our analysis confirms that serum β_2 microglobulin has a significant influence on survival.

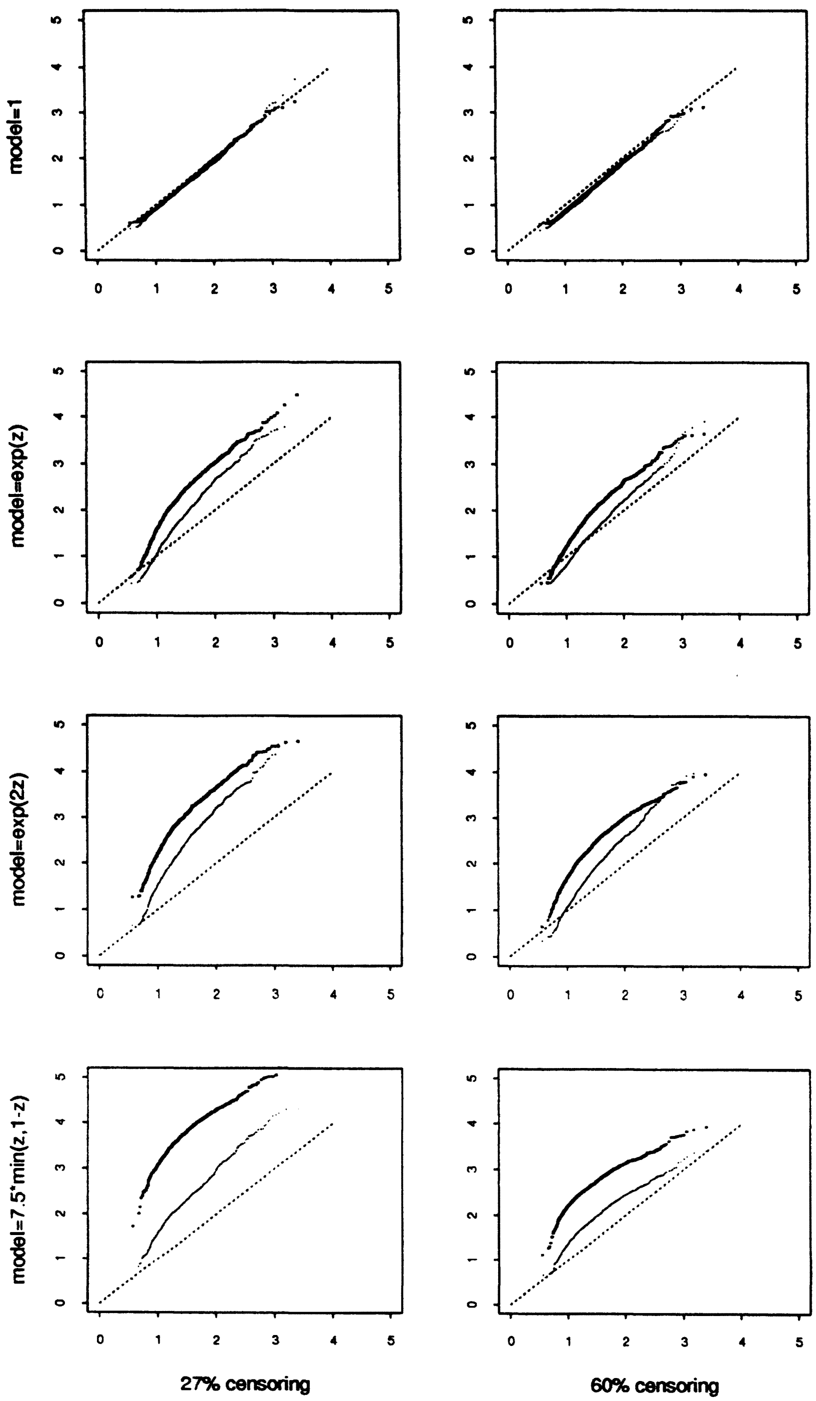

FIG. 1. *Q-Q plots of the observed distribution of S against the distribution of S^* at sample sizes* 500 (*thick lines*) *and* 1000 (*thin lines*).

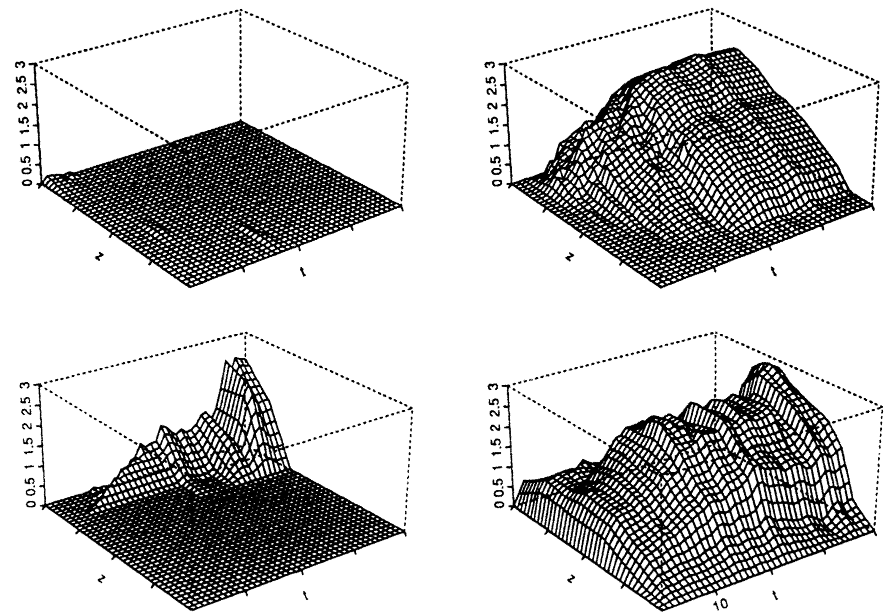

FIG. 2. *The test process $\hat{J}(X)$ for the BMRC data (first row) and the corresponding untransformed process X (second row). Positive parts are on the left; negative parts are on the right.*

Note that the test process achieved its supremum well away from the edge $z = 1$, so in this case S does not vary with ρ when ρ is close to 1 (although our results require $\rho < 1$). Nevertheless, the transformation $\hat{J}$ seems to have its greatest effect around $z = 1$ since the bump in the positive part of X is missing from $\hat{J}(X)$.

