# ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics*

http://www.econometricsociety.org/

*Econometrica*, Vol. 78, No. 3 (May, 2010), 1093–1125


## QUANTILE AND PROBABILITY CURVES WITHOUT CROSSING


VICTOR CHERNOZHUKOV
*Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.*

IVÁN FERNÁNDEZ-VAL
*Boston University, Boston, MA 02215, U.S.A.*

ALFRED GALICHON
*Ecole Polytechnique, 91128 Palaiseau, France*

*Econometrica*, Vol. 78, No. 3 (May, 2010), 1093–1125

# QUANTILE AND PROBABILITY CURVES WITHOUT CROSSING

BY VICTOR CHERNOZHUKOV, IVÁN FERNÁNDEZ-VAL,
AND ALFRED GALICHON[1]

This paper proposes a method to address the longstanding problem of lack of monotonicity in estimation of conditional and structural quantile functions, also known as the quantile crossing problem (Bassett and Koenker (1982)). The method consists in sorting or monotone rearranging the original estimated non-monotone curve into a monotone rearranged curve. We show that the rearranged curve is closer to the true quantile curve than the original curve in finite samples, establish a functional delta method for rearrangement-related operators, and derive functional limit theory for the entire rearranged curve and its functionals. We also establish validity of the bootstrap for estimating the limit law of the entire rearranged curve and its functionals. Our limit results are generic in that they apply to every estimator of a monotone function, provided that the estimator satisfies a functional central limit theorem and the function satisfies some smoothness conditions. Consequently, our results apply to estimation of other econometric functions with monotonicity restrictions, such as demand, production, distribution, and structural distribution functions. We illustrate the results with an application to estimation of structural distribution and quantile functions using data on Vietnam veteran status and earnings.

KEYWORDS: Conditional quantiles, structural quantiles, monotonicity problem, rearrangement, isotonic regression, functional delta method.

## 1. INTRODUCTION

THIS PAPER ADDRESSES the longstanding problem of lack of monotonicity in the estimation of conditional and structural quantile functions, also known as the quantile crossing problem (Bassett and Koenker (1982) and He (1997)). The most common approach to estimating quantile curves is to fit a curve, of-

[1]Previous, more extended, versions of this paper (September 2006, April 2007) are available at web.mit.edu/~vchern/www/ and www.ArXiv.org. The method developed in this paper has now been incorporated in the package `quantreg` (Koenker (2007)) in R. The title of this paper is (partially) borrowed from the work of Xuming He (1997), to whom we are grateful for the inspiration. We would like to thank the editor Oliver Linton, three anonymous referees, Alberto Abadie, Josh Angrist, Gilbert Bassett, Andrew Chesher, Phil Cross, James Durbin, Ivar Ekeland, Brigham Frandsen, Raymond Guiteras, Xuming He, Roger Koenker, Joonhwan Lee, Vadim Marmer, Ilya Molchanov, Francesca Molinari, Whitney Newey, Steve Portnoy, Shinichi Sakata, Art Shneyerov, Alp Simsek, and participants at BU, CEMFI, CEMMAP Measurement Matters Conference, Columbia Conference on Optimal Transportation, Columbia, Cornell, Cowles Foundation 75th Anniversary Conference, Duke-Triangle, Ecole Polytechnique, Frontiers of Microeconometrics in Tokyo, Georgetown, Harvard-MIT, MIT, Northwestern, UBC, UCL, UIUC, University of Alicante, and University of Gothenburg Conference "Nonsmooth Inference, Analysis, and Dependence," for comments that helped us to considerably improve the paper. We are grateful to Alberto Abadie for providing us the data for the empirical example. The authors gratefully acknowledge research support from the National Science Foundation and Chaire X-Dauphine "Finance et Développement Durable."

 DOI: 10.3982/ECTA7880

ten linear, pointwise for each probability index.[2] Researchers use this approach for a number of reasons, including parsimony of the resulting approximations and excellent computational properties. The resulting fits, however, may not respect a logical monotonicity requirement—that the quantile curve should be increasing as a function of the probability index.

This paper introduces a natural monotonization of the empirical curves by sampling from the estimated nonmonotone model, and then taking the resulting conditional quantile curves which by construction are monotone in the probability index. This construction of the monotone curve may be seen as a bootstrap and as a sorting or monotone rearrangement of the original nonmonotone curve (see Hardy, Littlewood, and Polya (1952), and references given below). We show that the rearranged curve is closer to the true quantile curve in finite samples than the original curve is, and derive functional limit distribution theory for the rearranged curve to perform simultaneous inference on the entire quantile function. Our theory applies to both dependent and independent data, and to a wide variety of original estimators, with only the requirement that they satisfy a functional central limit theorem. Our results also apply to many other econometric problems with monotonicity restrictions, such as distribution and structural distribution functions, as well as demand and production functions, option pricing functions, and yield curves.[3] As an example, we provide an empirical application to estimation of structural distribution and quantile functions based on Abadie (2002) and Chernozhukov and Hansen (2005, 2006).

There exist other methods to obtain monotonic fits for conditional quantile functions. He (1997), for example, proposed to impose a location–scale regression model, which naturally satisfies monotonicity. This approach is fruitful for location–scale situations, but in numerous cases the data do not satisfy the location–scale paradigm, as discussed in Lehmann (1974), Doksum (1974), and Koenker (2005). Koenker and Ng (2005) developed a computational method for quantile regression that imposes the noncrossing constraints in simultaneous fitting of a finite number of quantile curves. The statistical properties of this method have yet to be studied, and the method does not immediately apply to other quantile estimation methods. Mammen (1991) proposed two-step estimators, with mean estimation in the first step followed by isotonization in

[2]This includes all principal approaches to estimation of conditional quantile functions, such as the canonical quantile regression of Koenker and Bassett (1978) and censored quantile regression of Powell (1986). This also includes principal approaches to estimation of structural quantile functions, such as the instrumental quantile regression methods via control functions of Imbens and Newey (2009), Blundell and Powell (2003), Chesher (2003), and Ma and Koenker (2006), and instrumental quantile regression estimators of Chernozhukov and Hansen (2005, 2006).

