Econometric Theory, **22**, 2006, 614–632. Printed in the United States of America.
DOI: 10.1017/S0266466606060300

A CONSISTENT NONPARAMETRIC EQUALITY TEST OF CONDITIONAL QUANTILE FUNCTIONS

YIGUO SUN
University of Guelph

This paper proposes a consistent nonparametric test for testing for equality of unknown conditional quantile curves across subgroups within a survey data framework. Moreover, to improve the small-sample performance of the test, we propose a modified version of wild bootstrap procedure in a quantile context. Monte Carlo evidence shows that the performance of the test in small samples is adequate but is improved significantly when the bootstrap critical values are used.

1. INTRODUCTION

Survey data are frequently used in economics and have one prominent characteristic: data heterogeneity across population subgroups. The presence of heterogeneity naturally makes attractive the application of conditional quantile regressions; see Koenker and Bassett (1978, 1982), Buchinsky (1994), Powell (1986), and Newey and Powell (1990), among others. A major concern of modeling survey data is whether the data are poolable. To our knowledge, testing for poolability within the framework of a conditional quantile regression has not yet been studied in the literature, although within the framework of a conditional mean regression it has been investigated by Baltagi, Hidalgo, and Li (1996) and Lavergne (2001), among others. This paper aims to fill in the literature gap by providing a consistent nonparametric test for testing for the equality of unknown conditional quantile curves across population subgroups. The proposed test is a residual-based test derived by recognizing that an orthogonal moment condition only holds under the null hypothesis. Tests based on orthogonal moment(s) have been investigated extensively in other contexts also; related work includes Fan and Li (1996), Zheng (1996, 1998), and many others.

The proposed nonparametric test requires the availability of a consistent nonparametric estimator of the unknown conditional quantile curve. In the literature of nonparametric conditional quantile curve estimation, related work includes the kernel and k-nearest neighbor estimator of Bhattacharya and

I thank the co-editor and two referees for helpful comments that improved the paper. Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged. Address correspondence to Yiguo Sun, Department of Economics, University of Guelph, Guelph, Ontario, Canada N1G 2W1; e-mail: yisun@uoguelph.ca.

 © 2006 Cambridge University Press 0266-4666/06 $12.00

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Gangopadtyay (1990), the spline smoothing estimator of Koenker, Ng, and Portnoy (1994), the weighted Nadaraya–Watson estimator of Hall, Wolff, and Yao (1999), the local linear regression approach of Fan, Hu, and Truong (1994), the double kernel method of Yu and Jones (1998), and many others. Because the proposed test is designed to handle data containing both discrete (with or without natural ordering) and continuous regressors—usually the case of survey data—we extend the estimation method of Fan et al. (1994) to the mixed data case.

We show that the proposed nonparametric test not only is consistent against any fixed alternative but it also has asymptotic power of one against local alternatives converging to the null hypothesis at usual nonparametric rates. Monte Carlo evidence, however, indicates a size-distortion problem as well documented in the literature of nonparametric hypothesis tests in finite samples. In a mean regression context, the technique of wild bootstrap is used to improve the finite-sample performance of a nonparametric test; see Gozalo (1997), Härdle and Mammen (1993), and Li and Wang (1998), for example. The commonly used wild bootstrap, however, fails in a quantile context because the conditional mean of disturbance terms may not be zero for a general quantile regression model. Hence, we propose a modified version of the wild bootstrap procedure in the quantile context. Monte Carlo simulation results indicate that the proposed bootstrap procedure significantly improves the performance of the test in small samples.

The rest of the paper is organized as follows. Section 2 gives the test and its asymptotic results. Section 3 describes the bootstrap procedure and provides theoretical justification. Section 4 presents a Monte Carlo simulation study of the test. Section 5 concludes. All relevant proofs are given in the Appendixes.

2. EQUALITY TEST AND MAIN THEOREMS

Suppose we fit the following nonparametric quantile regression model (at a probability mass $\theta \in (0,1)$) to a set of survey data comprising J subgroups with total sample size n:

$$y_{ij} = g_{\theta,j}(x_{ij}) + v_{ij}, \tag{1}$$

where the θth-conditional quantile of v_{ij} given x_{ij} is 0 and $g_{\theta,j}(x)$ is an unknown function of x, $i = 1,2,\ldots,n_j; j = 1,2,\ldots,J$. Here J (≥ 2) is a finite number. Denote the sample size of the jth subgroup by n_j. We have $a_j = \lim_{n\to\infty}(n_j/n)$, where $a_j \in (0,1)$ and $\sum_{j=1}^{J} a_j = 1$. Let $x_{ij} = (x'_{1,ij}, x'_{2,ij})' \in R^{q-p} \times R^p$, where x_{ij} is a $q \times 1$ column vector of explanatory variables with p (≥ 1) continuously distributed variables $x_{2,ij}$ and $q - p (\geq 0)$ discretely distributed variables $x_{1,ij}$.

Let $F_j(y|x)$ be the conditional distribution of Y given $X = x$ of the jth subgroup. Then

$$g_{\theta,j}(x) \equiv \inf\{y : F_j(y|x) \geq \theta\}. \tag{2}$$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Throughout this paper we assume that $F_j(y|x)$ are absolutely continuous in y for almost all x, which implies that $F_j(g_{\theta,j}(x)|x) = \theta$ for almost all x. To simplify notation, we suppress the subscript θ of $g_{\theta,j}(\cdot)$. The aim of this paper is to test whether $g_j(\cdot)$ are the same across subgroups. If they are, the pooled data will be used to estimate the unknown conditional quantile curve to gain more efficiency. Hence, the null and alternative hypotheses under consideration are

$$H_0 : \Pr[g_j(x) = g_l(x)] = 1 \quad \text{for all } j \text{ and } l,$$

$$H_1 : \Pr[g_j(x) = g_l(x)] < 1 \quad \text{for some } j \neq l.$$

Using the pooled data to estimate the unknown θth-conditional quantile curve, $g(x)$, Lemma 1, which follows, shows that $g(x)$ is a weighted average of $g_j(x)$ and that $g(x) \equiv g_j(x)$ holds for all j only under H_0. Hence, the equivalent null and alternative hypotheses can be written as

$$H_0 : \Pr[g(x) = g_j(x)] = 1 \quad \text{for all } j\,,$$

$$H_1 : \Pr[g(x) = g_j(x)] < 1 \quad \text{for some } j.$$

As in Zheng (1998), we transfer model (1) to the following conditional mean regression model:

$$I(y_{ij} \leq g_j(x_{ij})) = \theta + \varepsilon_{ij}, \qquad E(\varepsilon_{ij}|x_{ij}) = 0, \qquad E(\varepsilon_{ij}^2|x_{ij}) = \theta(1-\theta). \tag{3}$$

Let $u_{ij} = I(y_{ij} \leq g(x_{ij})) - \theta$. Then under H_0, $u_{ij} = \varepsilon_{ij}$; under H_1, $u_{ij} = \varepsilon_{ij} + I(y_{ij} \leq g(x_{ij})) - I(y_{ij} \leq g_j(x_{ij}))$. Hence, $E[u_{ij}E(u_{ij}|x_{ij})] \geq 0$, and the equality holds if and only if H_0 is true. To avoid the random denominator problem in nonparametric kernel estimation, we use the density weighted version of $E[u_{ij}E(u_{ij}|x_{ij})]$ as in Powell, Stock, and Stoker (1989). The leave-one-out estimator of $E[u_{ij}E(u_{ij}|x_{ij})f_j(x_{ij})]$ is

$$\begin{aligned} J_n &= \frac{1}{n\sqrt{|H|}} \sum_{j=1}^{J} \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \sum_{k \neq i} \hat{u}_{ij}\hat{u}_{kj} K_{ij,kj} \\ &= \frac{1}{n\sqrt{|H|}} \sum_{j=1}^{J} \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \sum_{k \neq i} (u_{ij}u_{kj} + 2u_{ij}\hat{\omega}_{kj} + \hat{\omega}_{ij}\hat{\omega}_{kj}) K_{ij,kj} \\ &= J_{n1} + 2J_{n2} + J_{n3}, \end{aligned} \tag{4}$$

where $\hat{u}_{ij} = I(y_{ij} \leq \hat{g}(x_{ij})) - \theta = u_{ij} + \hat{\omega}_{ij}$ and $\hat{\omega}_{ij} = I(y_{ij} \leq \hat{g}(x_{ij})) - I(y_{ij} \leq g(x_{ij}))$. Also, for any fixed j,

$$K_{ij,kj} = K_1(x_{1,ij} - x_{1,kj}; \lambda) K_2(H^{-1/2}(x_{2,ij} - x_{2,kj})), \tag{5}$$

where $\lambda = (\lambda_1, \ldots, \lambda_{q-p})'$ are the smoothing parameters for the discrete variables and $H = \operatorname{diag}(h_1^2, \ldots, h_p^2)$ for the continuous variables, and $|H| = \prod_{s=1}^{p} h_s^2$. Moreover, $\hat{g}(x_{ij})$ defined in Lemma 1 is the nonparametric estimator

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

of $g(x_{ij})$ from the pooled data by using a different set of smoothing parameters $(\tilde{\lambda}, \tilde{H})$ in (7), where $\tilde{\lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_{q-p})'$ and $\tilde{H} = \text{diag}(\tilde{h}_1^2, \ldots, \tilde{h}_p^2)$.

