Statistica Sinica **28** (2018), 2609-2632
doi:https://doi.org/10.5705/ss.202017.0013

A BOOTSTRAP METHOD FOR CONSTRUCTING POINTWISE AND UNIFORM CONFIDENCE BANDS FOR CONDITIONAL QUANTILE FUNTIONS

Joel L. Horowitz and Anand Krishnamurthy

Northwestern University and Cornerstone Research

Abstract: This paper is concerned with inference about the conditional quantile function in a nonparametric quantile regression model. Any method for constructing a confidence interval or band for this function must deal with the asymptotic bias of nonparametric estimators of the function. In such estimation methods, as local polynomial estimation, this is usually done through undersmoothing or explicit bias correction. The latter usually requires oversmoothing. However, there are no satisfactory empirical methods for selecting bandwidths that under- or oversmooth. This paper extends the bootstrap method of Hall and Horowitz (2013) for conditional mean functions to conditional quantile functions. The paper also shows how the bootstrap method can be used to obtain uniform confidence bands. The bootstrap method uses only bandwidths that are selected by standard methods such as cross validation and plug-in. It does not use under- or oversmoothing. The results of Monte Carlo experiments illustrate the numerical performance of the bootstrap method.

Key words and phrases: Bias, bootstrap, confidence band, nonparametric estimation, quantile estimation

1. Introduction

This paper is concerned with inference about the unknown function g in the nonparametric quantile regression model

$$Y = g(X) + \epsilon, \qquad P(\epsilon \leq 0) = \tau, \tag{1.1}$$

where X is an observed continuously-distributed explanatory variable and ϵ is an unobserved continuously-distributed random variable that is independent of X and whose τ quantile $(0 < \tau < 1)$ is 0. Hall and Horowitz (2013) (hereinafter HH) describe a bootstrap method for constructing a pointwise confidence band for the unknown function $m(x) = E(Y|X = x)$ in a nonparametric mean regression. This paper extends the bootstrap method of HH to g in the quantile regression model (1.1). The paper also shows how the bootstrap method can be used to construct a uniform confidence band for g. The method for constructing a uniform confidence band for g can be used to construct a uniform confidence band for m, but this is not done here.

Any method for constructing a pointwise or uniform confidence band for g based on a nonparametric estimate must deal with the problem of asymptotic bias. For example, a local polynomial estimate of g with a bandwidth chosen by cross-validation or plug-in methods is asymptotically biased. Denote this estimate by $\hat{g}$. The expected value of $\hat{g}$ is not g, the asymptotic distribution of the scaled estimate is not centered at g, and the true coverage probability of an asymptotic confidence interval for g that is constructed from the normal distribution in the usual way is less than the nominal probability. This problem is usually overcome by undersmoothing or explicit bias reduction. Undersmoothing consists of making the bias asymptotically negligible by using a bandwidth whose rate of convergence is faster than the asymptotically optimal rate. In explicit bias reduction, an estimate of the asymptotic bias is used to construct an asymptotically unbiased estimate of g. Most explicit bias reduction methods involve some form of oversmoothing, using a bandwidth whose rate of convergence is slower than the asymptotically optimal rate. Undersmoothing and explicit bias correction methods are also available for the conditional mean function m.

Methods based on undersmoothing or oversmoothing require a bandwidth whose rate of convergence is faster or slower than the asymptotically optimal rate. As discussed by HH, there are no attractive, effective empirical ways to choose these bandwidths. In addition, undersmoothing can produce very wiggly confidence bands, even for smooth conditional quantile or conditional mean functions. Explicit bias correction methods that rely on estimation of derivatives can also produce wiggly confidence bands.

The method presented in this paper, like the method of HH, uses bandwidths chosen by standard empirical methods such as cross validation or a plug-in rule. It does not under- or oversmooth and does not use auxiliary or other non-standard bandwidths. Instead, the method uses the bootstrap to estimate the bias of $\hat{g}$. The bootstrap estimate of the bias has stochastic noise that is comparable in size to the bias itself. However, combining a suitable quantile of the distribution of the bootstrap bias estimate with $\hat{g}$ enables us to obtain a pointwise confidence band with an asymptotic coverage probability that equals or exceeds $1-\alpha$ for any given $\alpha > 0$ at all but a user-specified fraction of the possible values of x. The exceptional points are in regions where the function g has sharp peaks or troughs that cause the bias of $\hat{g}$ to be unusually large. These regions are typically visible in a plot of $\hat{g}$ and can also be found through a theoretical analysis. An asymptotic uniform confidence band that has no exceptional points is obtained by replacing the bootstrap bias estimate with an upper bound on the estimated bias.

This paper differs from HH in two important ways: we obtain confidence bands for g that are uniform in x, whereas HH obtain only pointwise bands; our bootstrap method is different from that of HH. To avoid complications caused by the non-smoothness of the quantile objective function, we apply the bootstrap to the leading term of the asymptotic bias of the quantile regression estimator. We do not estimate the conditional quantile function from the bootstrap sample. In contrast, the objective function of a mean-regression model is smooth. This enables HH to estimate the conditional mean function and its bias directly from the bootstrap sample.

Methods that use undersmoothing have been described by Bjerve, Doksum and Yandell (1985); Hall (1992); Hall and Owen (1993); Neumann (1995); Chen (1996); Neumann and Polzehl (1998); Picard and Tribouley (2000); Chen, Härdle and Li (2003); Claeskens and Van Keilegom (2003); Härdle et al. (2004); and McMurry and Politis (2008). Methods based on oversmoothing have been described by Härdle and Bowman (1988); Härdle and Marron (1991); Hall (1992); Eubank and Speckman (1993); Sun and Loader (1994); Härdle, Huet and Jolivet (1995); Xia (1998); and Schucany and Sommers (1977). Calonico, Cattaneo and Farrell (2016) describe an explicit bias correction method for conditional mean functions that does not require oversmoothing or an auxiliary bandwidth. It is not known whether this method can be extended to conditional quantile functions.

There is also a large literature on bootstrap methods for parametric quantile regression models. See, for example, De Angelis, Hall and Young (1993); Hahn (1995); Horowitz (1998); Feng, He and Hu (2011); Aguirre and Dominguez (2013); Galvao and Montes-Rojas (2015); and Hagemann (2017).

Section 2 of this paper presents an informal description of our method. The method is similar in some respects to that of HH for conditional mean functions, but the non-smoothness of quantile estimators presents problems that are different from those involved in estimating conditional mean functions. These require a separate treatment and modifications of parts of the method of HH. Section 2 also outlines the extension of our method to a heteroskedastic version of model (1.1). Section 3 presents theoretical results. Section 4 presents simulation results that illustrate the numerical performance of the method. Conclusions are in Section 5. The proofs of theorems are in an online supplementary appendix.

2. Informal Description of the Method

Let $\{Y_i, X_i : i = 1, \ldots, n\}$ denote an independent random sample of observations from the distribution of (Y, X) in model (1.1). Let $\hat{g}(x)$ denote a local polynomial nonparametric estimator of based on bandwidth h. Denote the scaled asymptotic bias and variance of $\hat{g}(x)$, respectively, by $\beta(x) = \lim_{n\to\infty} E(nh)^{1/2} \{\hat{g}(x) - g(x)\}$ and $\sigma^2_{\hat{g}}(x) = \lim_{n\to\infty} (nh)Var(\hat{g}(x))$. Assume that $[\hat{g}(x) - E\{\hat{g}(x)\}]/\sigma_{\hat{g}}(x) \overset{d}{\to} N(0,1)$ as $n \to \infty$. To minimize the complexity of the discussion in the remainder of this paper, we assume that X is a scalar random variable and $\hat{g}$ is a local linear quantile regression estimator. The main results of the paper continue to apply if X is a vector or $\hat{g}$ is a local polynomial estimator of odd degree different from 1. This paper does not treat series estimators. The local linear quantile estimation procedure is described in Step 1 in Section 2.1. To avoid boundary effects we restrict attention to a compact set S that is contained in an open subset of the support of X. Let h denote the bandwidth used in local linear estimation of g.