[3]See Matzkin (1994) for more examples and additional references, and Chernozhukov, Fernandez-Val, and Galichon (2009) for further theoretical results that cover the latter set of applications.

the second.[4] Similarly to Mammen (1991), we can employ quantile estimation in the first step followed by isotonization in the second, obtaining an interesting method whose properties have yet to be studied. In contrast, our method uses rearrangement rather than isotonization, and is better suited for quantile applications. The reason is that isotonization is best suited for applications with (near) flat target functions, while rearrangement is best suited for applications with steep target functions, as in typical quantile applications. Indeed, in a numerical example closely matching our empirical application, we find that rearrangement significantly outperforms isotonization. Finally, in an independent and contemporaneous work, Dette and Volgushev (2008) proposed to obtain monotonic quantile curves by applying an integral transform to a local polynomial estimate of the conditional distribution function, and derived pointwise limit theory for this estimator. In contrast, we directly monotonize any generic estimate of a conditional quantile function and then derive generic functional limit theory for the entire monotonized curve.[5]

In addition to resolving the problem of estimating quantile curves that avoid crossing, this paper develops a number of original theoretical results on rearranged estimators. It therefore makes both practical and theoretical contributions to econometrics and statistics. Before discussing these contributions more specifically, it is helpful to review some of the relevant literature and available results.

We begin the review by noting that the idea of rearrangement goes back at least to Chebyshev (see Bronshtein et al. (2003, p. 31)), Hardy, Littlewood, and Polya (1952), and Lorentz (1953), among others. Rearrangements have been extensively used in functional analysis and operations research (Villani (2003) and Carlier and Dana (2005)), but not in econometrics or statistics until recently. Recent research on rearrangements in statistics includes the work of Fougeres (1997), which used rearrangement to produce a monotonic kernel density estimator and derived its uniform rates of convergence; Davydov and Zitikis (2005), which considered tests of monotonicity based on rearranged kernel mean regression; Dette, Neumeyer, and Pilz (2006) and Dette and Scheder (2006), which introduced smoothed rearrangements for kernel mean regressions and derived pointwise limit theory for these estimators; and Chernozhukov, Fernandez-Val, and Galichon (2009), which used univariate and multivariate rearrangements on point and interval estimators of monotone functions based on series and kernel regression estimators. In the context of our problem, rearrangement is also connected to the quantile regression bootstrap of Koenker (1994). In fact, our research grew from the realization that

[4]Isotonization is also known as the pool-adjacent-violators algorithm in statistics and ironing in economics. It amounts to projecting the original estimate on the set of monotone functions.

[5]We refer to Dette and Volgushev (2008) for a more detailed comparison of the two approaches.

we could use this bootstrap for the purpose of monotonizing quantile regressions, and we discovered the link to the classical procedure of rearrangement later, while reading Villani (2003).

The theoretical contributions of this paper are threefold. First, our paper derives functional limit theory for rearranged estimators and functional delta methods for rearrangement operators, both of which are important original results. Second, the paper derives functional limit theory for estimators obtained by rearrangement-related operations, which are also original results. For example, our theory includes as a special case the asymptotics of the conditional distribution function estimator based on quantile regression, whose properties have long remained unknown (Bassett and Koenker (1982)). Moreover, our limit theory applies to functions, encompassing the pointwise results as a special case. An attractive feature of our theoretical results is that they do *not* rely on independence of data, the particular estimation method used, or any parametric assumption. They only require that a functional central limit theorem applies to the original estimator of the curve, and the population curves have some smoothness properties. Our results therefore apply to any quantile model and quantile estimator that satisfy these requirements. Third, our results immediately yield validity of the bootstrap for rearranged estimators, which is an important result for practice.

We organize the rest of the paper as follows. In Section 2 we present some analytical results on rearrangement and then present all the main results; in Section 3 we provide an application and a numerical experiment that closely matches the application; in Section 4 we give some concluding remarks; and in the Appendix we include the proofs of the results. The data and programs used for the examples are available in the on-line supplement (Chernozhukov, Fernandez-Val, and Galichon (2010)).

## 2. REARRANGEMENT: ANALYTICAL AND EMPIRICAL PROPERTIES

In this section, we describe rearrangement, derive some basic analytical properties of the rearranged curves in the population, establish functional differentiability results, and establish functional limit theorems and other estimation properties.

### 2.1. *Rearrangement*

We consider a target function $u \mapsto Q_0(u|x)$ that, for each $x \in \mathcal{X}$, maps $(0, 1)$ to the real line and is increasing in $u$. Suppose that $u \mapsto \widehat{Q}(u|x)$ is a parametric or nonparametric estimator of $Q_0(u|x)$. Throughout the paper, we use conditional and structural quantile estimation as the main application, where $u \mapsto Q_0(u|x)$ is the quantile function of a real response variable $Y$, given a vector of regressors $X = x$. Accordingly, we will usually refer to the functions $u \mapsto Q_0(u|x)$ as quantile functions throughout the paper. In other applications, such as estimation of conditional and structural distribution functions, other names would be appropriate and we need to accommodate different domains, as described in Remark 1 below.

Typical estimation methods fit the quantile function $\widehat{Q}(u|x)$ pointwise in $u \in (0,1)$.[6] A problem that might occur is that the map $u \mapsto \widehat{Q}(u|x)$ may not be increasing in $u$, which violates the logical monotonicity requirement.[7] Another manifestation of this issue, known as the quantile crossing problem, is that the conditional quantile curves $x \mapsto \widehat{Q}(u|x)$ may cross for different values of $u$ (He (1997)). Similar issues also arise in estimation of conditional and structural distribution functions (Hall, Wolff, and Yao (1999) and Abadie (2002)).