6. Proofs. In this section we prove Theorems 3.2 and 3.3. We begin by introducing some notation. Let M_i denote the $\mathscr{F}_t$-martingale $M_i(t) = N_i(t) - \int_0^t \lambda_i(s)\,ds$ and set

$$M^{(n)}(t, z) = \sum_{i=1}^{n} \int_0^t I(Z_i(s) \in \mathscr{I}_z)\, dM_i(s),$$

$$\lambda^{(n)}(t, z) = \sum_{i=1}^{n} I(Z_i(s) \in \mathscr{I}_z) Y_i(t)\, \lambda(t, Z_i(t)).$$

For a process $\xi(t, z)$, set $\xi_r(t) = \xi(t, x_r)$, where $x_r = rw_n$, $r = 1, \dots, d_n$.

We shall have frequent use for the following bounds from MU (Lemma 1):

$$\sup_{s,x,n} E\left[\frac{nw_n}{Y^{(n)}(s, x)}\right]^k < \infty \quad \text{for any positive integer } k, \tag{6.1}$$

$$\sup_{s,x} P\big(Y^{(n)}(s, x) = 0\big) \le \exp(-Cnw_n) \quad \text{for some } C > 0. \tag{6.2}$$

PROOF OF THEOREM 3.2. By Theorem 3.1, it is sufficient to show that under H_0,

$$\|(\bar{J} - \hat{J})(X)\| \to_P 0,$$

where $\|\cdot\|$ is the supremum norm on $D_2([0,1] \times [0, \rho])$. This will be done in the following two steps:

$$\left\| \int_0^{\cdot} \int_0^{\cdot} \bar{f}_1 \, dX \right\| \to_P 0, \tag{6.3}$$

$$\left\| \int_0^{\cdot} \int_0^1 \bar{f}_2(s, u, \cdot) \, dX(s, u) \right\| \to_P 0, \tag{6.4}$$

where $\bar{f}_i = \check{f}_i - f_i$ and $\check{f}_i$ is obtained by inserting $\check{h}$ in place of h in f_i, $i = 1, 2$.

Step 1. By the decomposition of X given in MU (proof of Theorem 4.1),

$$\begin{aligned}
\int_0^t \int_0^z \bar{f}_1 \, dX &= \sqrt{n} \int_0^z \int_0^t \bar{f}_1(s, x) \frac{M^{(n)}(ds, x)}{Y^{(n)}(s, x)} \, dx \\
&\quad - \sqrt{n} \int_0^z \int_0^t \bar{f}_1(s, x) \frac{d\overline{M}^{(n)}(s)}{\overline{Y}^{(n)}(s)} \, dx \\
&\quad + \int_0^t \int_0^z \bar{f}_1(s, x) \sqrt{n} \left(\frac{\lambda^{(n)}(s, x)}{Y^{(n)}(s, x)} - \lambda(s, x) \right) dx \, ds \\
&\quad + \sqrt{n} \int_0^t \int_0^z \bar{f}_1(s, x) \lambda(s) I\big(\overline{Y}^{(n)}(s) = 0\big) \, dx \, ds,
\end{aligned} \tag{6.5}$$

where $\overline{M}^{(n)}, \overline{Y}^{(n)}$ are defined by setting $\mathscr{I}_z = [0,1]$ in $M^{(n)}, Y^{(n)}$, respectively. We denote the four terms in the above decomposition by I_1, I_2, I_3 and I_4, respectively. Since K is continuous and has nonnegative support, we have $\hat{h}(\cdot, x)$, and therefore $\bar{f}_1(\cdot, x)$, is $\mathscr{F}_t$-predictable. Thus the stochastic integrals involved in I_1 and I_2 are square integrable martingales. Now $\|I_1\|$ is bounded by

$$\sup_t \eta(t) + \sqrt{n} \sup_{\substack{t, r \\ z \in \mathscr{I}_r}} \left| \int_{x_{r-1}}^z \int_0^t \bar{f}_1 \frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)} \, dx \right|, \tag{6.6}$$

where

$$\eta(t) = \sup_{1 \le j \le d_n} \left| \sum_{r=1}^j \xi(t, r) \right| \quad \text{and} \quad \xi(t, r) = \sqrt{n} \int_0^t \left(\int_{\mathscr{I}_r} \bar{f}_1 \, dx \right) \frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)}.$$

Since $\eta(t)$ is a positive submartingale, Doob's inequality gives $E \sup_t \eta^2(t) \le 4E\eta^2(1)$. Also, since $E\xi(1, r) = 0$, and $E\xi(1, j)\xi(1, k) = 0$ for all $1 \le j \ne k \le$

d_n, we can apply Menchoff's inequality [see, e.g., Shorack and Wellner (1986)] here to get

$$
\begin{aligned}
E\eta^2(1) &\le \left(\frac{\log 4d_n}{\log 2}\right)^2 \sum_{r=1}^{d_n} E\int_0^1 \left(\int_{\mathscr{I}_r} \bar{f}_1(s,x)\,dx\right)^2 \frac{n\lambda_r^{(n)}(s)}{\left(Y_r^{(n)}(s)\right)^2}\,ds \\
&\le O(\log d_n)^2 \sum_{r=1}^{d_n} \int_0^1 E\left[\left(\int_{\mathscr{I}_r} \bar{f}_1(s,x)\,dx\right)^2 \frac{n}{Y_r^{(n)}(s)}\right] ds \\
&\le O(\log d_n)^2 \sum_{r=1}^{d_n} \int_0^1 E\left[\int_{\mathscr{I}_r} \bar{f}_1^2(s,x)\,dx \frac{nw_n}{Y_r^{(n)}(s)}\right] ds \\
&\le O(\log d_n)^2 \left[\sum_{r=1}^{d_n} \int_0^1 E\left(\int_{\mathscr{I}_r} \bar{f}_1^2(s,x)\,dx\right)^{3/2} ds\right]^{2/3} \qquad (6.7)\\
&\quad \times \left[\sum_{r=1}^{d_n} \int_0^1 E\left(\frac{nw_n}{Y_r^{(n)}(s)}\right)^3 ds\right]^{1/3} \\
&\le O(\log d_n)^2 d_n^{1/3} \left[\sum_{r=1}^{d_n} \int_0^1 E\left(w_n^{1/2} \int_{\mathscr{I}_r} \left|\bar{f}_1(s,x)\right|^3 dx\right) ds\right]^{2/3} \\
&= O(\log d_n)^2 \left[\int_0^1 \int_0^1 E|\bar{f}_1|^3\,ds\,dx\right]^{2/3}.
\end{aligned}
$$