If $\beta(x)$ were known, an asymptotic $1 - \alpha_0$ confidence interval for $g(x)$ would be

$$\hat{g}(x) - \frac{z_{1-\alpha/2}\, \sigma_{\hat{g}}(x)}{(nh)^{1/2}} \leq g(x) \leq \hat{g}(x) + \frac{z_{1-\alpha/2}\, \sigma_{\hat{g}}(x)}{(nh)^{1/2}},$$

where $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution, $\alpha = \alpha(x, \alpha_0)$ satisfies

$$\Phi\left(z_{1-\alpha/2} - \frac{\beta(x)}{\sigma_{\hat{g}}(x)}\right) - \Phi\left(-z_{1-\alpha/2} - \frac{\beta(x)}{\sigma_{\hat{g}}(x)}\right) = 1 - \alpha_0, \qquad (2.1)$$

and Φ is the normal distribution function. In applications, $\beta(x)$ and $\sigma_{\hat{g}}(x)$ are unknown. Let $\hat{\sigma}_{\hat{g}}(x)$ be the estimate of $\sigma_{\hat{g}}(x)$ described in Section 2.1. Let $\hat{\lambda}(x)$ denote the bootstrap estimate of $\beta(x)/\sigma_{\hat{g}}(x)$ obtained in Step 5 of the procedure described in Section 2.1, and let $\hat{\alpha}(x, \alpha_0)$ denote the solution in α to

$$\Phi(z_{1-\alpha/2} - \hat{\lambda}(x)) - \Phi(-z_{1-\alpha/2} - \hat{\lambda}(x)) = 1 - \alpha_0,$$

For $\xi \in [0, 1]$, let $\hat{\alpha}_\xi(\alpha_0)$ be the ξ quantile of points in the set $\{\hat{\alpha}(x, \alpha_0) : x \in S\}$. Take $\hat{z}(\alpha_0) = z_{1-\hat{\alpha}_\xi(\alpha_0)/2}$. Construct the pointwise confidence band

$$B_n(\hat{\alpha}_\xi(\alpha_0)) = \left\{(x, y) :\, \hat{g}(x) - \frac{\hat{z}(\alpha_0)\, \hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}} \leq y \leq \hat{g}(x) + \frac{\hat{z}(\alpha_0)\, \hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}}\right\}.$$

It is shown in Section 3.3 that B_n has asymptotic coverage probability equal to or greater than $1 - \alpha_0$, except for a proportion ξ of points $x \in S$. An argument like that in Section 2.6 of HH shows that the exceptional points occur in regions where $|g''(x)|$ is large. These points typically occur near peaks and troughs of $g(x)$ and can be identified from a graph of $\hat{g}(x)$.

The critical value $\hat{z}(\alpha_0) = z_{1-\hat{\alpha}_\xi(\alpha_0)/2}$ is greater than $z_{1-\alpha_0/2}$. Therefore, the confidence band $B_n(\hat{\alpha}_\xi(\alpha_0))$ is wider than a band that uses the critical value $z_{1-\alpha_0/2}$. This enables $B_n(\hat{\alpha}_\xi(\alpha_0))$ to accommodate the asymptotic bias of $\hat{g}(x)$. The parameter ξ is selected by the user and controls the fraction of points $x \in S$ at which the asymptotic coverage probability of the confidence band $B_n(\hat{\alpha}_\xi(\alpha_0))$ is at least $1 - \alpha_0$. This control is an important advantage of our method over undersmoothing and explicit bias correction. As is illustrated in Section 4, the latter two methods have poor coverage accuracy but provide no information about the extent of this inaccuracy or the ability to control it. At the cost of a wider confidence band, the fraction of points at which our method undercovers can be reduced to zero asymptotically by constructing the uniform band described in Step 7 of Section 2.1.

To construct a uniform confidence band for g, define

$$\hat{\lambda}_{\max} = \max_{x \in S} \hat{\lambda}(x) \qquad \text{and} \qquad \hat{\lambda}_{\min} = \min_{x \in S} \hat{\lambda}(x).$$

Let W_1 be the mean-zero Gaussian process defined in Section 3.1, and let $\hat{t}_U$ satisfy

$$P\left(-\hat{t}_U - \hat{\lambda}_{\min} \leq W_1\left(\frac{x}{h}\right) \leq \hat{t}_U - \hat{\lambda}_{\max}, \quad \forall x \in S \right) = 1 - \alpha_0,$$

where h is the bandwidth used for local linear quantile estimation of g. It is shown in Section 3.3 that

$$B_U(\alpha_0) \equiv \left\{ (y, x) : \hat{g}(x) - \frac{\hat{t}_U \, \hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}} \leq y \leq \hat{g}(x) + \frac{\hat{t}_U \, \hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}}; x \in S \right\}$$

is an asymptotic uniform confidence band for g whose coverage probability equals or exceeds $1 - \alpha_0$.

2.1. The estimation procedure

This section provides a step-by-step explanation of the method for constructing B_n and B_U. Steps 1–6 are broadly similar to those of HH, but their implementation details differ because of differences between the mean-regression model of HH and the quantile regression model considered here. Step 7 constructs a uniform confidence band and is new.

Step 1: *Local linear estimation of g and estimation of $\sigma^2_{\hat{g}}$.* Let K be a kernel function and h be a possibly random bandwidth. For any real v, let $K_h(v) = K(v/h)$. Define the check function

$$\rho_\tau(v) = v(\tau - I(v \le 0)),$$

where $0 < \tau < 1$ and I is the indicator function. The local linear estimator of $g(x)$ is $\hat{g}(x) = \hat{b}_0$, where

$$(\hat{b}_0, \hat{b}_1) = \arg\min_{b_0, b_1} \sum_{i=1}^{n} \rho_\tau(Y_i - b_0 - b_1(X_i - x))K_h(X_i - x).$$

Fan, Hu and Truong (1994) and Yu and Jones (1998) describe properties of this estimator.

To obtain $\hat{\sigma}_{\hat{g}}(x)$, let $\hat{f}_X(x)$ be a consistent kernel nonparametric estimator of $f_X(x)$, the probability density function of X at x. Let $\tilde{\epsilon}_i = Y_i - \hat{g}(X_i)$ be the residuals from estimating model (1.1), and let $\hat{f}_\epsilon(0)$ be a consistent kernel nonparametric estimator of $f_\epsilon(0)$, the probability density of ϵ at 0. Specifically,

$$\hat{f}_\epsilon(0) = (nh_\epsilon)^{-1} \sum_{i=1}^{n} K_{h_\epsilon}(\tilde{\epsilon}_i),$$

where h_ϵ is a bandwidth. With

$$B_K = \int K^2(v)\, dv,$$

it is shown in Section 3.2 that the scaled variance of the asymptotic distribution of $\hat{g}(x)$ is

$$\sigma^2_{\hat{g}}(x) = (nh)Var(\hat{g}(x)) = \frac{\tau(1-\tau)B_K}{f_X(x)\{f_\epsilon(0)\}^2}.$$

The scaled variance can be estimated by replacing $f_X(x)$ and $f_\epsilon(0)$ with their consistent estimators to obtain

$$\hat{\sigma}^2_{\hat{g}}(x) = \frac{\tau(1-\tau)B_K}{\hat{f}_X(x)\{\hat{f}_\epsilon(0)\}^2}.$$

Step 2: *Compute centered residuals.* Let q_n be the τ quantile of the residuals $\{\tilde{\epsilon}_i\}$,

$$q_n = \inf\left(q : n^{-1} \sum_{i=1}^{n} I(\tilde{\epsilon}_i \le q) \ge \tau\right).$$

The centered residuals are $\hat{\epsilon}_i = \tilde{\epsilon}_i - q_n$, and the τ quantile of centered residuals $\{\hat{\epsilon}_i\}$ is 0.