We can transform the possibly nonmonotone function $u \mapsto \widehat{Q}(u|x)$ into a monotone function $u \mapsto \widehat{Q}^*(u|x)$ by quantile bootstrap or rearrangement. That is, we consider the random variable $Y_x := \widehat{Q}(U|x)$, where $U \sim \text{Uniform}(\mathcal{U})$ with $\mathcal{U} = (0,1)$, and take its quantile function denoted by $u \mapsto \widehat{Q}^*(u|x)$ instead of the original function $u \mapsto \widehat{Q}(u|x)$. This variable $Y_x$ has a distribution function

$$\widehat{F}(y|x) := \int_0^1 1\{\widehat{Q}(u|x) \leq y\}\,du, \tag{2.1}$$

which is naturally monotone in the level $y$, and a quantile function

$$\widehat{Q}^*(u|x) := \widehat{F}^{-1}(u|x) = \inf\{y : \widehat{F}(y|x) \geq u\}, \tag{2.2}$$

which is naturally monotone in the index $u$. Thus, starting with a possibly nonmonotone original curve $u \mapsto \widehat{Q}(u|x)$, the rearrangement (2.1)–(2.2) produces a monotone quantile curve $u \mapsto \widehat{Q}^*(u|x)$. Of course, the rearranged quantile function $u \mapsto \widehat{Q}^*(u|x)$ coincides with the original function $u \mapsto \widehat{Q}(u|x)$ if the original function is nondecreasing in $u$, but differs from it otherwise.

The mechanism (2.1)–(2.2) and its name have a direct relation to the rearrangement operator from functional analysis (Hardy, Littlewood, and Polya (1952)), since $u \mapsto \widehat{Q}^*(u|x)$ is the monotone rearrangement of $u \mapsto \widehat{Q}(u|x)$. Equivalently, as we stated earlier, rearrangement has a direct relation to the quantile bootstrap (Koenker (1994)), since the rearranged quantile curve is the quantile function of the bootstrap variable produced by the estimated quantile model. Moreover, we refer the reader to Dette, Neumeyer, and Pilz (2006, p. 470), who, using a closely related motivation, introduced the idea of

[6]See Koenker and Bassett (1978), Powell (1986), Chaudhuri (1991), Buchinsky (1994), Chamberlain (1994), Buchinsky and Hahn (1998), Yu and Jones (1998), Abadie, Angrist, and Imbens (2002), Honoré, Khan, and Powell (2002), and Chernozhukov and Hansen (2006), among others, for examples of exogenous, censored, endogenous, nonparametric, and other types of quantile regression estimators.

[7]Throughout the paper, by "monotone" we mean (weakly) increasing.

smoothed rearrangement, which produces smoothed versions of (2.1) and (2.2) and can be valuable in applications. Finally, for practical and computational purposes, it is helpful to think of rearrangement as sorting. Indeed to compute the rearrangement of a continuous function $u \mapsto \widehat{Q}(u|x)$, we simply set $\widehat{Q}^*(u|x)$ as the $u$th quantile of $\{\widehat{Q}(u_1|x), \dots, \widehat{Q}(u_k|x)\}$, where $\{u_1, \dots, u_k\}$ is a sufficiently fine net of equidistant indices in $(0, 1)$.

REMARK 1—Adjusting for Domains Different From the Unit Interval: Throughout the paper we assume that the domain of all the functions is the unit interval, $\mathcal{U} = (0, 1)$, but in many applications we may have to deal with different domains. For example, in quantile estimation problems, we may consider a subinterval $(a, b)$ of the unit interval as the domain, so as to avoid estimation of tail quantiles. In distribution estimation problems, we may consider the entire real line as the domain. In such cases, we can first transform these functions to have the unit interval as the domain. Concretely, suppose we have an original function $\bar{Q} : (a, b) \to \mathbb{R}$. Then using any increasing bijective mapping $\varphi : (a, b) \mapsto (0, 1)$, we can define $Q := \bar{Q} \circ \varphi^{-1} : (0, 1) \to \mathbb{R}$ and then proceed to obtain its rearrangement $Q^*$. In the case where $a \neq -\infty$ and $b \neq \infty$, we can take $\varphi$ to be an affine mapping. To obtain an increasing rearrangement $\bar{Q}^*$ of $\bar{Q}$, we then set $\bar{Q}^* = Q^* \circ \varphi$.

Let $Q$ denote the pointwise probability limit of $\widehat{Q}$, which we will refer to as the population curve. In the analysis we distinguish the following two cases:

CASE 1—Monotonic $Q$: The population curve $u \mapsto Q(u|x)$ is increasing, and thus satisfies the monotonicity requirement.

CASE 2—Nonmonotonic $Q$: The population curve $u \mapsto Q(u|x)$ is not increasing due to misspecification, and thus does not satisfy the monotonicity requirement.

In Case 1 the empirical curve $u \mapsto \widehat{Q}(u|x)$ may be nonmonotone due to estimation error, while in Case 2 it may be nonmonotone due to both misspecification and estimation error. A leading example of Case 1 is when the population curve $Q$ is correctly specified, so that it equals the target quantile curve, namely $Q(u|x) = Q_0(u|x)$ for all $u \in (0, 1)$. Case 1 also allows for some degree of misspecification, provided that the population curve, $Q \neq Q_0$, remains monotone. A leading example of Case 2 is when the population curve $Q$ is misspecified, $Q \neq Q_0$, to a degree that makes $u \mapsto Q(u|x)$ nonmonotone. For example, the common linear specification $u \mapsto Q(u|x) = p(x)^\top \beta(u)$ can be nonmonotone if the support of $X$ is sufficiently rich, while the set of transformations of $x$, $p(x)$, is not (Koenker (2005, Chap. 2.5)). Typically, by using a rich enough set $p(x)$ we can approximate the true function $Q_0(u|x)$ sufficiently well, and thus often avoid Case 2; see Koenker (2005, Chap. 2.5). This is the strategy that we generally recommend, since inference and limit theory under Case 1 is theoretically and practically simpler than under Case 2. However, in what follows we analyze the behavior of rearranged estimators in both Cases 1 and 2, since either of these cases could occur.