The second term in (6.6) is bounded by

$$\sqrt{n}\,\sup_r \int_{\mathscr{I}_r} \sup_t \left|\int_0^t \bar{f}_1 \frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)}\right| dx,$$

which has second moment bounded by

$$
\begin{aligned}
nw_n E \sup_r \int_{\mathscr{I}_r} \sup_t \left|\int_0^t \bar{f}_1 \frac{M_r^{(n)}(ds)}{Y_r^{(n)}(s)}\right|^2 dx &\le nw_n \int_0^1 E \sup_t \left|\int_0^t \bar{f}_1 \frac{M^{(n)}(ds,x)}{Y^{(n)}(s,x)}\right|^2 dx \\
&\le O(1) \int_0^1 \int_0^1 E\left[\bar{f}_1^2 \frac{nw_n}{Y^{(n)}(s,x)}\right] ds\,dx \qquad (6.8)\\
&\le O(1) \left[\int_0^1 \int_0^1 E|\bar{f}_1|^3\,ds\,dx\right]^{2/3},
\end{aligned}
$$

where Doob's inequality, Hölder's inequality and (6.1) are used. Therefore,

$$E\|I_1\|^2 \le O(\log d_n)^2 \left[\int_0^1 \int_0^1 E|\bar{f}_1|^3\,ds\,dx\right]^{2/3}. \qquad (6.9)$$

From (6.9) with $w_n = 1$,

$$E\|I_2\|^2 = O(1)\left[\int_0^1\int_0^1 E|\bar{f}_1|^3\, ds\, dx\right]^{2/3}. \tag{6.10}$$

Next,

$$\begin{aligned} E\|I_3\|^2 &\le \sup_{t,z} E\left[\sqrt{n}\left(\frac{\lambda^{(n)}(s,x)}{Y^{(n)}(s,x)} - \lambda(s,x)\right)\right]^2 \left[\int_0^1\int_0^1 E\bar{f}_1^2\, ds\, dx\right] \\ &= o(1)\int_0^1\int_0^1 E|\bar{f}_1|^2\, ds\, dx, \end{aligned} \tag{6.11}$$

by Lemma 6. Using (6.1) once more,

$$E\|I_4\|^2 = o(1)\int_0^1\int_0^1 E|\bar{f}_1|^2\, ds\, dx. \tag{6.12}$$

It can be checked that

$$|\bar{f}_1| \le O(c_n)|\hat{h} - h| I\left(c_n^{-1} \le \hat{h} \le c_n\right) + O\left(c_n^{1/2}\right) I\left(\hat{h} < c_n^{-1} \text{ or } \hat{h} > c_n\right),$$

uniformly in t, z. Thus,

$$E|\bar{f}_1|^3 \le O\left(c_n^3\right) E|\hat{h} - h|^3 + O\left(c_n^{3/2}\right) P\left(\hat{h} < c_n^{-1} \text{ or } \hat{h} > c_n\right),$$

so, from Lemmas 4 and 5,

$$\int_0^1\int_0^1 E|\bar{f}_1|^3\, ds\, dx = O(1)c_n^3\left[b_n + w_n^3 b_n^{-9} + \left(n w_n b_n^2\right)^{-3/2}\right]. \tag{6.13}$$

Combining the bounds (6.9)–(6.13), we find that the second moment of the lhs of (6.3) is of order $O(1)(\log d_n)^2 c_n^2[b_n^{2/3} + w_n^2 b_n^{-6} + (n w_n b_n^2)^{-1}]$, which tends to zero by Condition 3.1. This establishes (6.3).

Step 2. We now prove (6.4). Let

$$\delta(s,u,x) = \frac{\check{h}^{-1}(s,u)\check{h}^{-1/2}(s,x)}{\int_x^1 \check{h}^{-1}(s,v)\, dv} - \frac{h^{-1}(s,u)h^{-1/2}(s,x)}{\int_x^1 h^{-1}(s,v)\, dv}. \tag{6.14}$$

By the arguments of Step 1, the second moment of the lhs of (6.4) is bounded by

$$\begin{aligned} &\int_0^\rho E \sup_t \left|\int_0^t\int_x^1 \delta(s,u,x)\, dX(s,u)\right|^2 dx \\ &\quad\le \int_0^\rho \left[\int_0^1\int_0^1 E|\delta(s,u,x)|^3\, ds\, du\right]^{2/3} dx \\ &\quad\le \left[\int_0^\rho\int_0^1\int_0^1 E|\delta(s,u,x)|^3\, ds\, du\, dx\right]^{2/3} \\ &\quad= O(1)c_n^5\left[b_n^{2/3} + w_n^2 b_n^{-6} + \left(n w_n b_n^2\right)^{-1}\right] \to 0, \end{aligned}$$

where the bound on the triple integral is from Lemma 7. □

PROOF OF THEOREM 3.3. Define

$$\mathscr{A}^*(t, z) = z\int_0^t\int_0^1 \lambda(s, x) f(s, x)\, dx\, ds,$$

to which $\bar{\mathscr{A}}(t, z)$ converges in probability under the general alternative. From the definition of $\bar{J}$ in (3.2), it is easily checked that

$$\frac{\partial^2 \bar{J}(\mathscr{A} - \mathscr{A}^*)}{\partial t \partial z} = h^{-1/2}(t, z)\left[\lambda(t, z) - \frac{\int_z^1 h^{-1}(t, u)\lambda(t, u)\, du}{\int_z^1 h^{-1}(t, v)\, dv}\right].$$