Step 3: *Construct the bootstrap resample.* The bootstrap resample is $\{Y_i^*, X_i : i = 1, \dots, n\}$, where $Y_i^* = \hat{g}(X_i) + \epsilon_i^*$ and the ϵ_i^*s are obtained by sampling the $\hat{\epsilon}_i$s randomly with replacement. The X_is are not resampled.

Step 4: *Compute the bootstrap estimate of the asymptotic bias of* $\hat{g}(x)$. Let there be B bootstrap resamples indexed by $b = 1, \ldots, B$. For resample b, define

$$T^*_{nb}(x) = n^{-1} \sum_{i=1}^{n} \{1 - \tau^{-1} I(Y_i^* \le \hat{b}_0 + \hat{b}_1(X_i - x))\} K_h(X_i - x).$$

$T^*_{nb}(x)$ is a bootstrap analog of the first-order condition in equation (A.8) of the supplementary appendix. Let $E^*(T^*_{nb})$ denote the bootstrap expectation of T^*_{nb} conditional on the data. Estimate $E^*(T^*_{nb})$ by

$$\hat{T}_n(x) = B^{-1} \sum_{b=1}^{B} T^*_{nb}(x).$$

$\hat{T}_n(x)$ converges almost surely to $E^*(T^*_{nb})$ and can be made arbitrarily close to $E^*(T^*_{nb})$ by making B sufficiently large. As $n \to \infty$, the bias of $(nh)^{1/2}\{\hat{g}(x) - g(x)\}$ converges to $\beta(x) = \lim_{n\to\infty} (nh)^{1/2} E\{\hat{g}(x) - g(x)\}$. Equation (3.2) gives an analytic expression for $\beta(x)$. The bootstrap estimate of $\beta(x)$ is

$$\hat{\beta}(x) = h^{-1} \left\{ \frac{\tau}{(1-\tau)B_K} \right\}^{1/2} \frac{\hat{\sigma}_{\hat{g}}(x)}{\{\hat{f}_X(x)\}^{1/2}} \hat{T}_n(x).$$

In contrast to HH, we do not form a bootstrap estimate of $g(x)$. Instead, we form a bootstrap estimate of the asymptotic form of $E(nh)^{1/2}\{\hat{g}(x) - g(x)\}$ from the analytic expression for this form. This expression is given by equation (3.4).

The variance of $\hat{\beta}(x)$ converges to zero at the same rate as $\beta(x)^2$. Therefore, $\hat{\beta}(x)$ is not consistent for $\beta(x)$ in the sense that $\hat{\beta}(x)/\beta(x)$ does not converge in probability to one. The methods for finding pointwise and uniform confidence bands for g take account of this inconsistency. See Steps 6 and 7 below.

Step 5: *Obtain the normalized estimate of the bias and effective significance level.* The normalized bias of $\hat{g}(x)$ is defined as $\lambda(x) = \beta(x)/\sigma_{\hat{g}}(x)$ and is estimated by $\hat{\lambda}(x) = \hat{\beta}(x)/\hat{\sigma}_{\hat{g}}(x)$. The effective significance level at point x, $\hat{\alpha}(x, \alpha_0)$, is the solution in α to the equation

$$\Phi(z_{1-\alpha/2} - \hat{\lambda}(x)) - \Phi(-z_{1-\alpha/2} - \hat{\lambda}(x)) = 1 - \alpha_0.$$

Step 6: *Construct a pointwise confidence band for* g. Let $\xi \in [0, 1]$. Let $\hat{\alpha}_\xi(\alpha_0)$ be the ξ quantile of points in the set $\{\hat{\alpha}(x, \alpha_0) : x \in S\}$. Define $\hat{z}(\alpha_0) =$

$z_{1-\hat{\alpha}_\xi(\alpha_0)/2}$. Construct the pointwise confidence band

$$B_n(\hat{\alpha}_\xi(\alpha_0)) = \left\{(x,y) : \hat{g}(x) - \frac{\hat{z}(\alpha_0)\hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}} \le y \le \hat{g}(x) + \frac{\hat{z}(\alpha_0)\hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}}\right\}.$$

It is shown in Section 3.3 that the pointwise band $B_n[\hat{\alpha}_\xi(\alpha_0)]$ covers $g(x)$ with probability at least $1-\alpha_0$ except for a proportion ξ of points $x \in S$. Specifically, for $\alpha(x,\alpha_0)$ as in (2.1), let $\alpha_\xi(\alpha_0)$ denote the ξ quantile of the points $\{\alpha(x,\alpha_0) : x \in S\}$. Define the set

$$R_\xi(\alpha_0) = \{x \in S : \alpha(x,\alpha_0) > \alpha_\xi(\alpha_0)\}. \tag{2.2}$$

Then

$$\liminf_{n\to\infty} P((x, g(x)) \in B_n(\hat{\alpha}_\xi(\alpha_0))) \ge 1 - \alpha_0$$

for each $x \in R_\xi(\alpha_0)$. $R_\xi(\alpha_0)$ is the set of points $x \in S$ on which the pointwise confidence band has an asymptotic coverage probability of at least $1-\alpha_0$. It contains a fraction $1-\xi$ of points $x \in S$. Specifically, let $\|S\|$ and $\|R\|$, respectively, denote the Lebesgue measures of the sets S and $R_\xi(\alpha_0)$. Then $\|R\|/\|S\| = 1-\xi$.

Step 7: *Construct a uniform confidence band for g.* Define

$$\hat{\lambda}_{\max} = \max_{x\in S} \hat{\lambda}(x), \tag{2.3}$$

$$\hat{\lambda}_{\min} = \min_{x\in S} \hat{\lambda}(x). \tag{2.4}$$

Let W_1 denote the mean-zero Gaussian process defined in Theorem 3.1 in Section 3.1. Let $\hat{t}_U$ be the solution in t to

$$P\left(-t - \hat{\lambda}_{\min} \le W_1\left(\frac{x}{h}\right) \le t - \hat{\lambda}_{\max}, \quad \forall x \in S\right) = 1 - \alpha_0. \tag{2.5}$$

The asymptotic uniform confidence band is

$$B_U(\alpha_0) \equiv \left\{(y,x) :\ \hat{g}(x) - \frac{\hat{t}_U\hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}} \le y \le \hat{g}(x) + \frac{\hat{t}_U\hat{\sigma}_{\hat{g}}(x)}{(nh)^{1/2}};\ x \in S\right\}.$$

The quantities $\hat{\lambda}_{\max}$ and $\hat{\lambda}_{\min}$ can be computed by replacing S in (2.3) and (2.4) with a fine grid of equally spaced points. The critical value $\hat{t}_U$ can be computed by replacing S in (2.5) with the grid.

2.2. Heteroskedasticity

A heteroskedastic version of (1.1) is

$$Y = g(X) + \sigma(X)\epsilon, \qquad P(\epsilon \le 0) = \tau, \tag{2.6}$$

where $\sigma(\cdot)$ is a scale function and ϵ is independent of X. Identification of σ

requires normalizing the scale of ϵ. This is done by setting the interquartile range (IQR) of ϵ equal to 1. Then

$$\sigma(x) = IQR(Y|X = x).$$

Let $\hat{g}(x)$ be the local linear quantile regression estimate of $g(x)$ and $\hat{\sigma}(x)$ be a consistent nonparametric estimate of $IQR(Y|X = x)$. The residuals of model (2.5) are

$$\breve{\epsilon}_i = \frac{Y_i - \hat{g}(X_i)}{\hat{\sigma}(X_i)}.$$

The centered residuals of (2.6) are as in Step 2 after replacing $\tilde{\epsilon}_i$ with $\breve{\epsilon}_i$. The estimate of the scaled asymptotic variance of the estimate of $g(x)$ in (2.6) is

$$\hat{\sigma}^2_{\hat{g}}(x) = \frac{\hat{\sigma}^2(x)\tau(1-\tau)B_K}{\hat{f}_X(x)\{\hat{f}_\epsilon(0)\}^2},$$

where $\hat{f}_\epsilon$ is now based on the $\breve{\epsilon}_i$'s. Bootstrap sampling is done by setting

$$Y_i^* = \hat{g}(X_i) + \hat{\sigma}(X_i)\epsilon_i^*,$$

where the ϵ_i^*'s are sampled randomly with replacement from the centered $\breve{\epsilon}_i$'s. Steps 4–7 for construction of pointwise and uniform confidence bands remain as in Section 2.1 but with the foregoing modifications of $\hat{\sigma}^2_{\hat{g}}(x)$ and the bootstrap sampling procedure. We conjecture that our method can be extended to the generalization of model (1.1) in which $P(\epsilon \leq 0) = \tau$ is replaced by $P(\epsilon \leq 0|X) = \tau$, but we do not analyze this extension here.