Let $\|A\| = \sqrt{\text{tr}(AA')}$ for any matrix A and $x = (x_1', x_2')' \in R^{q-p} \times R^p$. We list all the relevant assumptions as follows.

Assumption 1. (x_{ij}, y_{ij}) are independent and identically distributed (i.i.d.) in the subscript i and independently distributed in the subscript j.

Assumption 2. Let $f_{1,j}(x_1)$ be the probability of $x_{1,ij} = x_1$ and $f_{2|1,j}(x_2|x_1)$ be the conditional density function of $x_{2,ij}$ given $x_{1,ij} = x_1$. Denote $f_j(x_1, x_2) = f_{1,j}(x_1) f_{2|1,j}(x_2|x_1)$. Then they satisfy the following restrictions:

(i) $x_{1,ij}$ can take any value from $\mathcal{X}_1$, which contains $\prod_{t=1}^{q-p} c_t$ different points $\{0, 1, \ldots, c_t - 1 : c_t \geq 2\}_{t=1}^{q-p}$. Also, $\sum_{x_1 \in \mathcal{X}_1} f_{1,j}(x_1) = 1$ with $f_{1,j}(x_1) > 0$ for any $x_1 \in \mathcal{X}_1$.

(ii) $f_{2|1,j}(x_2|x_1) > 0$, and its first- and second-order (partial) derivatives with respect to x_2 are all uniformly bounded.

Assumption 3. The conditional cumulative distribution function (c.d.f.) and probability density function (p.d.f.) of errors v_{ij} given $x_{ij} = x$ are $F_{v,j}(v|x)$ and $f_{v,j}(v|x)$, respectively. Also, $f_{v,j}(v|x)$ has uniformly bounded first- and second-order partial derivatives with respect to v and x_2. Moreover, $f_{v,j}(0|x) > 0$ and $F_{v,j}(0|x) = \Pr(v_{ij} \leq 0 | x_{ij} = x) = \theta$ hold for all j on its domain.

Assumption 4. The unknown quantile curves $g_j(x)$ satisfy the conditions in Fan et al. (1994). For instance, $g_j(x)$ are twice continuously differentiable with respect to x_2, and $g_j(x)$ and their first- and second-order partial derivatives with respect to x_2 are all uniformly bounded.

Assumption 5. The product kernel function is used, $K(u_1, u_2) = K_1(u_1; \lambda) \times K_2(u_2)$, where $K_1(u_1; \lambda) = \prod_{s=1}^{q-p} k_1(u_{1,s}; \lambda_s)$ and $K_2(u_2) = \prod_{s=1}^{p} k_2(u_{2,s})$, $u = (u_1', u_2')' \in R^{q-p} \times R^p$. (i) $k_1(u_{1,s}; \lambda_s) = 1$ if $u_{1,s} = 0, \lambda_s$ otherwise, where $0 \leq \lambda_s < 1$ for all s; (ii) $k_2(\cdot)$ is a nonnegative, symmetric, and bounded function taking values on $[-1, 1]$ such that $\int_{-1}^{1} k_2(u_{2,s})\, du_{2,s} = 1$ $(s = 1, \ldots, p)$, $\mu_2(K_2) = \int_{-1}^{1} \cdots \int_{-1}^{1} K_2(u_2) u_2 u_2'\, du_{2,1} \cdots du_{2,p}$, and $R_2(K_2) = \int_{-1}^{1} \cdots \int_{-1}^{1} K_2^2(u_2)\, du_{2,1} \cdots du_{2,p} < \infty$.

Assumption 6. As $n \to \infty$, the smoothing parameters satisfy the following restrictions: (i) $n|H|^{1/2} \to \infty$, $\|H\| \to 0$, $\lambda_s \to 0$ for all s; (ii) $n|\tilde{H}|^{1/2} \to \infty$, $\|\tilde{H}\| \to 0$, $\tilde{\lambda}_{\max} = \max_{1 \leq s \leq q-p} \tilde{\lambda}_s = O(\|\tilde{H}\|^2)$; (iii) $n|H|^{1/4} \|\tilde{H}\|^4 \to 0$; and (iv) $|H||\tilde{H}|^{-2} \to 0$.

The term $K_1(u_1; \lambda)$ in Assumption 5 is similar to what is defined by Racine and Li (2004). One advantage of this smoothing method relative to the conventional estimator with $\lambda_s = 0$ for all s lies in its ability to borrow information from nearby cells.

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

The construction of J_n in (4) requires consistent nonparametric estimation of $g(x)$. For the continuous case Fan et al. (1994) develop a local linear regression approach to estimate $g(x)$ and give the consistency and the asymptotic normality results of the estimator. Our paper extends their estimation method to the multivariate mixed categorical and continuous cases, and the result is presented in the following lemma.

LEMMA 1. *Suppose Assumptions 1–5 and (ii) of Assumption 6 hold. Then we have the following condition.*

(i) For any fixed $x_0 = (x'_{1,0}, x'_{2,0})' \in \mathcal{X}_1 \times R^p$, *there exists a set of weights* $\{\lambda_j(x_0)\}_{j=1}^J$ *satisfying* $\lambda_j(x_0) \in [0,1)$ *and* $\sum_{j=1}^J \lambda_j(x_0) = 1$ *so that* $g(x_0) = \sum_{j=1}^J \lambda_j(x_0) g_j(x_0)$.

(ii) Solving the following optimization problem:

$$(\hat{a}, \hat{b}) = \arg\min_{a,b} \hat{Q}_n(a,b)$$

$$= \sum_{j=1}^{J} \sum_{i=1}^{n_j} \rho_\theta(y_{ij} - a - b'(x_{2,ij} - x_{2,0}))\tilde{K}_{00,ij} \tag{6}$$

gives $\hat{g}(x_0) = \hat{a}$ *and* $\partial \hat{g}(x_0)/\partial x_2 = \hat{b}$, *where* $\rho_\theta(u) = u[\theta - I(u \le 0)]$ *is called the "check function" and*

$$\tilde{K}_{00,ij} = K_1(x_{1,0} - x_{1,ij}; \tilde{\lambda}) K_2(\tilde{H}^{-1/2}(x_{2,0} - x_{2,ij})); \tag{7}$$

and (iii)

$$\hat{g}(x_0) - g(x_0) = O_p(\|\tilde{H}\|^2) + O_p(n^{-1/2}|\tilde{H}|^{-1/4}). \tag{8}$$

The discrete variables are excluded from the check function in (6) because the validity of Lemma 1 requires the error term of the Taylor expansion (A.10) in Appendix A to vanish at some appropriate rate.

The following theorem gives the main asymptotic results of the proposed test and Lemmas 6–9 in Appendix A give the proof of the theorem.

THEOREM 2. *If Assumptions 1–6 hold, under* H_0 *we have*

$$n|H|^{1/4} J_n \xrightarrow{d} N(0, \sigma_0^2) \quad \text{as } n \to \infty, \tag{9}$$

where $\sigma_0^2 = 2\theta^2(1-\theta)^2 R_2(K_2) \sum_{j=1}^J E[f_j(x_{1,ij}, x_{2,ij})]$ *can be consistently estimated by*

$$\hat{\sigma}_0^2 = 2\theta^2(1-\theta)^2 \frac{1}{\sqrt{|H|}} \sum_{j=1}^{J} \frac{1}{n_j(n_j-1)} \sum_{i=1}^{n_j} \sum_{k \ne i} K_{ij,kj}^2. \tag{10}$$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Under H_1 we have

$$J_n \xrightarrow{p} \delta \quad as\ n \to \infty, \tag{11}$$

where $\delta = \sum_{j=1}^{J} a_j E\{[F_{v,j}(g(x_{ij}) - g_j(x_{ij})|x_{ij}) - \theta]^2 f_j(x_{1,ij}, x_{2,ij})\} > 0$. *Thus*

$$\Pr[n|H|^{1/4} J_n/\hat{\sigma}_0 > B_n] \to \infty \quad as\ n \to \infty, \tag{12}$$

where $B_n = o(n|H|^{1/4})$.

To investigate the local power of the test, we consider the following local alternative hypotheses:

$$H_1^a : g_j(x) = g_1(x) + d_n m_j(x) \quad \text{for } j = 1, \ldots, J, \tag{13}$$

where $d_n \to 0$ as $n \to \infty$, the unknown functions $m_j(x)$ are assumed to have continuous second-order partial derivatives with respect to x_2, and $m_j(x)$ and their first- and second-order partial derivatives are all uniformly bounded.

THEOREM 3. (Local power). *If Assumptions 1–6 hold, under local alternatives H_1^a with $d_n = n^{-1/2}|H|^{-1/8}$, we have*

$$n|H|^{1/4} J_n \xrightarrow{d} N(\mu, \sigma_0^2) \quad as\ n \to \infty, \tag{14}$$

where $\mu = \sum_{j=1}^{J} a_j E[f_{v,j}^2(0|x_{ij}) \delta_j^2(x_{ij}) f_j(x_{1,ij}, x_{2,ij})] > 0$, $\delta_j(x) = \sum_{l=1}^{J} \lambda_l(x)$ $[m_l(x) - m_j(x)]$, *and* $\lambda_l(x)$ *are as in Lemma 1.*

Theorems 2 and 3 indicate that the proposed test is consistent against any fixed alternative and against local alternatives converging to the null hypothesis at usual nonparametric rates. Let $T_n = n|H|^{1/4} J_n/\hat{\sigma}_0$; then $T_n \xrightarrow{d} N(0,1)$ under H_0. This is a one-sided test. If T_n is greater than the 95% critical value of the standard normal distribution, then the null hypothesis will be rejected at the significance level of 5%.