Suppose that $\bar{J}(\mathscr{A} - \mathscr{A}^*) = 0$. Then the expression inside the square brackets above vanishes, so that

$$\lambda(t, z)\int_z^1 h^{-1}(t, v)\, dv = \int_z^1 h^{-1}(t, u)\lambda(t, u)\, du.$$

Taking partial derivatives wrt z both sides gives that $\partial\lambda(t, z)/\partial z = 0$ for $(t, z) \in [0, 1] \times [0, \rho]$, so that H_0 holds, contrary to the premise of the theorem. Thus, $\bar{J}(\mathscr{A} - \mathscr{A}^*) \neq 0$. From arguments in the proof of Theorem 3.2, it can be seen that $\|(\hat{J} - \bar{J})(\mathscr{A} - \mathscr{A}^*)\| \to_P 0$. Hence,

$$\|\hat{J}(\mathscr{A} - \mathscr{A}^*)\| \to_P \|\bar{J}(\mathscr{A} - \mathscr{A}^*)\| > 0. \tag{6.15}$$

Along the lines of the proof of Theorem 3.2, it can be shown that $\hat{J}(\sqrt{n}(\bar{\mathscr{A}} - \mathscr{A}^*))$ converges weakly in $D_2([0, 1] \times [0, \rho])$, although not necessarily to a Brownian sheet; compare the proof of Proposition 4.3 of MU. Similarly, using MU (Proposition 3.2), it can be shown that $\hat{J}(\sqrt{n}(\hat{\mathscr{A}} - \mathscr{A}))$ converges in the same sense. The triangle inequality gives

$$\sqrt{n}\,\|\hat{J}(\mathscr{A} - \mathscr{A}^*)\| \le \left\|\hat{J}\left(\sqrt{n}\,(\hat{\mathscr{A}} - \mathscr{A})\right)\right\| + S + \left\|\hat{J}\left(\sqrt{n}\,(\mathscr{A} - \mathscr{A}^*)\right)\right\| = S + O_P(1),$$

the equality holding by the continuous mapping theorem, so that $S \to_P \infty$ by (6.15). Thus the test is consistent. □

APPENDIX

The following lemma is routine.

LEMMA 1. *Let h, h_1, h_2 be functions on $[0, 1]^2$ that have bounded variation and are Lipschitz, with h nonnegative and bounded away from zero. Then $1/h$, $\sqrt{h_1}$ and $h_1 h_2$ have bounded variation and are Lipschitz.*

The next two lemmas collect some properties of weak net integrals in the plane. The first is a version of the integration by parts formula. Let $\phi: \mathscr{T} \to \mathbb{R}$, where $\mathscr{T} = [a, a'] \times [b, b']$.

LEMMA 2. *If $\xi\colon \mathscr{T} \to \mathbb{R}$ is continuous and ϕ, $\phi(a, \cdot)$ and $\phi(\cdot, b)$ have bounded variation, then the weak net integral $\int\int_{\mathscr{T}} \phi \, d\xi$ exists and is equal to*

$$\begin{aligned}(\text{A.1}) \qquad &\int\int_{\mathscr{T}} d(\phi(s,x)\xi(s,x)) + \int\int_{\mathscr{T}} \xi(s,x)\, d\phi(s,x) \\ &\quad - \int_a^{a'} \xi(s,b')\, d\phi(s,b') \\ &\quad + \int_a^{a'} \xi(s,b)\, d\phi(s,b) - \int_b^{b'} \xi(a',x)\, d\phi(a',x) \\ &\quad + \int_b^{b'} \xi(a,x)\, d\phi(a,x).\end{aligned}$$

PROOF. Theorem III.9.3 of Hildebrandt (1963) gives that the weak net integral

$$(\text{A.2}) \qquad \int\int_{\mathscr{T}} (\xi(s,x) - \xi(s,b) - \xi(a,x) + \xi(a,b))\, d\phi(s,x)$$

exists, and coincides with the weak net integral

$$(\text{A.3}) \qquad \int\int_{\mathscr{T}} (\phi(s,x) - \phi(s,b') - \phi(a',x) + \phi(a',b'))\, d\xi(s,x),$$

which exists by Theorem III.8.8 of Hildebrandt. Theorem III.5.8 of Hildebrandt shows that $\phi(\cdot, x)$ and $\phi(s, \cdot)$ are of bounded variation for fixed s, x. (A.1) can then be obtained by rearranging the terms in (A.2) and (A.3). □

LEMMA 3. *Let ξ be a stochastic process on $\mathscr{T}$. If the weak integral $\int\int_{\mathscr{T}} \phi \, d\xi$ exists a.s., and the stochastic integral $\int\int_{\mathscr{T}} \phi \, d\xi$ exists in the L^2-sense, then they coincide a.s.*

PROOF. The result follows immediately from the definitions of the stochastic integral and the weak net integral, and the fact that an L^2-limit agrees almost surely with an a.s. limit. □

The next lemma is a refined version of Proposition 3.3 of MU, giving a *rate* of convergence of $\hat{h}$ to h.

LEMMA 4.

$$\int_0^1\int_0^1 E\left|\hat{h}(t,z) - h(t,z)\right|^3 dt\, dz = O(1)\left[b_n + w_n^3 b_n^{-9} + \left(n w_n b_n^2\right)^{-3/2}\right].$$

PROOF. We shall use much of the notation of MU (proof of Proposition 3.3), without redefining it here. As in MU,

$$\begin{aligned}(\text{A.4}) \qquad |\hat{h} - h|^3 \le O(1)\Big[&|\hat{h} - \tilde{h}|^3 + |h - h^0|^3 \\ &+ |h^0 - h^\dagger|^3 + |h^\dagger - h^*|^3 + |R|^3\Big].\end{aligned}$$