3. Theoretical Results

This section presents theorems giving conditions under which the pointwise and uniform confidence bands constructed in Steps 6–7 of Section 2.1 have the claimed coverage properties when $\hat{g}$ is a local linear quantile regression estimator. Theorem 3.1 shows that $(nh)^{1/2}\{\hat{g}(x) - g(x)\}$ is approximated sufficiently accurately by the sum of its asymptotic bias and a mean-zero Gaussian process. Theorem 3.2 shows that a similar approximation applies to the bootstrap bias estimator. These two approximations are combined in Theorem 3.4 to show that the bootstrap procedure of Section 2.1 yields pointwise confidence intervals with the coverage probabilities explained in Step 6 of Section 2.1. Theorem 3.5 shows that the pointwise confidence intervals of Theorem 3.4 can be widened to construct a uniform confidence band.

We need some assumptions:

Assumption 1: (i) The data $\{Y_i, X_i : i = 1, \ldots, n\}$ are an independent random sample from model (1.1);

(ii) X in (1.1) has compact support;

(iii) ϵ in (1.1) is independent of X and $P(\epsilon \leq 0) = \tau$ for some $\tau \in (0, 1)$.

Assumption 2: (i) The distribution of X is absolutely continuous with respect to Lebesgue measure with probability density function f_X;

(ii) f_X is bounded away from 0 on $\text{supp}(X)$ and twice continuously differentiable on the interior of $\text{supp}(X)$.

Assumption 3: (i) The distribution of ϵ is absolutely continuous with respect to Lebesgue measure with probability density function f_ϵ;

(ii) f_ϵ is twice continuously differentiable and $f_\epsilon(0) > 0$.

Assumption 4: (i) The function g in (1.1) is three times continuously differentiable on the interior of $\text{supp}(X)$;

(ii) $\hat{g}$ is a local linear quantile regression estimator of g.

(iii) There is a compact set $G \in \mathbb{R}^2$ such that $(g(x), g'(x)) \in G$ for each x.

Assumption 5: The kernel K is a probability density function with support $[-1, 1]$, symmetric around 0, and twice continuously differentiable on $(-1, 1)$.

Assumption 6: The bandwidth h used to construct $\hat{g}$ satisfies

(i) $h = \hat{d}n^{-1/5}$, where $\hat{d}$ is a function of the data $\{Y_i, X_i : i = 1, \ldots, n\}$ and $\hat{d} \xrightarrow{p} d_0$ as $n \to \infty$ for some finite constant $d_0 > 0$.

(ii) There exists a finite constant $D_1 > 0$ such that

$$P(|\hat{d} - d_0| > n^{-D_1}) \to 0 \text{ as } n \to \infty.$$

(iii) There are constants D_2 and D_3 such that $0 < D_2 < D_3 < 1$ and

$$P(n^{-D_3} \leq h \leq n^{-D_2}) = 1 - O(n^{-C})$$

as $n \to \infty$ for all finite $C > 0$.

Assumption 1 defines the data generation process. Assumptions 2–4 are smoothness assumptions. Assumption 5 specifies standard properties of K. Assumption 6 is satisfied by standard bandwidth choice methods such as cross-validation and plug-in methods. Under these assumptions, local linear estimates of g obtained using the random bandwidth h and the deterministic bandwidth $h_0 = d_0 n^{-1/5}$ are asymptotically equivalent. See Lemma A.1 in the supplementary appendix.

3.1. Asymptotic approximations to $(nh)^{1/2}\{\hat{g}(x) - g(x)\}$ and the bootstrap bias estimate

The asymptotic coverage probabilities of the pointwise and uniform confidence bands defined in Steps 6 and 7 of Section 2.1 depend on strong asymptotic approximations to $(nh)^{1/2}\{\hat{g}(x) - g(x)\}$ and the bootstrap estimate of $E(nh)^{1/2}$ $\{\hat{g}(x) - g(x)\}$, as follows.

Theorem 1. *Let Assumptions* 1–6 *hold. If*

$$\psi_0(x) = \frac{\{\tau(1-\tau)B_K\}^{1/2}}{f_X(x)^{1/2} f_\epsilon(0)},$$

$$\kappa_2 = \int v^2 K(v)dv,$$

and $h_0 = d_0 n^{-1/5}$, *there exists a Gaussian process* $W_1(x)$ *defined on the same probability space as the data such that* $E\{W_1(x)\} = 0$ *for all* $x \in S$, $E\{W_1(x)\}^2 = 1$ *for all* $x \in S$, *and for any* $\eta > 0$

$$\lim_{n\to\infty} P\left(\sup_{x\in S}\left|(nh)^{1/2}\{\hat{g}(x)-g(x)\}-\left\{\frac{d_0^{5/2}\kappa_2}{2}g''(x)+\psi_0(x)W_1\left(\frac{x}{h_0}\right)\right\}\right| > \eta\right) = 0, \quad x \in S.$$

It follows from Theorem 1 that for each $x \in S$,

$$(nh)^{1/2}\{\hat{g}(x) - g(x)\} \xrightarrow{d} N(\mu_g, V_g(x)), \tag{3.1}$$

where

$$\mu_g = \frac{d_0^{5/2}\kappa_2}{2} g''(x), \tag{3.2}$$

$$V_g(x) = \psi_0^2(x). \tag{3.3}$$

Moreover, asymptotically,

$$E\hat{g}(x) - g(x) = \frac{h_0^2 \kappa_2}{2} g''(x), \tag{3.4}$$

$$(nh_0)Var(\hat{g}(x)) = \sigma_{\hat{g}}^2(x) = \frac{\tau(1-\tau)B_K}{f_X(x)\{f_\epsilon(0)\}^2}. \tag{3.5}$$

Properties (3.1)–(3.5) were obtained previously by Fan, Hu and Truong (1994) and Yu and Jones (1998).

The quantity $(nh_0)Var(\hat{g}(x))$ can be estimated consistently by replacing $f_X(x)$ and $f_\epsilon(0)$ on the right-hand sides of (3.3) and (3.5) by the consistent estimators $\hat{f}_X(x)$ and $\hat{f}_\epsilon(0)$. Bootstrap estimation of $E(nh)^{1/2}\{\hat{g}(x)-g(x)\}$ relies on the following strong approximation.