3. HOW TO BOOTSTRAP

Bootstrap procedure is usually used to improve the performance of a test in small samples. Härdle and Mammen (1991, 1993) show that the two-point wild bootstrap is valid in the context of nonparametric estimation of conditional mean curves and nonparametric model specification tests. However, the commonly used wild bootstrap fails in a quantile context where the assumption of disturbances having zero conditional mean is not true. Therefore, this paper proposes a modified version of the wild bootstrap procedure for the quantile case.

In our context, we define the estimated residuals as

$$\hat{v}_{ij} = y_{ij} - \hat{g}(x_{ij}), \qquad i = 1,2,\ldots,n_j; \qquad j = 1,2,\ldots,J. \tag{15}$$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Let v_{ij}^b be the bootstrap resamples of $\hat{v}_{ij}$. For each pair of index (i,j), randomly draw with replacement the bootstrap residuals v_{ij}^b from a two-point distribution $\hat{F}_{ij}$ such that

$$E_{\hat{F}_{ij}}[I(Z \le 0)] = \theta, \tag{16}$$

$$E_{\hat{F}_{ij}} Z^2 = \hat{v}_{ij}^2, \tag{17}$$

$$E_{\hat{F}_{ij}} Z^3 = \hat{v}_{ij}^3, \tag{18}$$

where $\hat{F}_{ij} = p_{ij}\delta_{a_{ij}} + (1-p_{ij})\delta_{b_{ij}}$ and δ_x denotes a probability measure that puts mass one at x, $p_{ij} \in (0,1)$. If $a_{ij} \le 0 < b_{ij}$, (16) gives $p_{ij} = \theta$. Given $p_{ij} = \theta$, solving (17) and (18) gives $a_{ij} = c_{ij}\hat{v}_{ij}$ and $b_{ij} = d_{ij}\hat{v}_{ij}$ where $c_{ij} = -d_{ij}/(1+d_{ij})$ and d_{ij} is one of the real roots of the following function:

$$(1-\theta)x^4 + 2(1-\theta)x^3 - 2x - 1 = 0. \tag{19}$$

For example, when $\theta = 0.25$, we obtain[1]

$$a_{ij} = \hat{v}_{ij}[1.67967I(\hat{v}_{ij} \le 0) - 0.5269412I(\hat{v}_{ij} > 0)] \le 0,$$

$$b_{ij} = \hat{v}_{ij}[-0.6268197I(\hat{v}_{ij} < 0) + 1.113902I(\hat{v}_{ij} > 0)] > 0.$$

We now construct the bootstrap resamples $\{(x_{ij}, y_{ij}^b)\}$, where

$$y_{ij}^b = \hat{g}^*(x_{ij}) + v_{ij}^b \tag{20}$$

and $\hat{g}^*(x_{ij})$ are calculated from the pooled data $\{(x_{ij}, y_{ij})\}$ as described in Lemma 1 with smoothing parameters $\tilde{\lambda}^*$ and $\tilde{H}^*$. The proposed bootstrap method guarantees that $\Pr^b(y_{ij}^b \le \hat{g}^*(x_{ij})) = \theta$ for all i and j, where the probability is taken over the distribution of y_{ij}^b conditional on the true observations $\{(x_{ij}, y_{ij})\}$. Lemma 4, which follows, shows that $\hat{g}^*(x_{ij})$ needs to be oversmoothed, as found by Härdle and Mammen (1991) in the conditional mean regression context.

LEMMA 4. *Given the assumptions in Lemma 1, we have*

$$\hat{g}^b(x_0) - \hat{g}^*(x_0) = O_p(\|\tilde{H}\|^2) + O_p(n^{-1/2}|\tilde{H}|^{-1/4}). \tag{21}$$

Moreover, if $\|\tilde{H}\| = o(\|\tilde{H}^*\|)$, $\tilde{\lambda}_{\max} = O(\|\tilde{H}\|^2\|\tilde{H}^*\|^2)$, *and* $\tilde{\lambda}^*_{\max} = O(\|\tilde{H}^*\|^2)$, *we have*

$$E\{[\hat{g}^b(x_0) - \hat{g}^*(x_0)] - [\hat{g}(x_0) - g(x_0)]|x\}$$
$$= O_p(\|\tilde{H}\|^2\|\tilde{H}^*\|^2) + O_p(n^{-1/2}|\tilde{H}^*|^{-1/4}\|\tilde{H}^*\|^{-2}\|\tilde{H}\|^2). \tag{22}$$

Given that the bootstrap resamples $\{(x_{ij}, y_{ij}^b)\}$, we construct the bootstrap test T_n^*. We conduct B bootstrap repetitions to obtain the bootstrap test statistics $\{T_{n,b}^*\}_{b=1}^B$ so that we can construct the $(1-\alpha)$ quantile t_α^*. If T_n calculated from the true data is greater than t_α^*, we reject the null hypothesis of poolability. The validity of the proposed bootstrap procedure requires that the distribution of T_n^*

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

should consistently approximate that of T_n derived under the null hypothesis no matter whether the data are or are not generated from the null hypothesis. In what follows, we give the theoretical results and delay the proof to Appendix B.

THEOREM 5. *If the assumptions in Lemma 4 hold, then*

$$d_2(\mathfrak{L}^*(T_n^*), N(0,1)) \xrightarrow{p} 0, \tag{23}$$

where $\mathfrak{L}^*(T_n^*)$ *is the conditional distribution of the bootstrap test* T_n^* *given* $\{(x_{ij}, y_{ij})\}$ *and* $d_2(G,H)$*, the Mallows distance, is the infimum of* $E(X-Y)^2$ *over all joint distributions for the pair of random variables X and Y whose marginal distribution functions are G and H, respectively. The functions G and H belong to a functional space* $\Gamma_2 = \{G : \int x^2\, dG(x) < \infty\}$.

Bickel and Freedman (1981) show that convergence in Mallows distance is equivalent to the weak convergence; Pólya's theorem will extend the weak convergence to the strong convergence (Serfling 2002, p. 18).

4. MONTE CARLO SIMULATIONS

In this section we investigate the finite-sample properties of the test by means of Monte Carlo simulations. We consider $\theta = 0.25$, 0.5, and 0.75; $n = 200$ and 400. The number of Monte Carlo replications is $M = 1{,}000$ for each experiment. The data generating processes take the form of

$$y_{ij} = 1 + x_{1,ij} + x_{2,ij} + x_{1,ij}x_{2,ij} + x_{2,ij}^2 + d(1 + x_{2,ij})^2 + \sigma_j(x_{ij})u_{ij}, \tag{24}$$

where $u_{ij} \sim i.i.d.N(\xi_\theta, 1)$ such that the θth quantile of u_{ij} is zero, $x_{1,ij} \in \{0,1,2,3\}$ with natural ordering, and $\Pr(x_{1,ij} = l) = \frac{1}{4}$ for $l = 0,1,2,3$, for $i = 1,2,\ldots,n_j$, $j = 1,2$, and $n_1 = n_2$. When $j = 1$, we generate $x_{2,i1} \sim i.i.d.\ U[0,1]$, $\sigma_1(x) \equiv 0.5$, and $d = 0$. When $j = 2$, we generate $x_{2,i2} \sim i.i.d.\ \sqrt{5 - (U[0,1]-2)^2 - 1}$, $\sigma_2(x_{i2}) = cx_{2,i2}$ with c determined by $n_2^{-1}\sum_{i=1}^{n_2}\sigma_2(x_{i2}) = 0.5$, and $d \in \{0, 0.5, 1\}$ where $d = 0$ gives the null hypothesis.

Two pairs of smoothing parameters are used: (λ_j, h_j) for nonparametric estimation from subgroup data and $(\tilde{\lambda}, \tilde{h})$ for nonparametric estimation from the pooled data. We set $h_j = \hat{\sigma}_{x_{2,j}} n_j^{-\alpha}$, $\lambda_j = h_j^2$, $\tilde{h} = \hat{\sigma}_{x_2} n^{-\beta}$, and $\tilde{\lambda} = \tilde{h}^2$, where $\hat{\sigma}_{x_{2,j}}$ and $\hat{\sigma}_{x_2}$ are the standard deviations of $x_{2,ij}$ for a given j and for the pooled data, respectively. Assumption 6 implies that $\alpha > \frac{2}{5}$ and $\beta \in (\frac{1}{4} - (\alpha/8), (\alpha/2))$; we take $\beta = 0.2$ and take $\alpha = 0.6$ and 0.65 to measure the sensitivity of the test to the choice of bandwidth. Following Racine and Li (2004), we define $k_1(u;\lambda) = \lambda^s$ if $|u| = s$. Here $k_2(u) = \frac{3}{4}(1-u^2)I(|u| \le 1)$ is the Epanechnikov kernel. We acknowledge the use of the Koenker and d'Orey (1987) algorithm.