For the first term,

$$\begin{aligned}
&\sup_{t,z}\left|\hat{h}(t,z)-\tilde{h}(t,z)\right|^3 \\
\text{(A.5)}\qquad &\le \sup_{t,z}\left[\frac{1}{b_n^2}\sum_{j,r}K\left(\frac{t-\tau_j}{b_n}\right)\left|K\left(\frac{z-x_r}{b_n}\right)-\frac{1}{w_n}\int_{\mathscr{I}_r}K\left(\frac{z-x}{b_n}\right)dx\right|\Delta_{jr}\right]^3 \\
&= O\left(w_n^3 b_n^{-9}\right)\tilde{H}^3(1,1).
\end{aligned}$$

Also,

$$\begin{aligned}
E\tilde{H}^3(1,1) &= E\left[nw_n\int_0^1\int_0^1\frac{N^{(n)}(ds,x)}{\left(Y^{(n)}(s,x)\right)^2}\,dx\right]^3 \\
\text{(A.6)}\qquad &\le 8E\left[nw_n\int_0^1\int_0^1\frac{\lambda^{(n)}(s,x)}{\left(Y^{(n)}(s,x)\right)^2}\,ds\,dx\right]^3 \\
&\quad + 8(nw_n)^3E\left|\int_0^1\int_0^1\frac{M^{(n)}(ds,x)}{\left(Y^{(n)}(s,x)\right)^2}\,dx\right|^3.
\end{aligned}$$

The first term in (A.6) can be shown to be bounded using (6.1). The expectation in the second term in (A.6) is bounded by

$$\text{(A.7)}\qquad \int_0^1 E\left|\int_0^1\frac{M^{(n)}(ds,x)}{\left(Y^{(n)}(s,x)\right)^2}\right|^3 dx \le \int_0^1\left(EM_1^4(1)\right)^{3/4}dx,$$

where

$$M_1(t)=\int_0^t\frac{M^{(n)}(ds,x)}{\left(Y^{(n)}(s,x)\right)^2}$$

and we have suppressed the dependence of $M_1(t)$ on x. Let $[M_1]$ and $\langle M_1\rangle$ be the quadratic variation and the predictable quadratic variation of martingale M_1, respectively. We shall use the Burkholder–Davis–Gundy inequality [Dellacherie and Meyer (1982), page 287]

$$\text{(A.8)}\qquad E\sup_{v\in[0,t]}M_1^4(v)\le CE[M_1]_t^2.$$

Since the square integrable martingale M_1 is of integrable variation, it has no continuous part. Hence $[M_1]_t = \sum_{v \le t}(\Delta M_1(v))^2$. The process

$$\xi(t) = [M_1]_t - \langle M_1 \rangle_t = \int_0^t \frac{M^{(n)}(ds, x)}{\left(Y^{(n)}(s, x)\right)^4}$$

is a martingale, so $E[M_1]_1^2 \le 2E\langle M_1 \rangle_1^2 + 2E\xi^2(1) = 2E\langle M_1 \rangle_1^2 + 2E\langle \xi \rangle_1$. However, by (6.1)

$$E\langle M_1 \rangle_1^2 = E\left[\int_0^1 \frac{\lambda^{(n)}(s, x)}{\left(Y^{(n)}(s, x)\right)^4}\, ds\right]^2 = O\left(\frac{1}{nw_n}\right)^6,$$

$$E\langle \xi \rangle_1 = E\int_0^1 \frac{\lambda^{(n)}(s, x)}{\left(Y^{(n)}(s, x)\right)^8}\, ds = O\left(\frac{1}{nw_n}\right)^7.$$

It then follows from (A.7) and (A.8) that the second term in (A.6) is of order

$$O(nw_n)^3 O\left(\frac{1}{nw_n}\right)^{6\cdot(3/4)} \to 0,$$

so $E\tilde{H}^3(1, 1) < \infty$, and from (A.5),

$$E \sup_{t, z} \left|\hat{h}(t, z) - \tilde{h}(t, z)\right|^3 = O\left(w_n^3 b_n^{-9}\right). \tag{A.9}$$

Since $\sup_{t, z \ge b_n} |h(t, z) - h^0(t, z)| = O(b_n)$, by the Lipschitz condition on h,

$$\int_0^1 \int_0^1 \left|h(t, z) - h^0(t, z)\right|^3 dt\, dz = O(b_n). \tag{A.10}$$

For the third term in (A.4),

$$\begin{aligned}
&\sup_{t, z} |h^0(t, z) - h^\dagger(t, z)| \\
&\quad = \sup_{t, z} \left| \frac{1}{b_n^2} \int_0^1 K\left(\frac{t - s}{b_n}\right) \left[\int_0^1 K\left(\frac{z - x}{b_n}\right) h(s, x)\, dx \right.\right. \\
&\qquad\qquad \left.\left. - \frac{1}{d_n} \sum_{r=1}^{d_n} K\left(\frac{z - x_r}{b_n}\right) h(s, x_r)\right] ds \right| \\
&\quad \le \frac{1}{b_n} \sup_{s, z} \left[\sum_{r=1}^{d_n} \int_{I_r} \left| K\left(\frac{z - x}{b_n}\right) h(s, x) - K\left(\frac{z - x_r}{b_n}\right) h(s, x_r) \right| dx \right] \\
&\quad \le \frac{1}{b_n} \sup_{s, z} \left[\sum_{r=1}^{d_n} \int_{I_r} \left| K\left(\frac{z - x}{b_n}\right) - K\left(\frac{z - x_r}{b_n}\right) \right| h(s, x_r)\, dx \right] + O(w_n) \\
&\quad \le \frac{1}{b_n} O\left(w_n b_n^{-1}\right) + O(w_n) = O\left(w_n b_n^{-2}\right).
\end{aligned} \tag{A.11}$$