Theorem 2. *Let Assumptions* 1–6 *hold. Let* E^* *denote the bootstrap expectation conditional on the data. If*

$$A_1(x) = \tau^{-1} d_0^{5/2} f_\epsilon(0) f_X(x),$$
$$A_2(x) = \left\{ \frac{1-\tau}{\tau} B_K f_X(x) \right\}^{1/2},$$

then

(i) *For all* $x \in S$, *and any* $\eta > 0$

$$\lim_{n\to\infty} P\left(\sup_{x\in S} \left| \left(\frac{n}{h}\right)^{1/2} E^*\{T^*_{nb}(x)\} + \frac{A_1(x)\kappa_2}{2} g''(x) - A_2(x) W_1\left(\frac{x}{h_0}\right) \right| > \eta \right) = 0.$$

(ii) *There exists a Gaussian process* $\Delta(x)$ *such that* $E\{\Delta(x)\} = 0$ *for all* $x \in S$, $E\{\Delta(x)\}^2 = 1$ *for all* $x \in [0,1]$, *and for any* $\eta > 0$

$$\lim_{n\to\infty} P\left(\sup_{x\in S} \left| \hat{\lambda}(x) - \left\{ \frac{\beta(x)}{\sigma_{\hat{g}}(x)} + \Delta(x) \right\} \right| > \eta \right) = 0, \quad x \in S.$$

For any $\alpha \in (0,1)$ and $x \in S$ take

$$\hat{\pi}(x,\alpha) = \Phi(z_{1-\alpha/2} - \hat{\lambda}(x)) - \Phi(-z_{1-\alpha/2} - \hat{\lambda}(x)).$$

A corollary to Theorem 3.2 is used to establish the asymptotic coverage probabilities of the confidence bands constructed in Steps 6 and 7 of Section 2.1.

Corollary 1. *If Assumptions* 1–6 *hold, then for any* $\eta > 0$ *and* $0 < \alpha < 1$,

$$\lim_{n\to\infty} P\left(\sup_{x\in S} \left| \hat{\pi}(x,\alpha) - \left\{ \Phi\left(z_{1-\alpha/2} - \frac{\beta(x)}{\sigma_{\hat{g}}(x)} - \Delta(x) \right) - \Phi\left(-z_{1-\alpha/2} - \frac{\beta(x)}{\sigma_{\hat{g}}(x)} - \Delta(x) \right) \right\} \right| > \eta \right) = 0, \quad x \in S.$$

3.2. Coverage probabilities of confidence bands

This section shows that the pointwise and uniform confidence bands constructed in Steps 6 and 7 of Section 2.1 have asymptotic coverage probabilities of at least $1-\alpha_0$. We use the following notation. Let $\alpha_0 \in (0,1/2)$. Define $\hat{\alpha}(x,\alpha_0)$ as in Step 5. Take $T(x,\alpha_0)$ as the solution in T to

$$\Phi\left(T - \frac{\beta(x)}{\sigma_{\hat{g}}(x)} - \Delta(x)\right) - \Phi\left(-T - \frac{\beta(x)}{\sigma_{\hat{g}}(x)} - \Delta(x)\right) = 1 - \alpha_0,$$

and $A(x, \alpha_0) = 2\{1 - \Phi[T(x, \alpha_0)]\}$. Let $\alpha(x, \alpha_0)$ be the solution in a to

$$\Phi\left(z_{1-a/2} - \frac{\beta(x)}{\sigma_{\hat{g}}(x)}\right) - \Phi\left(-z_{1-a/2} - \frac{\beta(x)}{\sigma_{\hat{g}}(x)}\right) = 1 - \alpha_0,$$

and $\alpha_\xi(\alpha_0)$ denote the ξ-level quantile of points in the set $\{\alpha(x, \alpha_0) : x \in S\}$. Define $R_\xi(\alpha_0)$ as in (2.2). The following theorem gives the asymptotic coverage probability of the pointwise confidence band constructed in Step 6.

Theorem 3. *If Assumptions* 1–6 *hold, for all* $C > 0$ *and* $\eta > 0$,

(i) $\lim_{n\to\infty} P\left(\sup_{x\in S,\, |\Delta(x)|\le C} |\hat{\alpha}(x, \alpha_0) - A(x, \alpha_0)| > \eta\right) = 0.$

(ii) $\lim_{n\to\infty} P(\hat{\alpha_\xi}(\alpha_0) \le \alpha_\xi(\alpha_0)) = 1.$

(iii) $\liminf_{n\to\infty} P\big(\{x, g(x)\} \in B_n(\hat{\alpha}_\xi(\alpha_0))\big) \ge 1 - \alpha_0$ *for each* $x \in R_\xi(\alpha_0)$.

Theorem 3.4(iii) shows that the pointwise band $B_n[\hat{\alpha}_\xi(\alpha_0)]$ covers $g(x)$ with probability at least $1 - \alpha_0$ except for a proportion ξ of points $x \in S$.

Now consider the uniform confidence band constructed in Step 7 of Section 2.1. It follows from Theorem 3.1 that, up to asymptotically negligible terms

$$\frac{(nh)^{1/2}\{\hat{g}(x) - g(x)\}}{\sigma_{\hat{g}}(x)} = \frac{\beta(x)}{\sigma_{\hat{g}}(x)} + W_1\left(\frac{x}{h_0}\right)$$

uniformly over $x \in S$. If $\lambda(x) = \beta(x)/\sigma_{\hat{g}}(x)$ were known, an asymptotic uniform $1 - \alpha_0$ confidence band for g would be

$$-t_U \le \lambda(x) + W_1\left(\frac{x}{h_0}\right) \le t_U,$$

where t_U is the solution in t to

$$P\left(-t - \lambda(x) \le W_1\left(\frac{x}{h_0}\right) \le t - \lambda(x), \quad \forall\, x \in S\right) = 1 - \alpha_0,$$

If $\lambda_{\max} = \max_{x\in S} \lambda(x)$ and $\lambda_{\min} = \min_{x\in S} \lambda(x)$, then

$$\begin{aligned} &P\left(-t_U - \lambda_{\min} \le W_1\left(\frac{x}{h_0}\right) \le t_U - \lambda_{\max}, \quad \forall\, x \in S\right) \\ &\le P\left(-t_U - \lambda(x) \le W_1\left(\frac{x}{h_0}\right) \le t_U - \lambda(x), \quad \forall\, x \in S\right). \end{aligned}$$

Therefore, asymptotically,

$$P\left(-t_U \le \frac{(nh)^{1/2}\{\hat{g}(x) - g(x)\}}{\sigma_{\hat{g}}(x)} \le t_U, \quad \forall\, x \in S\right) \ge 1 - \alpha_0 \tag{3.6}$$

Table 1. Simulation results for $\tau = 0.25$.

Method	n	j	Prop. with Cov. Prob. ≥ 0.95	Av. Abs. Error of Cov. Prob.	Av. Width
Bootstrap method of Sec. 2.1	100	1	0.76	0.034	1.76
		2	0.73	0.022	1.57
		3	0.97	0.024	1.26
	500	1	0.89	0.049	1.01
		2	1.0	0.033	0.83
		3	1.0	0.028	0.59
	1,000	1	0.92	0.051	0.80
		2	0.95	0.032	0.62
		3	0.94	0.026	0.44
Undersmooth	100	1	0	0.097	1.24
		2	0	0.07	1.18
		3	0.03	0.04	1.16
	500	1	0.30	0.065	0.68
		2	0.22	0.016	0.71
		3	0.35	0.01	0.66
	1,000	1	0.32	0.062	0.52
		2	0.41	0.014	0.57
		3	0.4	0.009	0.56
Bias Corr.	100	1	0	0.18	1.38
		2	0	0.19	1.37
		3	0	0.15	1.31
	500	1	0	0.11	1.07
		2	0.08	0.052	0.91
		3	0.06	0.036	0.078
	1,000	1	0	0.068	0.88
		2	0.08	0.14	0.81
		3	0.09	0.018	0.059

if t_U is chosen so that

$$P\left(-t_U - \lambda_{\min} \leq W_1\left(\frac{x}{h_0}\right) \leq t_U - \lambda_{\max}, \quad \forall\, x \in S \right) = 1 - \alpha_0.$$

The quantities $\lambda_{\max}$ and $\lambda_{\min}$ are unknown in applications. A feasible confidence band can be obtained by replacing them with the bootstrap estimates $\lambda_{\max} = \max_{x\in S}\hat{\lambda}(x)$ and $\lambda_{\min} = \min_{x\in S}\hat{\lambda}(x)$. Similarly, the unknown quantities $\sigma_{\hat{g}}(x)$ and h_0 can be replaced with $\hat{\sigma}_{\hat{g}}(x)$ and h, respectively. The critical value t_U is replaced by $\hat{t}_U$, which is the solution in t to

$$P\left(-t - \hat{\lambda}_{\min} \leq W_1\left(\frac{x}{h}\right) \leq t - \hat{\lambda}_{\max}, \quad \forall\, x \in S \right) = 1 - \alpha_0. \tag{3.7}$$

Table 2. Simulation results for $\tau = 0.50$.