In the rest of the section, we establish the empirical properties of the rearranged estimated quantile functions and the corresponding distribution functions

$$u \mapsto \widehat{Q}^*(u|x) \quad \text{and} \quad y \mapsto \widehat{F}(y|x), \tag{2.3}$$

under Cases 1 and 2.

### 2.2. *Basic Analytical Properties of Population Curves*

We start by characterizing certain analytical properties of the probability limits or population versions of empirical curves (2.3), namely

$$\begin{aligned} & y \mapsto F(y|x) = \int_0^1 1\{Q(u|x) \leq y\}\, du, \\ & u \mapsto Q^*(u|x) := F^{-1}(u|x) = \inf\{y : F(y|x) \geq u\}. \end{aligned} \tag{2.4}$$

We need these properties to derive our main limit results stated in the following subsections.

Recall first the following definitions from Milnor (1965). Let $g : \mathcal{U} \subset \mathbb{R} \mapsto \mathbb{R}$ be a continuously differentiable function. A point $u \in \mathcal{U}$ is called a regular point of $g$ if the derivative of $g$ at this point does not vanish, that is, $\partial_u g(u) \neq 0$, where $\partial_u$ denotes the partial derivative operator with respect to $u$. A point $u$ which is not a regular point is called a critical point. A value $y \in g(\mathcal{U})$ is called a regular value of $g$ if $g^{-1}(y)$ contains only regular points, that is, if $\forall u \in g^{-1}(y)$, $\partial_u g(u) \neq 0$. A value $y$ which is not a regular value is called a critical value.

Define region $\mathcal{Y}_x$ as the support of $Y_x$, and define regions $\mathcal{YX} := \{(y, x) : y \in \mathcal{Y}_x, x \in \mathcal{X}\}$ and $\mathcal{UX} := \mathcal{U} \times \mathcal{X}$. We assume throughout that $\mathcal{Y}_x \subset \mathcal{Y}$, a compact subset of $\mathbb{R}$, and that $x \in \mathcal{X}$, a compact subset of $\mathbb{R}^d$. In some applications the curves of interest are not functions of $x$ or we might be interested in a particular value $x$. In this case, we can take the set $\mathcal{X}$ to be a singleton $\mathcal{X} = \{x\}$.

ASSUMPTION 1—Properties of $Q$: *We maintain the following assumptions on $Q$ throughout the paper*:

(a) $Q : \mathcal{U} \times \mathcal{X} \mapsto \mathbb{R}$ *is a continuously differentiable function in both arguments*.

(b) *The number of elements of* $\{u \in \mathcal{U} : \partial_u Q(u|x) = 0\}$ *is uniformly bounded on* $x \in \mathcal{X}$.

Assumption 1(b) implies that, for each $x \in \mathcal{X}$, $\partial_u Q(u|x)$ is not zero almost everywhere on $\mathcal{U}$ and can switch sign only a bounded number of times. Further, we define $\mathcal{Y}_x^*$ to be the subset of regular values of $u \mapsto Q(u|x)$ in $\mathcal{Y}_x$, and $\mathcal{YX}^* := \{(y, x) : y \in \mathcal{Y}_x^*, x \in \mathcal{X}\}$.

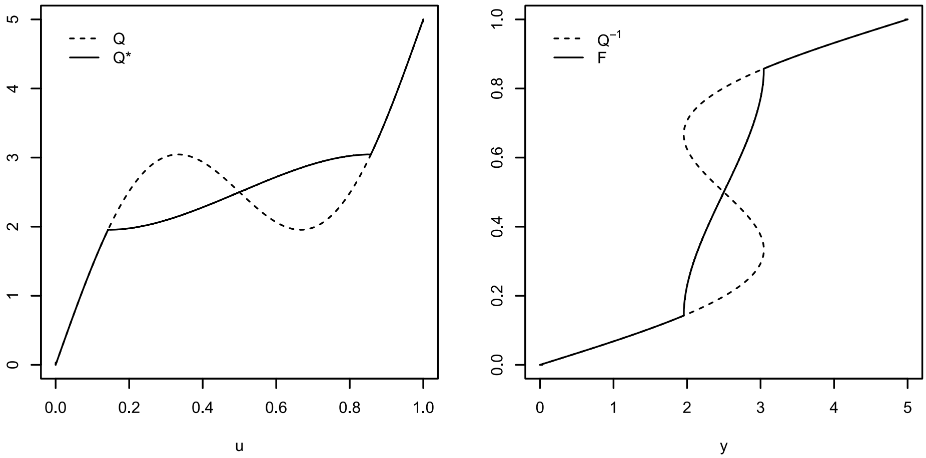


FIGURE 1.—Left: The pseudo-quantile function $Q$ and the rearranged quantile function $Q^*$. Right: The pseudo-distribution function $Q^{-1}$ and the distribution function $F$ induced by $Q$.