For the fourth term in (A.4),

$$
\begin{aligned}
&\sup_{t,z} E|h^{\dagger}(t,z) - h^{*}(t,z)|^3 \\
&\quad = \sup_{t,z} E\Bigg|\frac{1}{b_n^2 d_n}\sum_{r=1}^{d_n} K\bigg(\frac{z-x_r}{b_n}\bigg) \\
&\qquad\qquad \times \int_0^1 K\bigg(\frac{t-s}{b_n}\bigg)\bigg(h(s,x_r) - nw_n\frac{\lambda_r^{(n)}(s)}{\big(Y_r^{(n)}(s)\big)^2}\bigg)\,ds\Bigg|^3 \\
&\quad \le \bigg(\frac{1}{b_n^2 d_n}\bigg)^3 \sup_{t,z}\bigg[\sum_{r=1}^{d_n} K^{3/2}\bigg(\frac{z-x_r}{b_n}\bigg)\bigg]^{(2/3)\cdot 3} \\
&\qquad \times E\Bigg[\sum_{r=1}^{d_n}\bigg|\int_0^1 K\bigg(\frac{t-s}{b_n}\bigg)\bigg(h(s,x_r) - nw_n\frac{\lambda_r^{(n)}(s)}{\big(Y_r^{(n)}(s)\big)^2}\bigg)\,ds\bigg|^3\Bigg]^{(1/3)\cdot 3} \\
&\quad \le O\bigg(\frac{1}{b_n^2 d_n}\bigg)^3 (b_n d_n)^2 \sup_{t,z} E\Bigg[\sum_{r=1}^{d_n}\bigg|\int_0^1 K\bigg(\frac{t-s}{b_n}\bigg) \\
&\qquad\qquad \times \bigg(h(s,x_r) - nw_n\frac{\lambda_r^{(n)}(s)}{\big(Y_r^{(n)}(s)\big)^2}\bigg)\,ds\bigg|^3\Bigg] \\
&\quad \le O\bigg(\frac{1}{b_n^2 d_n}\bigg)^3 (b_n d_n)^2 \sup_{t,z}\sum_{r=1}^{d_n}\int_0^1 K^3\bigg(\frac{t-s}{b_n}\bigg) \\
&\qquad \times E\bigg|h(s,x_r) - nw_n\frac{\lambda_r^{(n)}(s)}{\big(Y_r^{(n)}(s)\big)^2}\bigg|^3\,ds \\
&\quad \le O\bigg(\frac{1}{b_n^2 d_n}\bigg)^3 (b_n d_n)^3 \sup_{s,r} E\bigg|h(s,x_r) - nw_n\frac{\lambda_r^{(n)}(s)}{\big(Y_r^{(n)}(s)\big)^2}\bigg|^3.
\end{aligned}
$$

Using the Lipschitz property of λ and (6.1), the supremum term above is bounded by

$$
\begin{aligned}
&O(w_n^3) + O(1)\sup_{s,r} E\bigg|\frac{Y_r^{(n)}(s)}{nw_n} - f(s,x_r)\bigg|^3 + O(1)\sup_{s,r} P\big(Y_r^{(n)}(s) = 0\big) \\
&\quad \le O(w_n^3) + O(1)\bigg[O\bigg(\frac{1}{nw_n}\bigg)^2\bigg]^{3/4} + O(\exp(-Cnw_n)) = O\bigg(\frac{1}{nw_n}\bigg)^{3/2},
\end{aligned}
$$

where we have used the Lipschitz property of f and the fact that $Y_r^{(n)}(s)$ is a binomial r.v. with mean of order $O(nw_n)$ to bound its fourth central moment. Therefore,

$$\text{(A.12)}\qquad \begin{aligned}\sup_{t,z} E|h^{\dagger}(t,z) - h^{*}(t,z)|^3 &= O\left(\frac{1}{b_n^3}\right) O\left(\frac{1}{nw_n}\right)^{3/2} \\ &= O\left(\frac{1}{nw_n b_n^2}\right)^{3/2}.\end{aligned}$$

Finally, for the fifth term in (A.4),

$$\text{(A.13)}\qquad \sup_{t,z} E|R(t,z)|^3 \le \sup_{t,z} \left[ER^4(t,z)\right]^{3/4}.$$

Let

$$R(t,z,u) = \frac{n}{b_n^2 d_n^2} \sum_{r=1}^{d_n} K\left(\frac{z - x_r}{b_n}\right) \int_0^u K\left(\frac{t-s}{b_n}\right) \frac{dM_r^{(n)}(s)}{\left(Y_r^{(n)}(s)\right)^2}.$$

Then $R(t, z, \cdot)$ is a martingale. Using the same arguments that were applied to M_1,

$$\text{(A.14)}\qquad \begin{aligned}ER^4(t,z) &\le E \sup_{0 \le u \le 1} R^4(t,z,u) \le CE[R(t,z,\cdot)]_1^2 \\ &\le O(1)\left(E\langle R(t,z,\cdot)\rangle_1^2 + E\xi^2(1)\right),\end{aligned}$$

where $\xi(u) = [R(t,z,\cdot)]_u - \langle R(t,z,\cdot)\rangle_u$ is a martingale. Since no two of the counting processes $N_r^{(n)}$, $r = 1, 2, \dots, d_n$ jump simultaneously,

$$\begin{aligned}[R(t,z,\cdot)]_u &= \sum_{v \le u} (\Delta R(t,z,v))^2 \\ &= \sum_{v \le u} \sum_{r=1}^{d_n} \left(\frac{n}{b_n^2 d_n^2}\right)^2 K^2\left(\frac{z - x_r}{b_n}\right) K^2\left(\frac{t - v}{b_n}\right) \frac{\Delta N_r^{(n)}(v)}{\left(Y_r^{(n)}(v)\right)^4} \\ &= \left(\frac{n}{b_n^2 d_n^2}\right)^2 \sum_{r=1}^{d_n} K^2\left(\frac{z - x_r}{b_n}\right) \int_0^u K^2\left(\frac{t-s}{b_n}\right) \frac{dN_r^{(n)}(s)}{\left(Y_r^{(n)}(s)\right)^4}.\end{aligned}$$

It follows that

$$\text{(A.15)}\qquad \xi(u) = \left(\frac{n}{b_n^2 d_n^2}\right)^2 \sum_{r=1}^{d_n} K^2\left(\frac{z - x_r}{b_n}\right) \int_0^u K^2\left(\frac{t-s}{b_n}\right) \frac{dM_r^{(n)}(s)}{\left(Y_r^{(n)}(s)\right)^4}.$$