Method	n	j	Prop. with Cov. Prob. ≥ 0.95	Av. Abs. Error of Cov. Prob.	Av. Width
Bootstrap method of Sec. 2.1	100	1	0.73	0.034	1.64
		2	0.76	0.022	1.48
		3	0.97	0.024	1.17
	500	1	0.89	0.048	0.95
		2	1.0	0.035	0.79
		3	0.95	0.027	0.56
	1,000	1	0.92	0.056	0.74
		2	0.97	0.035	0.58
		3	0.94	0.029	0.41
Undersmooth	100	1	0	0.075	1.20
		2	0	0.053	1.19
		3	0.06	0.026	1.17
	500	1	0.27	0.021	0.78
		2	0.41	0.016	0.62
		3	0.51	0.094	0.57
	1,000	1	0.49	0.025	0.56
		2	0.51	0.011	0.51
		3	0.63	0.099	0.43
Bias Corr.	100	1	0	0.14	1.38
		2	0	0.12	1.35
		3	0	0.092	1.26
	500	1	0	0.069	1.01
		2	0.11	0.028	0.86
		3	0.17	0.019	0.69
	1,000	1	0.05	0.038	0.80
		2	0.11	0.025	0.63
		3	0.20	0.013	0.51

where $\hat{\lambda}_{\max}$ and $\hat{\lambda}_{\min}$ are treated as non-stochastic constants, not random variables, when calculating the probability on the left-hand side of (3.7). The resulting uniform confidence band is

$$-\hat{t}_U \leq \frac{(nh)^{1/2}\{\hat{g}(x) - g(x)\}}{\hat{\sigma}_{\hat{g}}(x)} \leq \hat{t}_U, \qquad \forall x \in S.$$

The following theorem establishes the asymptotic coverage probability of this interval.

Theorem 4. *If Assumptions* 1–6 *hold, then*

$$\liminf_{n\to\infty} P\left(-\hat{t}_U \leq \frac{(nh)^{1/2}\{\hat{g}(x) - g(x)\}}{\hat{\sigma}_{\hat{g}}(x)} \leq \hat{t}_U, \quad \forall x \in S \right) \geq 1 - \alpha_0.$$

Table 3. Simulation results for $\tau = 0.75$.

Method	n	j	Prop. with Cov. Prob. ≥ 0.95	Av. Abs. Error of Cov. Prob.	Av. Width
Bootstrap method of Sec. 2.1	100	1	0.76	0.032	1.74
		2	0.89	0.023	1.58
		3	0.94	0.020	1.20
	500	1	0.86	0.045	0.98
		2	1.0	0.033	0.78
		3	0.91	0.022	0.57
	1,000	1	0.86	0.046	0.76
		2	0.97	0.035	0.62
		3	0.94	0.026	0.42
Undersmooth	100	1	0	0.092	1.20
		2	0	0.071	1.10
		3	0	0.045	0.90
	500	1	0.27	0.056	0.68
		2	0.27	0.025	0.62
		3	0	0.033	0.45
	1,000	1	0.35	0.055	0.52
		2	0.32	0.024	0.46
		3	0.11	0.028	0.34
Bias Corr.	100	1	0	0.011	1.07
		2	0	0.019	1.37
		3	0	0.015	1.31
	500	1	0	0.068	0.89
		2	0.08	0.05	0.91
		3	0.06	0.036	0.78
	1,000	1	0	0.068	0.89
		2	0.08	0.048	0.81
		3	0.09	0.018	0.59

The covariance function of W_1 can be estimated consistently. See equation (A.13) in the supplementary appendix. The probability on the left-hand side of (3.7) can be computed by simulation with arbitrary accuracy by replacing the covariance function of W_1 with its consistent estimate and the continuum S with a grid of equally spaced points.

4. Numerical Experiments

This section reports the results of a set of Monte Carlo experiments that illustrate the finite-sample performance of the method described in Section 2.

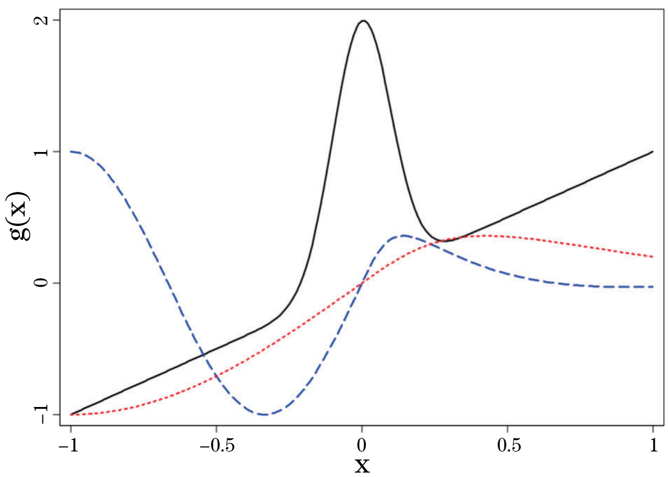

Figure 1. Conditional quantile functions. Solid line is $g_1(x)$. Long dashes are $g_2(x)$. Short dashes are $g_3(x)$.

4.1. Design

Data $\{Y_i, X_i : i = 1, \ldots, n\}$ were generated from the models

$$Y = g_j(X) + \epsilon, \qquad j = 1, 2, 3,$$

where

$$g_1(x) = x + 5\phi(10x),$$
$$g_2(x) = \frac{\sin(1.5\pi x)}{1 + 18x^2\{sgn(x) + 1\}},$$
$$g_3(x) = \frac{\sin(1.5\pi x)}{1 + 2x^2\{sgn(x) + 1\}},$$

ϕ is the standard normal probability density function, and $X \sim U[-1, 1]$. The distribution of ϵ is $N(\mu_\tau, 1)$ for $\tau = 0.25$, $\tau = 0.5$, and $\tau = 0.75$, where $\mu_{0.25} = 0.6745$, $\mu_{0.50} = 0$, and $\mu_{0.75} = -0.6745$. Thus, $P(\epsilon \le 0) = \tau$ for each value of τ. The functions g_j were used in numerical experiments by HH and other authors. Graphs of these functions are shown in Figure 1. The function g_1 has a sharp peak and is the most challenging for our method, g_2 is less challenging than g_1, and g_3 is the smoothest and least challenging.

The sample sizes in the experiments were $n = 100$, 500, and 1,000. The kernel function was

$$K(v) = 0.75(1 - v^2)I(|v| \le 1).$$

The bandwidth h for local linear estimation of the g_js was chosen using the

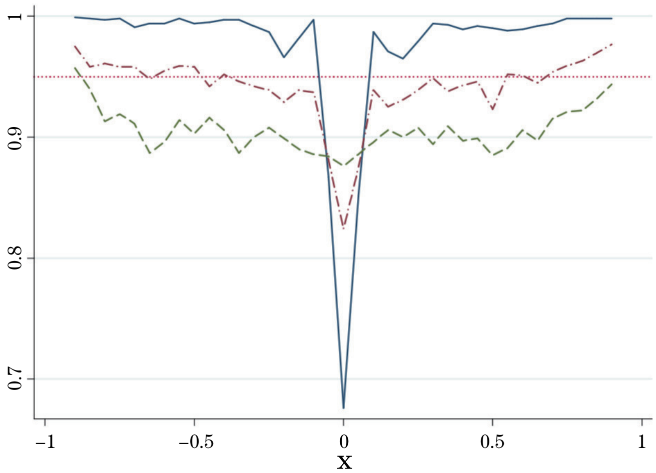

Figure 2a. Coverage probabilities of nominal 95% pointwise confidence bands with $g(x) = x + 5\phi(10x)$ and $n = 1{,}000$. Solid line: proposed bootstrap method. Dashes: explicit bias correction method. Dash-dots: undersmoothing method. Dots: 95% line.