Table 1 presents the percentage rejection rates of the test. Generally speaking, the power of the test improves as the sample size increases; the size of the test, however, does not improve much when the sample size increases to 400

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

TABLE 1. Percentage of rejections using critical values from the standard normal distribution

	θ		0.25			0.5			0.75		
α	d	n	1%	5%	10%	1%	5%	10%	1%	5%	10%
0.6	0.0	200	0.0	0.0	0.1	0.0	0.1	0.1	0.0	0.2	0.3
		400	0.0	0.1	0.4	0.0	0.0	0.3	0.0	0.5	1.0
	0.5	200	11.7	24.5	35.4	37.6	62.6	74.6	5.7	15.3	22.8
		400	74.1	89.8	95.6	99.8	99.9	100	52.1	78.4	87.1
	1.0	200	26.2	50.4	64.7	85.2	94.3	96.4	16.1	33.6	48.2
		400	95.0	99.1	99.8	100	100	100	83.1	94.6	97.3
0.65	0.0	200	0.0	0.1	0.2	0.1	0.1	0.3	0.0	0.2	0.5
		400	0.0	0.2	0.6	0.0	0.1	0.5	0.1	0.5	1.1
	0.5	200	9.8	20.7	30.5	29.4	53.8	66.7	5.0	13.5	20.8
		400	60.0	82.4	90.1	98.4	99.8	99.9	41.2	67.6	79.7
	1.0	200	21.6	41.3	57.2	75.0	90.7	94.8	12.8	29.6	40.3
		400	87.5	96.4	98.5	100	100	100	69.7	88.1	93.5

from 200, and this size distortion is observed in all cases; a smaller bandwidth ($\alpha = 0.65$) is good for size and a larger bandwidth ($\alpha = 0.6$) is good for power; the optimal bandwidth will be different for the test at the different probability masses.

To improve the performance of the test in small samples, we apply the proposed bootstrap procedure to the case of $\theta = 0.25$ and $n = 200$ and present the results in Table 2. By Lemma 4, we use $\tilde{h}^* = \hat{\sigma}_{x_2} n^{-1/9}$ and $\tilde{\lambda}^* = \tilde{h}^{*2}$ to calculate $\hat{g}^*(\cdot)$, and we replace $\tilde{\lambda}$ defined earlier by $\tilde{h}^2\tilde{h}^{*2}$. With 100 bootstrap

TABLE 2. Percentage of rejections using bootstrap critical values: $\theta = 0.25$ and $n = 200$

		$\alpha = 0.6$			$\alpha = 0.65$		
d	Critical values	1%	5%	10%	1%	5%	10%
0.0	Asy.	0.0	0.1	0.1	0.0	0.1	0.3
	Boot	0.5	4.7	9.5	0.5	4.1	8.9
0.5	Asy.	12.2	25.5	36.7	9.5	22.0	31.9
	Boot	39.4	70.4	82.0	29.6	59.7	74.0
1.0	Asy.	26.4	50.4	64.8	21.2	42.3	55.6
	Boot	52.3	82.6	91.8	41.8	72.4	84.1

Note: The rejection rates are calculated by using the critical values of the standard normal distribution in the "Asy." rows and by using the bootstrap critical values in the "Boot" rows.

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

replications, the results in Table 2 indicate that the performance of the test is improved significantly when the bootstrap critical values are used—smaller size distortions and stronger powers are observed.

5. CONCLUSION

This paper proposes a consistent nonparametric equality test for testing whether the unknown conditional quantile functions are the same across subgroups within a survey data framework. To improve the small-sample performance of the test, we propose a modified version of the two-point wild bootstrap in the quantile context. Monte Carlo evidence shows that the performance of the test in small samples is adequate but is improved significantly when the bootstrap critical values are used.

NOTE

1. When $\hat{v}_{ij} = 0$, we add a small positive number, δ, to $\hat{v}_{ij}$ such that $b_{ij} > 0$ will be obtained. For example, we take $\delta = 10^{-8}$.

REFERENCES

Baltagi, B.H., J. Hidalgo, & Q. Li (1996) A nonparametric test for poolability using panel data. *Journal of Econometrics* 75, 345–367.

Bhattacharya, P.K. & A.K. Gangopadtyay (1990) Kernel and nearest-neighbor estimation of a conditional quantile. *Annals of Statistics* 18, 1400–1415.

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

Buchinsky, M. (1994) Changes in the U.S. wage structure 1963–1987: Application of quantile regression. *Econometrica* 62, 405–458.

Fan, J., T. Hu, & Y.K. Truong (1994) Robust non-parametric function estimation. *Scandinavian Journal of Statistics* 21, 433–446.

Fan, Y. & Q. Li (1996) Consistent model specification tests: Omitted variables and semiparametric functional forms. *Econometrica* 64, 865–890.

Gozalo, P.L. (1997) Nonparametric bootstrap analysis with applications to demographic effects in demand functions. *Journal of Econometrics* 81, 357–393.

Hall, P. (1984) Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of Multivariate Analysis* 14, 1–16.

Hall, P., R.C.L. Wolff, & Q. Yao (1999) Methods for estimating a conditional distribution function. *Journal of the American Statistical Association* 94, 154–163.

Härdle, W. & E. Mammen (1991) Bootstrap simultaneous error bars for nonparametric regression. *Annals of Statistics* 19, 778–796.

Härdle, W. & E. Mammen (1993) Comparing nonparametric versus parametric regression fits. *Annals of Statistics* 21, 1926–1947.

Koenker, R. & G. Bassett (1978) Regression quantiles. *Econometrica* 46, 33–50.

Koenker, R. & G. Bassett (1982) Robust test for heteroskedasticity based on regression quantiles. *Econometrica* 50, 43–61.

Koenker, R. & V. d'Orey (1987) Computing regression quantiles. *Applied Statistics* 36, 383–393.

Koenker, R., P. Ng, & S. Portnoy (1994) Quantile smoothing spline. *Biometrika* 81, 673–680.

Lavergne, P. (2001) An equality test across nonparametric regressions. *Journal of Econometrics* 103, 307–340.

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Li, Q. & S. Wang (1998) A simple consistent bootstrap test for a parametric regression function. *Journal of Econometrics* 87, 145–165.
Newey, W.K. & D. McFadden (1994) Large sample estimation and hypothesis testing. In R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, vol. 4. Elsevier Science B.V., 2111–2245.
Newey, W.K. & J.L. Powell (1990) Efficient estimation of linear and type I censored regression models under conditional quantile restrictions. *Econometric Theory* 6, 295–317.
Pollard, D. (1991) Asymptotic for least absolute deviation regression estimators. *Econometric Theory* 7, 186–199.
Powell, J.L. (1986) Censored regression quantiles. *Journal of Econometrics* 32, 143–155.
Powell, J.L., J.H. Stock, & T.M. Stoker (1989) Semiparametric estimation of index coefficients. *Econometrica* 57, 1403–1430.
Racine, J. & Q. Li (2004) Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* 119, 99–130.
Serfling, R.J. (2002) Approximation theorems of mathematical statistics. Wiley.
Yu, K. & M.C. Jones (1998) Local linear quantile regression. *Journal of the American Statistical Association* 93, 228–237.
Zheng, J.X. (1996) A consistent test of functional form via nonparametric estimation techniques. *Journal of Econometrics* 75, 263–289.
Zheng, J.X. (1998) A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* 14, 123–138.

APPENDIX A

Proof of Lemma 1. First, for any given $x_0 \in \mathcal{X}_1 \times R^p$, we will show that $\hat{g}(x_0)$ calculated from the pooled data converges to a weighted average of $g_1(x_0), \ldots, g_J(x_0)$. To simplify the proof, we consider the linear constant approach with an objective function

$$\hat{Q}_n(a) = n^{-1}|\tilde{H}|^{-1/2} \sum_{j=1}^{J} \sum_{i=1}^{n_j} \rho_\theta(y_{ij} - a)\tilde{K}_{00,ij}. \tag{A.1}$$

If Assumptions 1–5 hold and $n|\tilde{H}|^{1/2} \to \infty$, $\tilde{\lambda}_s \to 0$, $\tilde{h}_s \to 0$ for all s, as $n \to \infty$, we obtain

$$\begin{aligned}
&|\tilde{H}|^{-1/2} E[\rho_\theta(y_{ij} - a)\tilde{K}_{00,ij}] \\
&\quad = f_{1,j}(x_{1,0}) \int K_2(u) f_{2|1,j}(x_{2,0} + \sqrt{\tilde{H}}u | x_{1,0}) \\
&\qquad \times E[\rho_\theta(g_j(x_{1,0}, x_{2,0} + \sqrt{\tilde{H}}u) + v_{ij} - a) | x_{1,0}, x_{2,0} + \sqrt{\tilde{H}}u]\, du \\
&\qquad + O(\tilde{\lambda}_{\max}) \sum_{x_{1,ij} \neq x_{1,0}} f_{1,j}(x_{1,ij}) \int K_2(u) f_{2|1,j}(x_{2,0} + \sqrt{\tilde{H}}u | x_{1,ij}) \\
&\qquad \times E[\rho_\theta(g_j(x_{1,ij}, x_{2,0} + \sqrt{\tilde{H}}u) + v_{ij} - a) | x_{1,ij}, x_{2,0} + \sqrt{\tilde{H}}u]\, du \\
&\quad = f_j(x_{1,0}, x_{2,0}) E[\rho_\theta(g_j(x_0) + v_{ij} - a)] + O(\|\tilde{H}\|^2) + O(\tilde{\lambda}_{\max}),
\end{aligned} \tag{A.2}$$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

and similarly, $\text{var}[\hat{Q}_n(a)] = O(n^{-1}|\tilde{H}|^{-1/2})$. Hence,

$$\hat{Q}_n(a) \xrightarrow{p} Q_0(a) = \sum_{j=1}^{J} a_j f_j(x_{1,0}, x_{2,0}) E[\rho_\theta(g_j(x_0) + v_{ij} - a)]. \tag{A.3}$$