We use a simple example to describe some basic analytical properties of (2.4), which we state more formally in the proposition given below. Consider the pseudo-quantile function $Q(u) = 5\{u + \sin(2\pi u)/\pi\}$, which is highly nonmonotone in $(0, 1)$ and therefore fails to be a proper quantile function. The left panel of Figure 1 shows $Q$ together with its monotone rearrangement $Q^*$. We see that $Q^*$ partially coincides with $Q$ on the areas where $Q$ behaves like a proper quantile function, and that $Q^*$ is continuous and increasing. Note also that 1/3 and 2/3 are the critical points of $Q$, and 3.04 and 1.96 are the corresponding critical values. The right panel of Figure 1 shows the pseudo-distribution function $Q^{-1}$, which is multivalued, and the distribution function $F = Q^{*-1}$ induced by sampling from $Q$. We see that $F$ is continuous and does not have point masses. The left panel of Figure 2 shows $\partial_u Q^*$, the sparsity function for $Q^*$. We see that the sparsity function is continuous at the $Q^{*-1}$ image of the regular values of $Q$ and has jumps at the $Q^{*-1}$ image of the critical values of $Q$. The right panel of Figure 2 shows $\partial_y F$, the density function for $F$. We see that $\partial_y F$ is continuous at the regular values of $Q$ and has jumps at the critical values of $Q$.

The following proposition states more formally the properties of $Q^*$ and $F$:

PROPOSITION 1—Basic Properties of $F$ and $Q^*$: *The functions* $y \mapsto F(y|x)$ *and* $u \mapsto Q^*(u|x)$ *satisfy the following properties, for each* $x \in \mathcal{X}$:

(i) *The set of critical values*, $\mathcal{Y}_x \setminus \mathcal{Y}_x^*$, *is finite, and* $\int_{\mathcal{Y}_x \setminus \mathcal{Y}_x^*} dF(y|x) = 0$.

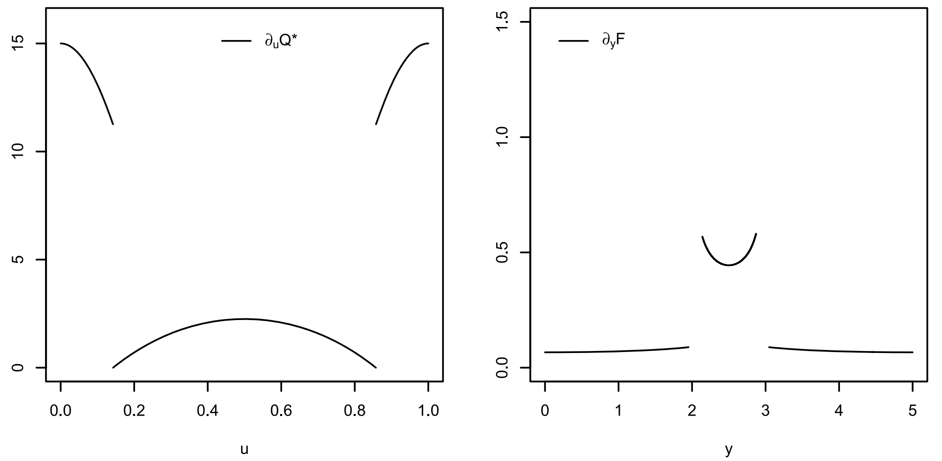


FIGURE 2.—Left: The density (sparsity) function of the rearranged quantile function $Q^*$. Right: The density function of the distribution function $F$ induced by $Q$.

(ii) *For any* $y \in \mathcal{Y}_x^*$,

$$F(y|x) = \sum_{k=1}^{K(y|x)} \operatorname{sign}\{\partial_u Q(u_k(y|x)|x)\} u_k(y|x) + 1\{\partial_u Q(u_{K(y|x)}(y|x)|x) < 0\},$$

*where* $\{u_k(y|x),$ *for* $k = 1, 2, \ldots, K(y|x) < \infty\}$ *are the roots of* $Q(u|x) = y$ *in increasing order*.

(iii) *For any* $y \in \mathcal{Y}_x^*$, *the ordinary derivative* $f(y|x) := \partial_y F(y|x)$ *exists and takes the form*

$$f(y|x) = \sum_{k=1}^{K(y|x)} \frac{1}{|\partial_u Q(u_k(y|x)|x)|},$$

*which is continuous at each* $y \in \mathcal{Y}_x^*$. *For any* $y \in \mathcal{Y} \setminus \mathcal{Y}_x^*$, *we set* $f(y|x) := 0$. $F(y|x)$ *is absolutely continuous and strictly increasing in* $y \in \mathcal{Y}_x$. *Moreover*, $y \mapsto f(y|x)$ *is a Radon–Nikodym derivative of* $y \mapsto F(y|x)$ *with respect to the Lebesgue measure*.

(iv) *The quantile function* $u \mapsto Q^*(u|x)$ *partially coincides with* $u \mapsto Q(u|x)$; *namely* $Q^*(u|x) = Q(u|x)$, *provided that* $u \mapsto Q(u|x)$ *is increasing at* $u$, *and the preimage of* $Q^*(u|x)$ *under* $Q$ *is unique*.

(v) *The quantile function* $u \mapsto Q^*(u|x)$ *is equivariant to monotone transformations of* $u \mapsto Q(u|x)$, *in particular, to location and scale transformations*.

(vi) *The quantile function* $u \mapsto Q^*(u|x)$ *has an ordinary continuous derivative* $\partial_u Q^*(u|x) = 1/f(Q^*(u|x)|x)$, *when* $Q^*(u|x) \in \mathcal{Y}_x^*$. *This function is also a Radon–Nikodym derivative with respect to the Lebesgue measure.*

(vii) *The map* $(y, x) \mapsto F(y|x)$ *is continuous on* $\mathcal{YX}$ *and the map* $(u, x) \mapsto Q^*(u|x)$ *is continuous on* $\mathcal{UX}$.

### 2.3. *Functional Derivatives for Rearrangement-Related Operators*

Here we derive functional derivatives for the rearrangement operator $Q \mapsto Q^*$ and the pre-rearrangement operator $Q \mapsto F$ defined by equation (2.4). These results constitute the first set of original main theoretical results obtained in this paper. In the subsequent sections, these results allow us to establish a generic functional central limit theorem for the estimated functions $\widehat{Q}^*$ and $\widehat{F}$, as well as to establish validity of the bootstrap for estimating their limit laws.