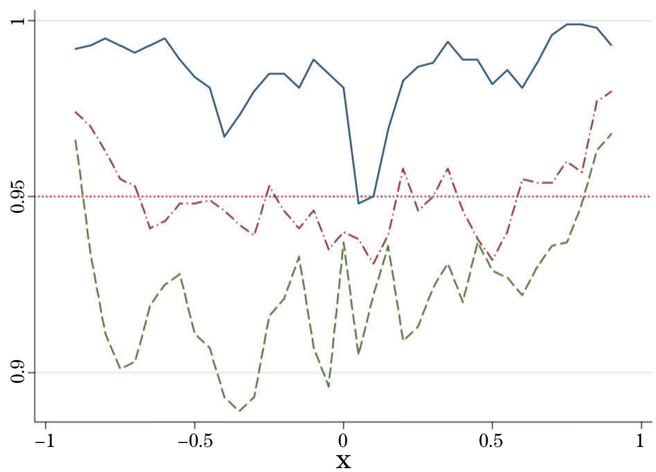

Figure 2b. Coverage probabilities of nominal 95% pointwise confidence bands with $g(x) = \sin(1.5\pi x)/[1 + 18x^2\{sgn(x) + 1\}]$ and $n = 1{,}000$. Solid line: proposed bootstrap method. Dashes: explicit bias correction method. Dash-dots: undersmoothing method. Dots: 95% line.

plug-in method of Yu and Jones (1998). Bandwidths for estimating $f_X(x)$ and $f_\epsilon(0)$ were chosen by Silverman's rule of thumb. To avoid boundary effects, the set S was chosen so that its boundaries were at least one bandwidth from the boundaries of $[-1, 1]$. This resulted in $S = [-0.9, 0.9]$ for experiments with g_1 and g_2 and $S = [-0.85, 0.85]$ for experiments with g_3. S is narrower for the

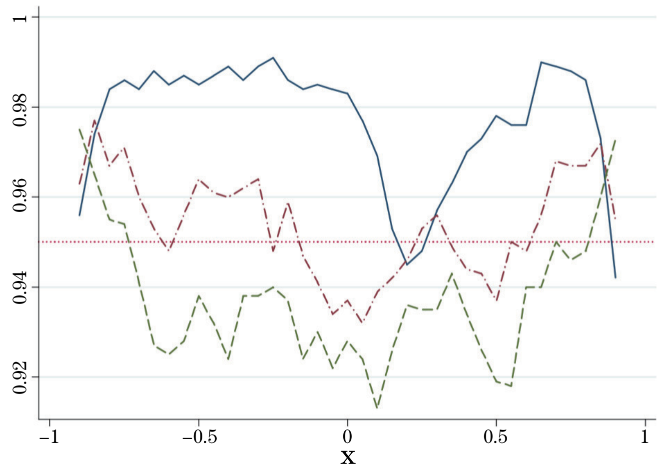

Figure 2c. Coverage probabilities of nominal 95% pointwise confidence bands with $g(x) = \sin(1.5\pi x)/[1+2x^2\{sgn(x)+1\}]$ and $n = 1{,}000$. Solid line: proposed bootstrap method. Dashes: explicit bias correction method. Dash-dots: undersmoothing method. Dots: 95% line.

experiments with g_3 because that function is smoother than g_1 and g_2 and has a larger bandwidth. We set $\alpha_0 = 0.05$ and $1 - \xi = 0.95$. Pointwise confidence bands were computed using an equally spaced grid of points $x \in S$ with a spacing of 0.05. The grid spacing was 0.02 for uniform confidence bands. The proportion of points x at which the coverage probability is at least 0.95 was estimated by the proportion of grid points at which the coverage probability equals or exceeds this value. There were 1,000 Monte Carlo replications in each experiment.

We also computed pointwise confidence bands using undersmoothing and the explicit bias correction method of Schucany and Sommers (1977). There are no satisfactory empirical methods for choosing an undersmoothing bandwidth or the auxiliary bandwidth required for explicit bias correction. For undersmoothing, we set the bandwidth equal to $\gamma_1 h$, where h is the bandwidth selected by the method of Yu and Jones (1998) and $\gamma_1 \leq 1$ is a constant. For explicit bias correction, we set the auxiliary bandwidth equal to $\hat{d}n^{-\gamma_2/5}$, where $\gamma_2 \leq 1$ is a constant and $\hat{d}$ is as in Assumption 6. The values of γ_1 and γ_2 were chosen to achieve coverage probabilities of at least 0.95 for as large a proportion of values of x in the grid as possible. This approach cannot be used in applications and gives an advantage to undersmoothing and explicit bias correction. Nonetheless, it will be seen in Section 4.2 that the performance of these methods is poor compared to that of the method of Section 2.1.

Table 4. The bootstrap method's coverage probabilities for g_1 when the interval $[-0.05, 0.05]$ is removed from S.

n	τ	Prop. with Cov. Prob. ≥ 0.95	Av. Abs. Error of Cov. Prob.	Av. Width
100	0.25	0.76	0.020	1.71
	0.50	0.79	0.022	1.62
	0.75	0.85	0.021	1.75
500	0.25	0.91	0.033	0.92
	0.50	1.0	0.039	0.88
	0.75	0.94	0.035	0.94
1,000	0.25	0.94	0.034	0.70
	0.50	1.0	0.041	0.67
	0.75	0.94	0.035	0.71

4.2. Results of the experiments

Tables 1-3 show properties of pointwise confidence bands for $\tau = 0.25$, 0.50 and 0.75, respectively. At all quantiles and sample sizes, the bootstrap method described in Section 2.1 has much higher proportions of values of x for which the probability of covering of $g(x)$ exceeds 0.95 than do the undersmoothing and explicit bias correction methods. When $n = 100$, the bootstrap method's proportions exceed 0.70 for $j = 1$ and 2, and 0.95 for $j = 3$. When $n = 1{,}000$, the bootstrap method's proportions exceed 0.92 for all values of j. By contrast, the proportion of values of x for which undersmoothing achieves a coverage probability of at least 0.95 is below 0.65 for all values of n and j. The absolute error in the coverage probability (column 5 of Tables 1-3) is the absolute value of the difference between the actual coverage probability and the nominal probability of 0.95. Thus, the absolute error increases when the actual coverage probability exceeds 0.95, as well as when the actual coverage probability is less than 0.95.

The proportion of values of x for which explicit bias correction achieves a coverage probability of at least 0.95 is below 0.20 for all values of n and j. Undersmoothing and explicit bias correction perform poorly despite choosing the bandwidth for undersmoothing and the auxiliary bandwidth for explicit bias correction to achieve optimal performance of these methods.

Although confidence intervals based on undersmoothing and explicit bias correction rarely achieve the nominal coverage probability of 0.95, intervals based on these methods often have coverage probabilities of at least 0.93 if $n = 1{,}000$ and γ_1 and γ_2 are chosen to maximize the proportions of x values at which the coverage probability equals or exceeds this value. For example, with $g_1(x)$, $\tau =$

Table 5. Coverage probabilities of uniform confidence bands obtained by the bootstrap method.