Because $\hat{Q}_n(a)$ is convex in a and $Q_0(a)$ is continuous in a, by the convexity lemma of Pollard (1991) and Theorem 2.1 of Newey and McFadden (1994), we have that $\hat{g}(x_0)$ converges in probability to $g(x_0)$, the unique minimizer of $Q_0(a)$ and the solution of the following equation:

$$\frac{dQ_0(g(x_0))}{da} = \sum_{j=1}^{J} a_j f_j(x_{1,0}, x_{2,0})[\theta - F_{v,j}(g(x_0) - g_j(x_0))] = 0. \tag{A.4}$$

By Assumption 3, taking Taylor expansion at 0 yields

$$F_{v,j}(g(x_0) - g_j(x_0)) = \theta + f_{v,j}(\zeta_{0,j})(g(x_0) - g_j(x_0)), \tag{A.5}$$

where $\zeta_{0,j}$ is between 0 and $g(x_0) - g_j(x_0)$. Substituting (A.5) into (A.4) gives

$$g(x_0) = \sum_{j=1}^{J} \lambda_j(x_0) g_j(x_0), \tag{A.6}$$

where $\lambda_j(x_0) = a_j f_{v,j}(\zeta_{0,j}) f_j(x_{1,0}, x_{2,0}) / \sum_{j=1}^{J} a_j f_{v,j}(\zeta_{0,j}) f_j(x_{1,0}, x_{2,0}) \in [0,1)$ for all j and $\sum_{j=1}^{J} \lambda_j(x_0) = 1$. Assumption 2 and the continuity of v_{ij} conditional on x_{ij} indicate that $g(x_0) = g_j(x_0)$ for all j if and only if H_0 is true.

Next, closely following Fan et al. (1994), we give the proof of the rest of Lemma 1. For the sake of convenience, in what follows, the summations begin with 1 and end with n, and the heterogeneity in the j subscript is suppressed without hurting the essence of the proof of this lemma. Hence, the modified linear quantile regression approach minimizes the following objective function:

$$\hat{Q}_n(a, b) = \sum_{i=1}^{n} \rho_\theta(y_i - a - b'(x_{2,i} - x_{2,0}))K_i, \tag{A.7}$$

where $K_i = K_1(x_{1,i} - x_{1,0}; \tilde{\lambda}) K_2(\tilde{H}^{-1/2}(x_{2,i} - x_{2,0}))$. Here we define

$$\hat{\delta}_n = n^{1/2}|\tilde{H}|^{1/4}(\hat{a} - g(x_0), \quad (\hat{b} - \nabla_{x_2} g(x_0))'\tilde{H}^{1/2})', \tag{A.8}$$

$$z_i = (1, (x_{2,i} - x_{2,0})'\tilde{H}^{-1/2})', \tag{A.9}$$

and $y_i^* = y_i - g(x_0) - \nabla_{x_2'} g(x_0)(x_{2,i} - x_{2,0})$, where $\nabla_{x_2} g(x_0)$ is the first-order partial derivative of $g(x)$ with respect to x_2 evaluated at x_0, the subscripts 1 correspond to the discrete variables, and the subscripts 2 refer to the continuous variables. Then $\hat{\delta}_n$ minimizes the following function:

$$G_n(\delta) = \sum_{i=1}^{n} [\rho_\theta(y_i^* - n^{-1/2}|\tilde{H}|^{-1/4}\delta' z_i) - \rho_\theta(y_i^*)]K_i$$

$$= E(G_n(\delta)|x) + n^{-1/2}|\tilde{H}|^{-1/4}\delta' \sum_{i=1}^{n} [\rho_\theta'(y_i^*) - E(\rho_\theta'(y_i^*|x_i))]K_i z_i + R_n.$$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Our proof will repeatedly use the following three results: (a) Taylor expansion of $g(x_{1,0}, x_{2,i})$ at $x_0 = (x_{1,0}, x_{2,0})'$ implies that

$$\zeta_i = g(x_0) + \nabla_{x_2'} g(x_0)(x_{2,i} - x_{2,0}) - g(x_{1,0}, x_{2,i}) = O(\|\tilde{H}\|^2) \tag{A.10}$$

holds uniformly over $I_0 = \{i : \|\tilde{H}^{-1/2}(x_{2,i} - x_{2,0})\| \le 1\}$; (b) the law of large numbers gives $\Pr\{\sum_{i=1}^n I(y_i^* = 0) = O(1)\} = 1$ (a.s.), where $I(y_i^* = 0) = 1$ if $y_i^* = 0$, 0 otherwise; and (c) for any $x \neq 0$

$$\rho_\theta(x - y) - \rho_\theta(x) = y[\theta - I(x < 0)] + \int_0^y [I(x \le t) - I(x \le 0)]\, dt. \tag{A.11}$$

As a result, if $\tilde{\lambda}_{\max} \to 0, \|\tilde{H}\| \to 0$, and $n|\tilde{H}|^{1/2} \to \infty$ as $n \to \infty$, we have $E(G_n(\delta)|x) = n^{-1}|\tilde{H}|^{-1/2}\delta' \sum_{i=1}^n E[\rho_\theta'(y_i^*)|x_i]K_i z_i + (\delta' S_n \delta/2) + (\tilde{R}_n/2) + o_p(1)$, where

$$S_n = n^{-1}|\tilde{H}|^{-1/2} \sum_{i=1}^n f_v(\xi_i|x_i) K_i z_i z_i' \xrightarrow{p} S(x_0),$$

$$S(x_0) = f_v(0|x_0) f(x_{1,0}, x_{2,0}) \begin{pmatrix} 1 & 0' \\ 0 & \mu_2(K_2) \end{pmatrix}, \tag{A.12}$$

and $\xi_i = g(x_0) + \nabla_{x_2'} g(x_0)(x_{2,i} - x_{2,0}) - g(x_{1,i}, x_{2,i})$; also, $\tilde{R}_n = \sum_{i=1}^n K_i \int_0^{n^{-1/2}|\tilde{H}|^{-1/4}\delta' z_i} f_v'(\bar{\xi}_i|x_i) t^2\, dt = o_p(1)$ with $\bar{\xi}_i$ lying between $\xi_i + t$ and ξ_i. Similarly, we can show that those conditions are also sufficient for $R_n = o_p(1)$.

Therefore, by the convexity lemma of Pollard (1991), we have $G_n(\delta) = (\delta' S(x_0)\delta/2) + \widetilde{W}_n'\delta + o_p(1)$ holds uniformly. Then simple calculations yield

$$\hat{\delta}_n = -S(x_0)^{-1}\widetilde{W}_n + o_p(1), \tag{A.13}$$

where

$$\widetilde{W}_n = n^{-1/2}|\tilde{H}|^{-1/4} \sum_{i=1}^n \rho_\theta'(y_i^*) K_i z_i. \tag{A.14}$$

It is easy to calculate that $E\widetilde{W}_n = \sqrt{n|\tilde{H}|^{1/2}}\,(O(\|\tilde{H}\|^2) + O(\tilde{\lambda}_{\max}))$ and $\operatorname{var}(\widetilde{W}_n) = O(1) + O(n|\tilde{H}|^{1/2}\|\tilde{H}\|^4) + O(n|\tilde{H}|^{1/2}\tilde{\lambda}_{\max}^2)$. This will complete the proof of (8). ■

LEMMA 6. *If Assumptions 1–5 and (i) of Assumption 6 hold, then under* H_0, $n|H|^{1/4} J_{n1} \xrightarrow{d} N(0, \sigma_0^2)$ *as* $n \to \infty$.

Proof. Under the null hypothesis, $u_{ij} = \varepsilon_{ij}$. For a given $j (1 \le j \le J)$, denote $J_{n1,j} = (2/n_j(n_j - 1)) \sum_{i=1}^{n_j - 1} \sum_{k=i+1} \varepsilon_{ij}\varepsilon_{kj} K_{ij,kj}$. Let $z_{ij} = (\varepsilon_{ij}, x_{ij})$; then $H_{n,j}(z_{ij}, z_{kj}) = \varepsilon_{ij}\varepsilon_{kj} K_{ij,kj}$ is symmetric and $E[H_{n,j}(z_{ij}, z_{kj})|z_{ij}] = 0$ for any $i \neq k$. The central limit theorem of Hall (1984) for a degenerate U-statistic will be used to derive the asymptotic normality result of $J_{n1,j}$. Then the asymptotic normality of $J_{n1} = (1/n\sqrt{|H|}) \sum_{j=1}^J n_j J_{n1,j}$ will follow.