To describe the results, let $\ell^\infty(\mathcal{UX})$ denote the set of bounded and measurable functions $h : \mathcal{UX} \mapsto \mathbb{R}$, let $C(\mathcal{UX})$ denote the set of continuous functions $h : \mathcal{UX} \mapsto \mathbb{R}$, and let $\ell^1(\mathcal{UX})$ denote the set of measurable functions $h : \mathcal{UX} \mapsto \mathbb{R}$ such that $\int_{\mathcal{U}} \int_{\mathcal{X}} |h(u|x)|\, du\, dx < \infty$, where $du$ and $dx$ denote the integration with respect to the Lebesgue measure on $\mathcal{U}$ and $\mathcal{X}$, respectively.

PROPOSITION 2—Hadamard Derivatives of $F$ and $Q^*$ With Respect to $Q$:

(i) *Define* $F(y|x, h_t) := \int_0^1 1\{Q(u|x) + t h_t(u|x) \leq y\}\, du$. *As* $t \to 0$,

$$D_{h_t}(y|x, t) := \frac{F(y|x, h_t) - F(y|x)}{t} \to D_h(y|x), \tag{2.5}$$

$$D_h(y|x) := -\sum_{k=1}^{K(y|x)} \frac{h(u_k(y|x)|x)}{|\partial_u Q(u_k(y|x)|x)|}. \tag{2.6}$$

*The convergence holds uniformly in any compact subset of* $\mathcal{YX}^* := \{(y, x) : y \in \mathcal{Y}_x^*, x \in \mathcal{X}\}$ *for every* $\|h_t - h\|_\infty \to 0$, *where* $h_t \in \ell^\infty(\mathcal{UX})$ *and* $h \in C(\mathcal{UX})$.

(ii) *Define* $Q^*(u|x, h_t) := F^{-1}(y|x, h_t) = \inf\{y : F(y|x, h_t) \geq u\}$. *As* $t \to 0$,

$$\widetilde{D}_{h_t}(u|x, t) := \frac{Q^*(u|x, h_t) - Q^*(u|x)}{t} \to \widetilde{D}_h(u|x), \tag{2.7}$$

$$\widetilde{D}_h(u|x) := -\frac{1}{f(Q^*(u|x)|x)} D_h(Q^*(u|x)|x). \tag{2.8}$$

*The convergence holds uniformly in any compact subset of* $\mathcal{UX}^* = \{(u, x) : (Q^*(u|x), x) \in \mathcal{YX}^*\}$ *for every* $\|h_t - h\|_\infty \to 0$, *where* $h_t \in \ell^\infty(\mathcal{UX})$ *and* $h \in C(\mathcal{UX})$.

This proposition establishes the Hadamard (compact) differentiability of the rearrangement operator $Q \mapsto Q^*$ and the pre-rearrangement operator $Q \mapsto F$ with respect to $Q$, tangentially to the subspace of continuous functions. Note that the convergence holds uniformly on regions that exclude the critical values of the mapping $u \mapsto Q(u|x)$. These results are new and could be of independent interest. Rearrangement operators include inverse (quantile) operators as a special case. In this sense, our results generalize the previous results of Gill and Johansen (1990), Doss and Gill (1992), and Dudley and Norvaisa (1999) on functional delta method (Hadamard differentiability) for the quantile operator. There are two main difficulties in establishing the Hadamard differentiability in our case: first, as in the quantile case, we allow the perturbations $h_t$ to $Q$ to be discontinuous functions, though converging to continuous functions; second, unlike in the quantile case, we allow the perturbed functions $Q + th_t$ to be nonmonotone even when $Q$ is monotone. We need to allow for such rich perturbations in order to match applications where the empirical perturbations $h_t = (\widehat{Q} - Q)/t$, for $t = 1/a_n$ and $a_n$ a growing sequence with the sample size $n$, are discontinuous functions, though converging to continuous functions by the means of a functional central limit theorem; moreover, the empirical (pseudo-) quantile functions $\widehat{Q} = Q + th_t$ are not monotone even when $Q$ is monotone.

The following result deals with the monotonic case. It is worth emphasizing separately, because functional derivatives are particularly simple and we do not have to exclude any nonregular regions from the domains.

COROLLARY 1—Hadamard Derivatives of $F$ and $Q^*$ With Respect to $Q$ in the Monotonic Case: *Suppose $u \mapsto Q(u|x)$ has $\partial_u Q(u|x) > 0$ for each $(u, x) \in \mathcal{UX}$. Then $\mathcal{YX}^* = \mathcal{YX}$ and $\mathcal{UX}^* = \mathcal{UX}$. Therefore, the convergence in Proposition 2 holds uniformly over the entire $\mathcal{YX}$ and $\mathcal{UX}$, respectively. Moreover, $\widetilde{D}_h(u|x) = h$, that is, the Hadamard derivative of the rearranged function with respect to the original function is the identity operator.*

Next we consider the linear functionals obtained by integration,

$$(y', x) \mapsto \int_{\mathcal{Y}} g(y|x, y') F(y|x)\, dy,$$

$$(u', x) \mapsto \int_{\mathcal{U}} g(u|x, u') Q^*(u|x)\, du,$$

with the restrictions on $g$ specified below. These functionals are of interest because they are useful building blocks for various statistics, for example, Lorenz curves with function $g(u|x, u') = 1\{u \leq u'\}$, as discussed in the next section. The following proposition calculates the Hadamard derivative of these functionals.