The last two terms in (A.14) can be bounded above as follows:

$$\sup_{t,z} E\langle R(t,z,\cdot)\rangle_1^2 = \left(\frac{n}{b_n^2 d_n^2}\right)^4 \sup_{t,z} E\left[\sum_{r=1}^{d_n} K^2\left(\frac{z-x_r}{b_n}\right) \times \int_0^1 K^2\left(\frac{t-s}{b_n}\right) \frac{\lambda_r^{(n)}(s)}{\left(Y_r^{(n)}(s)\right)^4}\,ds\right]^2$$

$$\text{(A.16)} \qquad \le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 (nw_n)^{-6} \sup_{t,z} \sum_{r,l=1}^{d_n} K^2\left(\frac{z-x_r}{b_n}\right) K^2\left(\frac{z-x_l}{b_n}\right) \times \int_0^1\int_0^1 K^2\left(\frac{t-s}{b_n}\right) K^2\left(\frac{t-v}{b_n}\right) ds\,dv$$

$$\le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 (nw_n)^{-6} b_n^2 (b_n d_n)^2$$

$$= O(1)\left(n^2 b_n^4\right)^{-1},$$

$$\sup_{t,z} E\xi^2(1) = \sup_{t,z} E\langle \xi \rangle_1$$

$$= \left(\frac{n}{b_n^2 d_n^2}\right)^4 \sup_{t,z} \sum_{r=1}^{d_n} K^4\left(\frac{z-x_r}{b_n}\right) \int_0^1 K^4\left(\frac{t-s}{b_n}\right) E\frac{\lambda_r^{(n)}(s)}{\left(Y_r^{(n)}(s)\right)^8}\,ds$$

$$\text{(A.17)} \qquad \le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 \sup_{t,z} \sum_{r=1}^{d_n} K^4\left(\frac{z-x_r}{b_n}\right) \int_0^1 K^4\left(\frac{t-s}{b_n}\right) E\left(Y_r^{(n)}(s)\right)^{-7} ds$$

$$\le O\left(\frac{n}{b_n^2 d_n^2}\right)^4 (nw_n)^{-7} (b_n d_n) b_n$$

$$= O(1)\left(n^3 b_n^6\right)^{-1}.$$

From (A.13)–(A.17), we get

$$\text{(A.18)} \qquad \sup_{t,z} E|R(t,z)|^3 \le O(1)\left[n^2 b_n^4\right]^{-3/4}.$$

Finally combining the bounds for the five terms in (A.4), we have

$$\int_0^1\int_0^1 E|\hat{h}(t,z) - h(t,z)|^3\,dt\,dz$$

$$= O\left[w_n^3 b_n^{-9} + b_n + w_n^3 b_n^{-6} + \left(nw_n b_n^2\right)^{-3/2} + \left(nb_n^2\right)^{-3/2}\right]$$

$$= O(1)\left[b_n + w_n^3 b_n^{-9} + \left(nw_n b_n^2\right)^{-3/2}\right]. \qquad \square$$

LEMMA 5.

$$\int_0^1\int_0^1 P\left(\hat{h}(t,z) < c_n^{-1} \text{ or } \hat{h}(t,z) > c_n\right) dt\,dz$$
$$= O(1)\left[b_n + w_n^3 b_n^{-9} + \left(n w_n b_n^2\right)^{-3/2}\right].$$

PROOF. Let $0 < c \le C < \infty$ be lower and upper bounds for h,

$$\begin{aligned} P\left(\hat{h}(t,z) < c_n^{-1} \text{ or } \hat{h}(t,z) > c_n\right) &= P\left(\hat{h}(t,z) \le c_n^{-1}\right) + P\left(\hat{h}(t,z) \ge c_n\right) \\ &\le P\left(\hat{h}(t,z) - h(t,z) \le c_n^{-1} - c\right) + P\left(\hat{h}(t,z) - h(t,z) \ge c_n - C\right) \\ &= P\left(h(t,z) - \hat{h}(t,z) \ge c - c_n^{-1}\right) + P\left(\hat{h}(t,z) - h(t,z) \ge c_n - C\right) \\ &\le \frac{E|\hat{h}(t,z) - h(t,z)|^3}{\left(c - c_n^{-1}\right)^3} + \frac{E|\hat{h}(t,z) - h(t,z)|^3}{(c_n - C)^3} \end{aligned}$$

so the result follows by Lemma 4. □

LEMMA 6.

$$\sup_{t,z} E\left[\sqrt{n}\left(\frac{\lambda^{(n)}(t,z)}{Y^{(n)}(t,z)} - \lambda(t,z)\right)\right]^2 \to 0.$$

PROOF. The expression on the lhs above is bounded by

$$\begin{aligned} &\sup_{t,z} E\left[\sqrt{n}\,O(w_n) I\left(Y^{(n)}(t,z) \ne 0\right) + O(\sqrt{n}) I\left(Y^{(n)}(t,z) = 0\right)\right]^2 \\ &\qquad \le O\left(\sqrt{n w_n^2}\right) + O(n) \sup_{t,z} P\left(Y^{(n)}(t,z) = 0\right) \\ &\qquad \le O\left(n^{(1-2\alpha)/2}\right) + O\left(n \exp(-Cn^{1-\alpha})\right) \to 0, \end{aligned}$$

where the last inequality comes from (6.2). □

LEMMA 7. *For* $\delta(s,u,x)$ *defined by* (6.14),

$$\int_0^\rho\int_0^1\int_0^1 E|\delta(s,u,x)|^3\,ds\,du\,dx = O(1)c_n^{7.5}\left[b_n + w_n^3 b_n^{-9} + \left(n w_n b_n^2\right)^{-3/2}\right].$$

PROOF.