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

For $k \neq i \neq s$, given the assumption, we have

$$E[G_{n,j}^2(z_{ij}, z_{kj})] = E\{\{E[H_n(z_{sj}, z_{ij})H_n(z_{sj}, z_{kj})|z_{ij}, z_{kj}]\}^2\}$$
$$= \theta^2(1-\theta)^2 E\{\varepsilon_{ij}^2 \varepsilon_{kj}^2 [E(K_{sj,ij} K_{sj,kj}|z_{ij}, z_{kj})]^2\}$$
$$= O(|H|^{3/2}) + O(\lambda_{\max}^2 |H|^{3/2}), \tag{A.15}$$

$$E[H_{n,j}^2(z_{ij}, z_{kj})] = \theta^2(1-\theta)^2 E[K_1^2(x_{1,ij} - x_{1,kj}; \lambda) K_2^2(H^{-1/2}(x_{2,ij} - x_{2,kj}))]$$
$$= \sigma_j^2 \sqrt{|H|} + O(\sqrt{|H|}\|H\|) + O(\lambda_{\max}^2 \sqrt{|H|}), \tag{A.16}$$

where $\sigma_j^2 = \theta^2(1-\theta)^2 R_2(K_2) E[f_j(x_{1,ij}, x_{2,ij})]$, $\lambda_{\max} = \max_{1 \leq s \leq q-p} \lambda_s$, and

$$E[H_{n,j}^4(z_{ij}, z_{kj})] \leq E[K_2^4(H^{-1/2}(x_{2,ij} - x_{2,kj}))] = O(\sqrt{|H|}). \tag{A.17}$$

Hence $0 \leq \{E[G_{n,j}^2(z_{ij}, z_{kj})] + n_j^{-1} E[H_{n,j}^4(z_{ij}, z_{kj})]\}/\{E[H_{n,j}^2(z_{ij}, z_{kj})]\}^2 \to 0$ if $n_j \sqrt{|H|} \to \infty$, $\|H\| \to 0$, and $\lambda_{\max} \to 0$ as $n_j \to \infty$. Then Theorem 1 of Hall (1984) gives $n_j |H|^{-1/4} J_{n1,j} \xrightarrow{d} N(0, 2\sigma_j^2)$. Immediately, we have $n|H|^{1/4} J_{n1} = (n|H|^{1/4}/n\sqrt{|H|}) \sum_{j=1}^J n_j J_{n1,j} \xrightarrow{d} N(0, \sigma_0^2)$, where $\sigma_0^2 = 2\sum_{j=1}^J \sigma_j^2 = 2\theta^2(1-\theta)^2 R_2(K_2) \sum_{j=1}^J E[f_j(x_{1,ij}, x_{2,ij})]$. ■

LEMMA 7. *If Assumptions 1–6 hold, then under H_0 we have $n|H|^{1/4} J_{n2} = o_p(1)$.*

Proof. Under H_0, $u_{ij} = \varepsilon_{ij} = I(v_{ij} \leq 0) - \theta$ and $\hat{\omega}_{kj} = I(v_{kj} \leq \hat{\xi}_{kj}) - I(v_{kj} \leq 0)$ where $\hat{\xi}_{kj} = \hat{g}(x_{kj}) - g(x_{kj})$. Simple calculation gives

$$E(n|H|^{1/4} J_{n2})^2$$
$$= J|H|^{-1/2}[E(\varepsilon_{ij}^2 \hat{\omega}_{kj}^2 K_{ij,kj}^2) + O(n)E(\varepsilon_{ij}^2 \hat{\omega}_{kj} \hat{\omega}_{mj} K_{ij,kj} K_{ij,mj})$$
$$+ O(n)E(\varepsilon_{ij} \varepsilon_{sj} \hat{\omega}_{kj}^2 K_{ij,kj} K_{sj,kj})$$
$$+ O(n^2)E(\varepsilon_{ij} \varepsilon_{sj} \hat{\omega}_{kj} \hat{\omega}_{tj} K_{ij,kj} K_{sj,tj})], \tag{A.18}$$

where $i \neq k \neq m \neq s \neq t$. The computational complexity of (A.18) stems from the dependency between $\hat{\omega}_{kj}$ and ε_{ij} because $\hat{\xi}_{kj}$ depends on v_{ij}. To overcome this problem, we rewrite $\hat{\xi}_{kj}$ by (A.12)–(A.14): $\hat{\xi}_{kj} = \sum_{l=1}^J \sum_{i=1}^{n_l} \Delta_{kj,il}$, where

$$\Delta_{kj,il} = n^{-1}|\tilde{H}|^{-1/2}(f_{v,j}(0|x_{kj}) f_j(x_{kj}))^{-1} \tilde{K}_{kj,il}(I(v_{il} \leq \xi_{kj,il}) - \theta), \tag{A.19}$$

$\tilde{K}_{kj,il}$ is the product kernel with smoothing parameters $(\tilde{\lambda}, \tilde{H})$, and $\xi_{kj,il} = g(x_{kj}) + \nabla_{x_2'} g(x_{kj}) \times (x_{2,il} - x_{2,kj}) - g(x_{il})$. The rest of the proofs follow by recognizing that only $\Delta_{kj,ij}$ is correlated with ε_{ij} for all $k \neq i$ conditional on x. Without hurting the essence of the proof, we also assume that $\hat{g}(x_{kj})$ is the leave-one-out estimate of $g(x_{kj})$ to remove the dependence between $\hat{\xi}_{kj}$ and v_{kj}; this is only assumed to shorten the proof.

First, we obtain

$$E(\varepsilon_{ij}^2 \hat{\omega}_{kj}^2 K_{ij,kj}^2) = E[\varepsilon_{ij}^2 K_{ij,kj}^2 |F_{v,j}(\hat{\xi}_{kj}|x) - \theta|]$$
$$= E\left[\varepsilon_{ij}^2 K_{ij,kj}^2 \left| f_{v,j}(0|x)\hat{\xi}_{kj} + \frac{1}{2} f_{v,j}'(\bar{\xi}_{kj}|x)\hat{\xi}_{kj}^2 \right|\right]$$
$$= O(\sqrt{|H|})[O(\|\tilde{H}\|^2) + O(n^{-1/2}|\tilde{H}|^{-1/4})], \tag{A.20}$$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

where $\tilde{\xi}_{kj}$ lies between 0 and $\hat{\xi}_{kj}$. Second, define $\tilde{\omega}_{kj} = I(v_{kj} \le \tilde{\xi}_{kj}) - I(v_{kj} \le 0)$ where $\tilde{\xi}_{kj} = \hat{\xi}_{kj} - \Delta_{kj,ij} - \Delta_{kj,sj}$. The terms $\tilde{\omega}_{kj}$ and $\tilde{\xi}_{kj}$ are independent of ε_{ij} and ε_{sj} for all $k \neq s \neq i$ conditional on x. Some calculations give $E(\varepsilon_{ij}^2 \hat{\omega}_{kj} \hat{\omega}_{mj} K_{ij,kj} K_{ij,mj}) = O(|H|\|\tilde{H}\|^4) + O(n^{-1}|H||\tilde{H}|^{-1/2})$, $E(\varepsilon_{ij}\varepsilon_{sj}\hat{\omega}_{kj}^2 K_{ij,kj} K_{sj,kj}) = O(n^{-2}|H||\tilde{H}|^{-1})$, and $E(\varepsilon_{ij}\varepsilon_{sj}\hat{\omega}_{kj}\hat{\omega}_{tj} K_{ij,kj} K_{sj,tj}) = O(n^{-2}|H||\tilde{H}|^{-1/2})$. Hence $E(n|H|^{1/4}J_{n2})^2 = o(1)$ under Assumption 6.

Similarly, we obtain $n|H|^{1/4}EJ_{n2} = O(H^{1/4}|\tilde{H}|^{-1/2}) = o(1)$ if $|H||\tilde{H}|^{-2} = o(1)$ because $E(\varepsilon_{ij}\hat{\omega}_{kj}K_{ij,kj}) = E[\varepsilon_{ij}K_{ij,kj}(\hat{\omega}_{kj} - \tilde{\omega}_{kj})] = O(n^{-1}|H|^{1/2}|\tilde{H}|^{-1/2})$ holds for $k \neq i$. This will complete the proof of this lemma. ■

LEMMA 8. *If Assumptions 1–6 hold, then under H_0 we have $n|H|^{1/4}J_{n3} = o_p(1)$.*

Proof. First we have

$$n|H|^{1/4}EJ_{n3} = |H|^{-1/4}\sum_{j=1}^{J}\frac{1}{n_j - 1}\sum_{i=1}^{n_j}\sum_{k\neq i}E(\hat{\omega}_{ij}\hat{\omega}_{kj}K_{ij,kj})$$

$$= n|H|^{1/4}[O(\|\tilde{H}\|^4) + O(n^{-1}|\tilde{H}|^{-1/2})] = o(1), \tag{A.21}$$

if $n|H|^{1/4}\|\tilde{H}\|^4 \to 0$ and $|H||\tilde{H}|^{-2} \to 0$ as $n \to \infty$. Second, we obtain

$$E(n^2|H|^{1/2}J_{n3}^2)$$

$$= J|H|^{-1/2}[E(\hat{\omega}_{ij}^2\hat{\omega}_{kj}^2K_{ij,kj}^2) + O(n)E(\hat{\omega}_{ij}^2\hat{\omega}_{kj}\,\hat{\omega}_{mj}K_{ij,kj}\,K_{ij,mj})$$

$$+ O(n^2)E(\hat{\omega}_{ij}\hat{\omega}_{sj}\hat{\omega}_{kj}\,\hat{\omega}_{tj}K_{ij,kj}\,K_{sj,tj})]$$

$$= o(1), \tag{A.22}$$

where $i \neq s \neq m \neq k \neq t$ under Assumptions 1–6, because

$$E(\hat{\omega}_{ij}^2\hat{\omega}_{kj}^2K_{ij,kj}^2) = O(\sqrt{|H|})[O(\|\tilde{H}\|^4) + O(n^{-1}|\tilde{H}|^{-1/2})],$$

$$E(\hat{\omega}_{ij}^2\hat{\omega}_{kj}\,\hat{\omega}_{mj}K_{ij,kj}K_{ij,mj}) = O(|H|)O[\|\tilde{H}\|^6 + n^{-1}|\tilde{H}|^{-1/2}\|\tilde{H}\|^2 + n^{-2}|\tilde{H}|^{-1}],$$

$$E(\hat{\omega}_{ij}\hat{\omega}_{sj}\hat{\omega}_{kj}\,\hat{\omega}_{tj}K_{ij,kj}K_{sj,tj}) = O(|H|)O(\|\tilde{H}\|^8 + n^{-2}|\tilde{H}|^{-1} + n^{-1}|\tilde{H}|^{-1/2}\|\tilde{H}\|^4).$$