τ	n	j	Cov. Prob.
0.25	100	1	0.84
		2	0.94
		3	0.95
	500	1	0.92
		2	0.99
		3	0.97
	1,000	1	0.94
		2	1
		3	0.98
0.50	100	1	0.85
		2	0.95
		3	0.93
	500	1	0.95
		2	1
		3	0.97
	1,000	1	0.95
		2	1
		3	0.98
0.75	100	1	0.86
		2	0.96
		3	0.96
	500	1	0.94
		2	1
		3	0.97
	1,000	1	0.95
		2	1
		3	0.97

0.50, and undersmoothing, the proportion of x values with coverage probabilities of at least 0.93 is 0.97. With explicit bias correction, the proportion of x values with coverage probabilities of at least 0.93 is 0.65. Figure 2 shows the coverage probabilities obtained by the three methods as functions of x with $n = 1{,}000$.

The relatively low proportions of points at which the coverage probability of the bootstrap method equals or exceeds 0.95 for g_1 are due to the sharp peak of this function in the vicinity of $x = 0$, which causes the bias of $\hat{g}_1$ to be especially large. HH provide a theoretical explanation for why the bootstrap method performs poorly in regions of unusually high bias. The phenomenon is illustrated in Table 4, which shows the proportion of points for which the bootstrap confidence band covers $g_1(x)$ with probability exceeding 0.95 when the interval $[-0.05, 0.05]$ containing the peak is excluded from S. The proportion of points for which the coverage probability equals or exceeds 0.95 is at least 0.94 when $n = 500$ or $n = 1{,}000$ except for $n = 500$ and $\tau = 0.25$, when the proportion is 0.91. The function g_2 has peaks and troughs at $x = 0.15$ and $x = -0.35$, but these are not as sharp as the peak of g_1. Consequently, they have little effect on the coverage probabilities for g_2 when $n \geq 500$.

Table 5 shows the coverage probabilities of uniform confidence bands obtained with the bootstrap method. The coverage probabilities for g_2 and g_3 at all quantiles equal or exceed 0.95 if $n \geq 500$ and 0.93 if $n = 100$. The coverage probabilities for g_1 with $n \geq 500$ equal or exceed 0.94 except for $n = 500$ and $\tau = 0.25$, when the coverage probability is 0.92.

5. Conclusions

This paper has described a bootstrap method for constructing pointwise and uniform confidence bands for a conditional quantile function that is estimated nonparametrically. The method is based on local polynomial estimation and uses only a bandwidth that can be selected using standard methods such as cross validation or plug-in. In contrast to other methods for constructing confidence bands, the bootstrap method does not require bandwidths that under- or oversmooth the nonparametric function estimator. This is an important advantage of the bootstrap method, because there are no satisfactory empirical methods for selecting bandwidths that under- or oversmooth a nonparametric estimator. The bootstrap method presented here is an extension of the method of Hall and Horowitz (2013) for conditional mean functions to conditional quantile functions and uniform confidence bands. The results of Monte Carlo experiments have illustrated the good finite-sample performance of the bootstrap method and the poor performance of methods based on under- or oversmoothing.

Supplementary Materials

The online supplementary materials provide the proofs of Theorems 3.1 and 3.2, Corollary 3.3, and Theorem 3.4.

References

Aguirre, V. M. and Dominguez M. A. (2013). New inference methods for quantile regression based on resampling. *The Econometrics Journal* **16**, 278–283.

Bjerve, S., Doksum, K. A. and Yandell, B. S. (1985). Uniform confidence bounds for regression based on a simple moving average. *Scandinavian Journal of Statistics* **12**, 159–169.

Calonico, S., Cattaneo, M. D. and Farrell, M. H. (2016). On the effect of bias estimation on coverage accuracy in nonparametric inference. Working paper, Department of Economics, University of Michigan.

Chen, S. X. (1996). Empirical likelihood confidence intervals for nonparametric density estimation. *Biometrika* **83**, 329–341.

Chen, S. X., Härdle, W. and Li, M. (2003). An empirical likelihood goodness-of-fit test for time series. *Journal of the Royal Statistical Society Series B Statistical Methodology* **65**, 663–678.

Claeskens, G. and Van Keilegom, I. (2003). Bootstrap confidence bands for regression curves and their derivatives. *The Annals of Statistics* **31**, 1852–1884.

De Angelis, D., Hall, P. and Young, G. A. (1993). Analytical and bootstrap approximations to estimator distributions in regression. *Journal of the American Statistical Association* **88**, 1310–1316.

Eubank, R. L. and Speckman, P. L. (1993). Confidence regions in nonparametric regression. *Journal of the American Statistical Association* **88**, 1287–1301.

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

Feng, X., He, X. and Hu, J. (2011). Wild bootstrap for quantile regression. *Biometrika* **98**, 995–999.

Galvao, A. F. and Montes-Rojas, G. (2015). On bootstrap inference for quantile regression panel data. *Econometrics* **3**, 654–666.

Hagemann, A. (2017). Cluster-robust bootstrap inference in quantile regression models. *Journal of the American Statistical Association* **112**, 446–456.

Hahn, J. (1995). Bootstrapping quantile regression estimators. *Econometric Theory* **11**, 105–121.

Hall, P. (1992). Effect of bias estimation on coverage accuracy of bootstrap confidence intervals for a probability density. *The Annals of Statistics* **20**, 675–694.

Hall, P. and Horowitz, J. L. (2013). A simple bootstrap method for constructing nonparametric confidence bands for functions. *The Annals of Statistics* **41**, 1892–1921.

Hall, P. and Owen, A. (1993). Empirical likelihood confidence bands in density estimation. *Journal of Computational and Graphical Statistics* **2**, 273–289.

Härdle, W. and Bowman, A. (1988). Bootstrapping in nonparametric regression: Local adaptive smoothing and confidence bands. *Journal of the American Statistical Association* **83**, 102–110.

Härdle, W., Huet, S. and Jolivet, E. (1995). Better bootstrap confidence intervals for regression curve estimation. *Statistics* **26**, 287–306.

Härdle, W., Huet, S., Mammen, E. and Sperlich, S. (2004). Bootstrap inference in semiparametric generalized additive models. *Econometric Theory* **20**, 265–300.

Härdle, W. and Marron, J. (1991). Bootstrap simultaneous error bars for nonparametric regression. *The Annals of Statistics* **19**, 778–796.

Horowitz, J. L. (1998). Bootstrap methods for median regression models. *Econometrica* **66**, 1327–1351.

McMurry, T. L. and Politis, D. N. (2008). Bootstrap confidence intervals in nonparametric regression with built-in bias correction. *Statistics and Probability Letters* **78**, 2463–2469.

Neumann, M. H. (1995). Automatic bandwidth choice and confidence intervals in nonparametric regression. *The Annals of Statistics* **23**, 1937–1959.

Neumann, M. H. and Polzehl, J. (1998). Simultaneous bootstrap confidence bands in nonparametric regression. *Journal of Nonparametric Statistics* **9**, 307–333.

Picard, D. and Tribouley, K. (2000). Adaptive confidence interval for pointwise curve estimation. *The Annals of Statistics* **28**, 298–335.

Schucany, W. R. and Sommers, J. P. (1977). Improvement of kernel type density estimators. *Journal of the American Statistical Association* **72**, 420–423.

Sun, J. and Loader, C. (1994). Simultaneous confidence bands for linear regression and smoothing. *The Annals of Statistics* **22**, 1328–1345.

Xia, Y. (1998). Bias-corrected confidence bands in nonparametric regression. *Journal of the Royal Statistical Society Series B Statistical Methodology* **60**, 797–811.

Yu, K. and Jones, M. (1998). Local linear quantile regression. *Journal of the American Statistical Association* **93**, 228–237.

Department of Economics, Northwestern University, Evanston, IL 60201, USA.

E-mail: joel-horowitz@northwestern.edu

Cornerstone Research, 599 Lexington Ave. New York, NY 10022, USA.

E-mail: akrishnamurthy@cornerstone.com

(Received January 2017; accepted January 2018)