$$
\begin{aligned}
|\delta(s,u,x)| &= O(c_n)\Big|\big(\breve{h}^{-1}(s,u) - h^{-1}(s,u)\big)\breve{h}^{-1/2}(s,x)\int_x^1 h^{-1}(s,v)\,dv \\
&\qquad + \big(\breve{h}^{-1/2}(s,x) - h^{-1/2}(s,x)\big)h^{-1}(s,u)\int_x^1 h^{-1}(s,v)\,dv \\
&\qquad + h^{-1/2}(s,x)h^{-1}(s,u)\int_x^1\big(h^{-1}(s,v) - \breve{h}^{-1}(s,v)\big)\,dv\Big| \\
&= O(c_n)\Big[c_n^{1/2}\big|\breve{h}^{-1}(s,u) - h^{-1}(s,u)\big| \\
&\qquad + \big|\breve{h}^{-1/2}(s,x) - h^{-1/2}(s,x)\big| \\
&\qquad + \int_x^1 |h^{-1}(s,v) - \breve{h}^{-1}(s,v)|\,dv\Big] \\
&= O(c_n)\Big[c_n^{3/2}\big|\hat{h}(s,u) - h(s,u)\big| \\
&\qquad + c_n^{3/2} I\big(\hat{h}(s,u) < c_n^{-1} \text{ or } \hat{h}(s,u) > c_n\big) \\
&\qquad + c_n\big|\hat{h}(s,x) - h(s,x)\big| \\
&\qquad + c_n^{1/2} I\big(\hat{h}(s,x) < c_n^{-1} \text{ or } \hat{h}(s,x) > c_n\big) \\
&\qquad + c_n\int_x^1 \big|\hat{h}(s,v) - h(s,v)\big|\,dv \\
&\qquad + c_n\int_x^1 I\big(\hat{h}(s,v) < c_n^{-1} \text{ or } \hat{h}(s,v) > c_n\big)\,dv\Big],
\end{aligned}
$$

where we have used the fact that $\breve{h} = \hat{h}$ when $c_n^{-1} \le \hat{h} \le c_n$. Thus,

$$
\begin{aligned}
E|\delta(s,u,x)|^3 \le O(c_n^3)\Big[&c_n^{9/2} E\big|\hat{h}(s,u) - h(s,u)\big|^3 \\
&+ c_n^{9/2} P\big(\hat{h}(s,u) < c_n^{-1} \text{ or } \hat{h}(s,u) > c_n\big) \\
&+ c_n^3 E\big|\hat{h}(s,x) - h(s,x)\big|^3 \\
&+ c_n^3\int_0^1 E\big|\hat{h}(s,v) - h(s,v)\big|^3\,dv \\
&+ c_n^3\int_0^1 P\big(\hat{h}(s,v) < c_n^{-1} \text{ or } \hat{h}(s,v) > c_n\big)\,dv\Big].
\end{aligned}
$$

Now apply Lemmas 4 and 5 to complete the proof. □

Acknowledgments. The authors thank Estate Khmaladze for his encouragement, Jayaram Sethuraman for help in refining the proof of Theorem 2.1 and I. C. M. MacLennan for furnishing data from the British Medical Research Council's 4th myelomatosis trial. They also thank the Associate Editor and a referee for many helpful comments.

REFERENCES

AALEN, O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.* **6** 701–726.

ADLER, R. J. (1991). *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*. IMS, Hayward, CA.

ANDERSEN, P. K., BORGAN, O., GILL, R. D. and KEIDING, N. (1982). Linear non-parametric tests for comparison of counting processes, with applications to censored survival data (with discussion). *Internat. Statist. Rev.* **50** 219–258 [Correction (1984) **52** 225].

ANDERSEN, P. K., BORGAN, O., GILL, R. and KEIDING, N. (1993). *Statistical Models Based on Counting Processes*. Springer, Berlin.

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

COX, D. R. (1972). Regression models and life tables (with discussion). *J. Roy. Statist. Soc. Ser. B* **34** 187–220.

CUZICK, J., COOPER, E. H. and MACLENNAN, I. C. M. (1985). The prognostic value of serum β_2 microglobulin compared with other presentation features in myelomatosis. *British Journal of Cancer* **52** 1–6.

DELLACHERIE, C. and MEYER, P. A. (1982). *Probabilities and Potential B*. North-Holland, Amsterdam.

DOOB, J. (1949). Heuristic approach to the Kolmogorov–Smirnov theorems. *Ann. Math. Statist.* **20** 393–403.

DURBIN, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* **1** 279–290.

HILDEBRANDT, T. H. (1963). *Introduction to the Theory of Integration*. Academic Press, New York.

HJORT, N. L. (1990). Goodness of fit tests in models for life history data based on cumulative hazard rates. *Ann. Statist.* **18** 1221–1258.

KEIDING, N. and GILL, D. (1990). Random truncation models and Markov processes. *Ann. Statist.* **18** 582–602.

KHMALADZE, E. V. (1981). Martingale approach to the theory of goodness-of-fit tests. *Theory Probab. Appl.* **26** 246–265.

KHMALADZE, E. V. (1993). Goodness of fit problem and scanning innovation martingales. *Ann. Statist.* **21** 798–829.

LIN, D. Y., WEI, L. J. and YING, Z. (1993). Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika* **80** 557–572.

LIPSTER, R. S. and SHIRYAYEV, A. N. (1977). *Statistics of Random Processes I*. Springer, Berlin.

MCKEAGUE, I. W. and UTIKAL, K. J. (1990). Identifying nonlinear covariate effects in semimartingale regression models. *Probab. Theory Related Fields* **87** 1–25.

NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** 1285–1295.

NIKABADZE, A. M. (1987). On a method for constructing goodness-of-fit tests for parametric hypotheses in $\mathbb{R}^m$. *Theory Probab. Appl.* **32** 539–544.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

WONG, E. and ZAKAI, M. (1974). Martingales and stochastic integrals for processes with a multi-dimensional parameter. *Z. Wahrsch. Verw. Gebiete* **29** 109–122.

IAN W. MCKEAGUE
DEPARTMENT OF STATISTICS
FLORIDA STATE UNIVERSITY
TALLAHASSEE, FLORIDA 32306-3033

A. M. NIKABADZE
A. RAZMADZE MATHEMATICAL INSTITUTE
RUKHADZE ST. 1
380093 TBILISI
REPUBLIC OF GEORGIA

YANQING SUN
DEPARTMENT OF BIOSTATISTICS
UNIVERSITY OF ROCHESTER
ROCHESTER, NEW YORK 14642