Equations (A.21) and (A.22) together give $n|H|^{1/4}J_{n3} = o_p(1)$. ■

LEMMA 9. *If Assumptions 1–6 hold, then $\hat{\sigma}_0^2 \xrightarrow{p} \sigma_0^2$ as $n \to \infty$.*

Proof. For a given j, let $\hat{\sigma}_j^2 = 2/n_j(n_j - 1)\sum_{i=1}^{n_j-1}\sum_{k=i+1}K_{ij,kj}^2$, a U-statistic with $H_{n,j}(x_{ij}, x_{kj}) = K_{ij,kj}^2$, which is a symmetric function of (x_{ij}, x_{kj}) and $E(H_{n,j}^2(x_{ij}, x_{kj})) = E(K_{ij,kj}^4) = O(\sqrt{|H|}(1 + \lambda_{\max})) = o(n_j)$. Hence, Lemma 3.1 of Powell et al. (1989) yields

$$\hat{\sigma}_j^2 = E(K_{ij,kj}^2) + o_p(n_j^{-1/2}) = \sqrt{|H|}R_2(K_2)E[f_j(x_{1,ij}, x_{2,ij})]$$

$$+ O(\sqrt{|H|}\|H\|) + O(\lambda_{\max}\sqrt{|H|}) + o_p(n_j^{-1/2}).$$

Hence $\hat{\sigma}_0^2 = 2\theta^2(1-\theta)^2(1/\sqrt{|H|})\sum_{j=1}^{J}\hat{\sigma}_j^2 = \sigma_0^2 + o_p(1)$ if $\lambda_{\max} = o(1)$. ■

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

LEMMA 10. *If Assumptions 1–6 hold, then under H_1, we have*

$$J_n \xrightarrow{p} \sum_{j=1}^{J} a_j E\{[F_{v,j}(g(x_{ij}) - g_j(x_{ij})|x) - \theta]^2 f_j(x_{1,ij}, x_{2,ij})\}. \tag{A.23}$$

Proof. Lemma 1 shows that $E[\hat{g}(x) - g(x)]^2 = O(\|\tilde{H}\|^4) + O(n^{-1}|\tilde{H}|^{-1/2})$ also holds under H_1, and it is easy to show that J_{n2} and J_{n3} are both $o_p(1)$. Hence, by (4), we only need to show that

$$J_{n1} = \frac{1}{n\sqrt{|H|}} \sum_{j=1}^{J} \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \sum_{k \neq i} u_{ij} u_{kj} K_{ij,kj} = \frac{1}{n\sqrt{|H|}} \sum_{j=1}^{J} n_j J_{n1,j}$$

$$\xrightarrow{p} \sum_{j=1}^{J} a_j E\{[F_{v,j}(g(x_{ij}) - g_j(x_{ij})|x_{ij}) - \theta]^2 f_j(x_{1,ij}, x_{2,ij})\}, \tag{A.24}$$

where $u_{ij} = I(y_{ij} \leq g(x_{ij})) - \theta$.

Because $E(u_{ij}^2 u_{kj}^2 K_{ij,kj}^2) \leq E(K_{ij,kj}^2) = O(\sqrt{|H|}) + O(\lambda_{\max}\sqrt{|H|}) = o(n)$, by Lemma 3.1 of Powell et al. (1989), we obtain

$$\begin{aligned} J_{n1,j} &= E(u_{ij} u_{kj} K_{ij,kj}) + o_p\left(\frac{1}{\sqrt{n_j}}\right) \\ &= E[E(w_{ij} w_{kj}|x) K_{ij,kj}] + o_p(1) \\ &= \sqrt{|H|} E\{[F_{v,j}(g(x_{ij}) - g_j(x_{ij})|x_{ij}) - \theta]^2 f_j(x_{1,ij}, x_{2,ij})\} \\ &\quad + O(\sqrt{|H|}\|H\|^2) + O(\lambda_{\max}\sqrt{|H|}) + o_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned} \tag{A.25}$$

where $w_{ij} = I(y_{ij} \leq g(x_{ij})) - I(y_{ij} \leq g_j(x_{ij}))$. This gives (A.24) and will complete the proof of this lemma. ■

Proof of Theorem 3. Under H_1^a, $g(x) - g_j(x) = \sum_{l=1}^{J} \lambda_l(x)(g_l(x) - g_j(x)) = d_n \sum_{l=1}^{J} \lambda_l(x) \times (m_l(x) - m_j(x)) \to 0$ as $n \to \infty$. Similar to the proofs in Lemmas 7 and 8, we obtain that $n|H|^{1/4} J_{n2}$ and $n|H|^{1/4} J_{n3}$ are both $o_p(1)$. So we only need to show that $n|H|^{1/4} J_{n1} \xrightarrow{d} N(\mu, \sigma_0^2)$.

By (4), we have

$$\begin{aligned} n|H|^{1/4} J_{n1} &= |H|^{-1/4} \sum_{j=1}^{J} \frac{2n_j}{n_j(n_j - 1)} \sum_{i=1}^{n_j - 1} \sum_{k=i+1}^{n_j} u_{ij} u_{kj} K_{ij,kj} \\ &= |H|^{-1/4} \sum_{j=1}^{J} \frac{2n_j}{n_j(n_j - 1)} \sum_{i=1}^{n_j - 1} \sum_{k=i+1}^{n_j} (\varepsilon_{ij}\varepsilon_{kj} + 2\varepsilon_{ij} w_{kj} + w_{ij} w_{kj}) K_{ij,kj} \\ &= A_n + B_n + C_n, \end{aligned} \tag{A.26}$$

where Lemma 6 gives $A_n \xrightarrow{d} N(0, \sigma_0^2)$ and simple calculations give $B_n = o_p(1)$.

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Let $\delta_j(x) = \sum_{l=1}^{J} \lambda_l(x)[m_l(x) - m_j(x)]$. We obtain

$$\begin{aligned} E(C_n) &= |H|^{-1/4} \sum_{j=1}^{J} \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \sum_{k \neq i} E(w_{ij} w_{kj} K_{ij,kj}) \\ &= |H|^{1/4} \sum_{j=1}^{J} n_j E\{[F_{v,j}(g(x_{ij}) - g_j(x_{ij})|x_{ij}) - \theta]^2 f_j(x_{1,ij}, x_{2,ij})\} + o_p(1) \\ &= d_n^2 |H|^{1/4} n \sum_{j=1}^{J} \frac{n_j}{n} E[f_{v,j}^2(0|x_{ij}) \delta_j^2(x_{ij}) f_j(x_{1,ij}, x_{2,ij})] + o_p(1) \\ &\xrightarrow{n\to\infty} \mu \quad \text{if } d_n = n^{-1/2}|H|^{-1/8}, \end{aligned}$$

where $\mu = \sum_{j=1}^{J} a_j E[f_{v,j}^2(0|x_{ij}) \delta_j^2(x_{ij}) f_j(x_{1,ij}, x_{2,ij})] > 0$. Following the proof of Lemma 8, we have $\text{var}(C_n) = o(1)$. Hence, $C_n = \mu + o_p(1)$, which will complete the proof of this theorem. ■

APPENDIX B

We have bootstrap resamples $\{(x_{ij}, y_{ij}^b)\}$, where

$$y_{ij}^b = \hat{g}^*(x_{ij}) + v_{ij}^b, \tag{B.1}$$

$$v_{ij}^b = \begin{cases} \hat{v}_{ij}[a_1 I(\hat{v}_{ij} \le 0) + a_2 I(\hat{v}_{ij} > 0)] & \text{with probability } \theta \\ \hat{v}_{ij}[b_1 I(\hat{v}_{ij} < 0) + b_2 I(\hat{v}_{ij} > 0)] & \text{with probability } 1 - \theta, \end{cases} \tag{B.2}$$

$$\hat{v}_{ij} = y_{ij} - \hat{d}_{ij}, \qquad \hat{d}_{ij} = \hat{g}(x_{ij}) - g(x_{ij}), \tag{B.3}$$

where $a_1, b_2 > 0$ and $a_2, b_1 < 0$. To simplify the proofs, we assume that the leave-one-out nonparametric estimation of $g(\cdot)$ is applied.

Proof of Lemma 4. The proof of this lemma closely follows that of Lemma 1 in Appendix A, replacing y_i by y_i^b, y_i^* by y_i^{b*}, and $\hat{\delta}_n$ by $\hat{\delta}_n^b$ in the proof of Lemma 1, where $y_i^{b*} = y_i^b - \hat{g}^*(x_0) - \nabla_{x_2'} \hat{g}^*(x_0)(x_{2,i} - x_{2,0}) = v_i^b - \hat{\xi}_i^*$, $\hat{\xi}_i^* = \hat{g}^*(x_0) + \nabla_{x_2'} \hat{g}^*(x_0)(x_{2,i} - x_{2,0}) - \hat{g}^*(x_i)$, and $\hat{\delta}_n^b = n^{1/2}|\tilde{H}|^{1/4}(\hat{a} - \hat{g}^*(x_0)(\hat{b} - \nabla_{x_2} \hat{g}^*(x_0))' \tilde{H}^{1/2})'$.

By (B.1)–(B.3), $a_2 < 0 < a_1$, and $b_1 < 0 < b_2$, simple calculations give

$$\begin{aligned} E[I(y_i^{b*} \le t)|x] &= E[I(v_i^b \le \hat{\xi}_i^* + t)|x] \\ &= \theta E\{I(v_i \le \hat{d}_i, v_i \le a_1^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i) \\ &\qquad + I(v_i > \hat{d}_i, v_i \ge a_2^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i)|x\} \\ &\quad + (1 - \theta) E\{I(v_i \le \hat{d}_i, v_i \ge b_1^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i) \\ &\qquad + I(v_i \ge \hat{d}_i, v_i \le b_2^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i)|x\} \\ &= \theta - \theta E\{I(\hat{\xi}_i^* \le -t)[F_v(a_2^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i|x) \\ &\qquad - F_v(a_1^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i|x)]|x\} \\ &\quad + (1 - \theta) E\{I(\hat{\xi}_i^* > -t)[F_v(b_2^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i|x) \\ &\qquad - F_v(b_1^{-1}(\hat{\xi}_i^* + t) + \hat{d}_i|x)]|x\}. \end{aligned} \tag{B.4}$$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

Treating the last equation as a function of t and applying Taylor expansion to this function at $t = 0$ yields

$$E[I(y_i^{b*} \le t) - I(y_i^{b*} \le 0)|x]$$
$$= ctE\{f_v(\hat{d}_i|x)[\theta I(\hat{\xi}_i^* \le 0) + (1-\theta)I(\hat{\xi}_i^* > 0)]\,|x\} + O(t^2), \tag{B.5}$$

where $c = a_2^{-1} - a_1^{-1} = b_1^{-1} - b_2^{-1}$. We then obtain $\hat{\delta}_n^b = -\tilde{S}_n^{-1}\widetilde{W}_n^b + o_p(1)$, where

$$\widetilde{W}_n^b = n^{-1/2}|\tilde{H}|^{-1/4}\sum_{i=1}^n \rho_\theta'(y_i^{b*})\tilde{K}_i z_i,$$

$$\tilde{S}_n = cn^{-1}|\tilde{H}|^{-1/2}\sum_{i=1}^n \tilde{K}_i z_i z_i' E\{f_v(\hat{d}_i|x)[\theta I(\hat{\xi}_i^* \le 0) + (1-\theta)I(\hat{\xi}_i^* > 0)]|x\}.$$

If Assumptions 1–5 and (ii) of Assumption 6 hold, we have

$$\tilde{S}_n \xrightarrow{p} c\xi(x_0,\theta)f_v(x_0)f(x_{1,0}x_{2,0})\begin{pmatrix}1 & 0'\\ 0 & \mu_2(K_2)\end{pmatrix}, \tag{B.6}$$

where $\xi(x_0,\theta) = \theta$ if $\partial^2 g(x_0)/\partial x_2 \partial x_2'$ is a nonnegative definite matrix, otherwise $1-\theta$. It can be shown that $\widetilde{W}_n^b$ has the same order as $\widetilde{W}_n$ in the proof of Lemma 1, and this will complete the proof of (21).

Next, we are going to show that $[c\xi(x_0,\theta)]^{-1}e_1'\widetilde{W}_n^b = e_1'\widetilde{W}_n + o_p(1)$, where e_1 is a $(p+1)\times 1$ vector with the first element being one and zero for the rest. By (B.4), we obtain

$$E(e_1'\widetilde{W}_n^b|x) = n^{-1/2}|\tilde{H}|^{-1/4}\sum_{i=1}^n \tilde{K}_i E[\theta - I(y_i^{b*} \le 0)|x]$$
$$= \theta n^{-1/2}|\tilde{H}|^{-1/4}\sum_{i=1}^n \tilde{K}_i E\{I(\hat{\xi}_i^* \le 0)[F_v(a_2^{-1}\hat{\xi}_i^* + \hat{d}_i|x) - F_v(a_1^{-1}\hat{\xi}_i^* + \hat{d}_i|x)]|x\}$$
$$+ (1-\theta)n^{-1/2}|\tilde{H}|^{-1/4}\sum_{i=1}^n \tilde{K}_i E\{I(\hat{\xi}_i^* > 0)[F_v(b_1^{-1}\hat{\xi}_i^* + \hat{d}_i|x) - F_v(b_2^{-1}\hat{\xi}_i^* + \hat{d}_i|x)]|x\}$$
$$= cn^{-1/2}|\tilde{H}|^{-1/4}\sum_{i=1}^n \tilde{K}_i E\{f_v(\hat{d}_i|x_i)\hat{\xi}_i^*[\theta I(\hat{\xi}_i^* \le 0) + (1-\theta)I(\hat{\xi}_i^* > 0)]|x\}$$
$$+ n^{-1/2}|\tilde{H}|^{-1/4}\sum_{i=1}^n \tilde{K}_i E\{w(\hat{\xi}_i^{*2})[\theta I(\hat{\xi}_i^* \le 0) + (1-\theta)I(\hat{\xi}_i^* > 0)]|x\},$$

where $w(\cdot)$ is a uniformly bounded function over its domain. If $\|\tilde{H}\| = o(\|\tilde{H}^*\|)$ and $\tilde{\lambda}_{\max} = O(\|\tilde{H}\|^2\|\tilde{H}^*\|^2)$, we have $e_1'E\{[c\xi(x_0,\theta)]^{-1}\widetilde{W}_n^b - \widetilde{W}_n\} = O(n^{1/2}|\tilde{H}|^{1/4}\|\tilde{H}\|^2\|\tilde{H}^*\|^2)$ and $\mathrm{var}[e_1'E([c\xi(x_0,\theta)]^{-1}\widetilde{W}_n^b - \widetilde{W}_n|x)] = n|\tilde{H}|^{1/2}\|\tilde{H}\|^4 O(\|\tilde{H}^*\|^4 + n^{-1}|\tilde{H}^*|^{-1/2}\|\tilde{H}^*\|^{-4})$ because it can be shown that $[\partial^2\hat{g}^*(x_0)/\partial x_2\partial x_2'] - [\partial^2 g(x_0)/\partial x_2\partial x_2'] = O_p(\|\tilde{H}^*\|^2) +$

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press

$O_p(n^{-1/2}|\tilde{H}^*|^{-1/4}\|\tilde{H}^*\|^{-2})$ based on the results from the proof of Lemma 1. This will complete the proof of this lemma. ■

Proof of Theorem 5. The bootstrap test is given by

$$J_n^b = \frac{1}{n\sqrt{|H|}} \sum_{j=1}^{J} \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \sum_{k \neq i} \hat{u}_{ij}^b \hat{u}_{kj}^b K_{ij,kj}$$

$$= \frac{1}{n\sqrt{|H|}} \sum_{j=1}^{J} \frac{1}{n_j - 1} \sum_{i=1}^{n_j} \sum_{k \neq i} (u_{ij}^b u_{kj}^b + 2u_{ij}^b \hat{\omega}_{kj}^b + \hat{\omega}_{ij}^b \hat{\omega}_{kj}^b) K_{ij,kj}$$

$$= J_{n1}^b + 2J_{n2}^b + J_{n3}^b, \tag{B.7}$$

where $\hat{u}_{ij}^b = I(y_{ij}^b \leq \hat{g}^b(x_{ij})) - \theta = \varepsilon_{ij}^b + \hat{\omega}_{ij}^b$, $u_{ij}^b = \varepsilon_{ij}^b = I(y_{ij}^b \leq \hat{g}^*(x_{ij})) - \theta$, and $\hat{\omega}_{ij}^b = I(y_{ij}^b \leq \hat{g}^b(x_{ij})) - I(y_{ij}^b \leq \hat{g}^*(x_{ij}))$. The proposed bootstrap procedure ensures that $E^b(\varepsilon_{ij}^b) = 0$, $E^b(\varepsilon_{ij}^{b2}) = \theta(1-\theta)$, and $E^b(\varepsilon_{ij}^b \varepsilon_{kj}^b) = 0$ for $i \neq k$. Following the proofs of Lemma 6, the asymptotic normality of $n|H|^{1/4}J_{n1}^b$ can be easily obtained. Following the proofs in Lemmas 7 and 8, we can show that $n|H|^{1/4}J_{n2}^b = o_p(1)$ and $n|H|^{1/4}J_{n3}^b = o_p(1)$, respectively, although the proofs are more complicated because more terms are involved. ■

https://doi.org/10.1017/S0266466606060300 Published online by Cambridge University Press