PROPOSITION 3—Hadamard Derivative of Linear Functionals of $Q^*$ and $F$ With Respect to $Q$: *The following results are true with the limits being continuous on the specified domains*:

(i) *For any measurable $g$ that is bounded uniformly in its arguments and such that $(x, y') \mapsto g(y|x, y')$ is continuous for almost everywhere (a.e.) $y$,*

$$\int_{\mathcal{Y}} g(y|x, y')D_{h_t}(y|x, t)\,dy \to \int_{\mathcal{Y}} g(y|x, y')D_h(y|x)\,dy,$$

*uniformly in $(y', x) \in \mathcal{YX}$.*

(ii) *For any measurable $g$ such that $\sup_{u',x} |g(u|x, u')| \in \ell^1(\mathcal{U})$ and such that $(x, u') \mapsto g(u|x, u')$ is continuous for a.e. $u$,*

$$\int_{\mathcal{U}} g(u|x, u')\widetilde{D}_{h_t}(u|x, t)\,du \to \int_{\mathcal{U}} g(u|x, u')\widetilde{D}_h(u|x)\,du, \tag{2.9}$$

*uniformly in $(u', x) \in \mathcal{UX}$.*

It is important to note that Proposition 3 applies to integrals defined over entire domains, unlike Proposition 2 which states uniform convergence of integrands over domains excluding nonregular neighborhoods. (Thus, Proposition 3 does not immediately follow from Proposition 2.) Here integration acts like a smoothing operation and allows us to ignore these nonregular neighborhoods. To prove convergence of integrals defined over entire domains, we couple the almost everywhere convergence implied by Proposition 2 with the uniform integrability of Lemma 3 in the Appendix, and then interchange limits and integrals. We should also note that an alternative way of proving result (2.9), but *not* other results in the paper, can be based on the convexity of the functional in (2.9) with respect to the underlying curve, following the approach of Mossino and Temam (1981) and Alvino, Lions, and Trombetti (1989). Due to this limitation, we do not pursue this approach in this paper.

It is also worth emphasizing the properties of the following smoothed functionals. For a measurable function $f : \mathbb{R} \mapsto \mathbb{R}$, define the smoothing operator $S$ as

$$Sf(y') := \int k_\delta(y' - y)f(y)\,dy, \tag{2.10}$$

where $k_\delta(v) = 1\{|v| \le \delta\}/2\delta$ and $\delta > 0$ is a fixed bandwidth. Accordingly, the smoothed curves $SF$ and $SQ^*$ are given by

$$SF(y'|x) := \int k_\delta(y' - y)F(y|x)\,dy,$$

$$SQ^*(u'|x) := \int k_\delta(u' - u)Q^*(u|x)\,du.$$

Note that given the quantile function $Q^*$, the smoothed function $SQ^*$ has a convenient interpretation as a local average quantile function or fractile. Since we form these curves as differences of the elementary functionals in Proposition 3 divided by $2\delta$, the following corollary is immediate:

COROLLARY 2—Hadamard Derivative of Smoothed $Q^*$ and $F$ With Respect to $Q$: *We have that* $SD_{h_t}(y'|x,t) \to SD_h(y'|x)$ *uniformly in* $(y',x) \in \mathcal{YX}$, *and* $S\widetilde{D}_{h_t}(u'|x,t) \to S\widetilde{D}_h(u'|x)$ *uniformly in* $(u',x) \in \mathcal{UX}$. *The results hold uniformly in the smoothing parameter* $\delta \in [\delta_1, \delta_2]$, *where* $\delta_1$ *and* $\delta_2$ *are positive constants.*

Note that smoothing allows us to achieve uniform convergence over the entire domain, without excluding nonregular neighborhoods.

### 2.4. *Empirical Properties and Functional Limit Theory for Rearranged Estimators*

Here we state a finite sample result and then derive functional limit laws for rearranged estimators. These results constitute the second set of original main theoretical results obtained in this paper.

The following proposition shows that the rearranged quantile curves have smaller estimation error than the original curves whenever the latter are not monotone.

PROPOSITION 4—Improvement in Estimation Property Provided by Rearrangement: *Suppose that* $\widehat{Q}$ *is an estimator (not necessarily consistent) for some true quantile curve* $Q_0$. *Then, the rearranged curve* $\widehat{Q}^*$ *is closer to the true curve than* $\widehat{Q}$ *in the sense that, for each* $x \in \mathcal{X}$,

$$\|\widehat{Q}^* - Q_0\|_p \le \|\widehat{Q} - Q_0\|_p, \quad p \in [1, \infty],$$

*where* $\|\cdot\|_p$ *denotes the* $L^p$ *norm of a measurable function* $Q : \mathcal{U} \mapsto \mathbb{R}$, *namely* $\|Q\|_p = \{\int_{\mathcal{U}} |Q(u)|^p \, du\}^{1/p}$. *The inequality is strict for* $p \in (1, \infty)$ *whenever* $u \mapsto \widehat{Q}(u|x)$ *is strictly decreasing on a subset of* $\mathcal{U}$ *of positive Lebesgue measure, while* $u \mapsto Q_0(u|x)$ *is strictly increasing on* $\mathcal{U}$. *The above property is independent of the sample size and of the way the estimate of the curve is obtained, and thus continues to hold in the population.*

This property suggests that the rearranged estimators should be preferred over the original estimators. Moreover, this property does not depend on the way the quantile model is estimated or any other specifics, and is thus applicable quite generally. Regarding the proof of this property, the weak reduction in estimation error follows from an application of a classical rearrangement inequality of Lorentz (1953) and the strict reduction follows from its appropriate strengthening (Chernozhukov, Fernandez-Val, and Galichon (2009)).[8]

[8]Similar contractivity properties have been shown for the pool-adjacent-violators algorithm in different contexts. See, for example, Robertson, Wright, and Dykstra (1988) for isotonic re-

To derive the functional limit laws of rearranged estimators, we maintain the following assumptions on $\widehat{Q}$ throughout the paper:

ASSUMPTION 2—Properties of $\widehat{Q}$: *The empirical curve $\widehat{Q}$ takes its values in the space of bounded measurable functions defined on $\mathcal{U}\mathcal{X}$, and in $\ell^\infty(\mathcal{U}\mathcal{X})$,*

$$(2.11) \quad a_n(\widehat{Q}(u|x) - Q(u|x)) \quad \Rightarrow \quad G(u|x)$$

*as a stochastic process indexed by $(u, x) \in \mathcal{U}\mathcal{X}$, where $(u, x) \mapsto G(u|x)$ is a stochastic process (typically Gaussian) with continuous paths. Here $a_n$ is a sequence of constants such that $a_n \to \infty$ as $n \to \infty$, where $n$ is the sample size.*

This assumption requires that the original quantile estimator satisfies a functional central limit theorem with a continuous limit stochastic process over the domain $\mathcal{U} = (0, 1)$ for the index $u$. If (2.11) holds only over a subinterval of $(0, 1)$, we can accommodate the reduced domain following Remark 1. This key condition is rather weak and holds for a wide variety of conditional and structural quantile estimators.[9] With an appropriate normalization rate and a fixed $x$, this assumption holds for series quantile regressions. For example, Belloni and Chernozhukov (2007) extended the results of He and Shao (2000) to the process case and established the functional central limit theorem for $a_n(\widehat{Q}(u|x) - Q(u|x))$ for a fixed $x$. At the same time, we should also point out that this assumption does not need to hold in all estimation problems with monotonicity restrictions.[10]

The following proposition derives functional limit laws for the rearranged quantile estimator $\widehat{Q}^*$ and the corresponding distribution estimator $\widehat{F}$, using the functional differentiation results for the rearrangement-related operators from the previous section.

PROPOSITION 5—Functional Limit Laws for $\widehat{F}$ and $\widehat{Q}^*$: *In $\ell^\infty(K)$, where $K$ is any compact subset of $\mathcal{Y}\mathcal{X}^*$,*

$$(2.12) \quad a_n(\widehat{F}(y|x) - F(y|x)) \quad \Rightarrow \quad D_G(y|x)$$

gression, and Eggermont and LaRiccia (2000) for monotone density estimation. Glad, Hjort, and Ushakov (2003) showed that a density estimator corrected to be a proper density satisfies a similar property.

[9]For sufficient conditions, see, for example, Gutenbrunner and Jurečková (1992), Portnoy (1991), Angrist, Chernozhukov, and Fernandez-Val (2006), and Chernozhukov and Hansen (2006).

[10]For example, Assumption 2 does not hold when we estimate monotone production or demand functions $u \mapsto f(u)$, where $u$ is input or price, using nonparametric kernel or series regression. We refer the reader to Chernozhukov, Fernandez-Val, and Galichon (2009) for appropriate additional results that enable us to perform uniform inference in such cases. On the other hand, Assumption 2 does hold when we estimate monotone production or demand functions $u \mapsto f(u)$ using parametric or semiparametric regression.

*as a stochastic process indexed by* $(y, x) \in \mathcal{YX}^*$; *and in* $\ell^\infty(\mathcal{UX}_K)$, *with* $\mathcal{UX}_K = \{(u, x) : (Q^*(u|x), x) \in K\}$,

$$a_n(\widehat{Q}^*(u|x) - Q^*(u|x)) \quad \Rightarrow \quad \widetilde{D}_G(u|x) \tag{2.13}$$

*as a stochastic process indexed by* $(u, x) \in \mathcal{UX}_K$, *where the maps* $h \mapsto D_h$ *and* $h \mapsto \widetilde{D}_h$ *are defined in equations* (2.6) *and* (2.8).

This proposition provides the basis for inference using rearranged quantile estimators and corresponding distribution estimators.

Let us first discuss inference for the case with a monotonic population curve $Q$. It is useful to emphasize the following corollary of Proposition 5:

COROLLARY 3—Functional Limit Laws for $\widehat{F}$ and $\widehat{Q}^*$ in the Monotonic Case: *Suppose* $u \mapsto Q(u|x)$ *has* $\partial_u Q(u|x) > 0$ *for each* $(u, x) \in \mathcal{UX}$. *Then* $\mathcal{YX}^* = \mathcal{YX}$ *and* $\mathcal{UX}^* = \mathcal{UX}$. *Accordingly, the convergence in Proposition* 5 *holds uniformly over the entire* $\mathcal{YX}$ *and* $\mathcal{UX}$. *Moreover,* $\widetilde{D}_G(u|x) = G(u|x)$, *that is, the rearranged quantile curves have the same first order asymptotic distribution as the original estimated quantile curves.*

Thus, if the population curve is monotone, we can rearrange the original nonmonotone quantile estimator to be monotonic *without* affecting its (first order) asymptotic properties. Hence, all the inference tools that apply to the original quantile estimator $\widehat{Q}$ also apply to the rearranged quantile estimator $\widehat{Q}^*$. In particular, if the bootstrap is valid for the original estimator, it is also valid for the rearranged estimator, by the functional delta method for the bootstrap. Thus, when $Q$ is monotone, Corollary 3 enables us to perform uniform inference on $Q$ and $F$ based on the rearranged estimators $\widehat{Q}^*$ and $\widehat{F}$.

REMARK 2—Detecting and Avoiding Cases With Nonmonotone $Q$: Before discussing inference for the case with a nonmonotonic population curve $Q$, let us first emphasize that since nonmonotonicity of $Q$ is a rather obvious sign of specification error, it is best to try to detect and avoid this case. For this purpose we should use sufficiently flexible functional forms and reject the ones that fail to pass monotonicity tests. For example, we can use the following generic test of monotonicity for $Q$: If $Q$ is monotone, the first order behavior of $\widehat{Q}^*$ and $\widehat{Q}$ coincides, and if $Q$ is not monotone, $\widehat{Q}^*$ and $\widehat{Q}$ converge to different probability limits $Q^*$ and $Q$. Therefore, we can reject the hypothesis of monotone $Q$ if a uniform confidence region for $Q$ based on $\widehat{Q}$ does not contain $\widehat{Q}^*$, for at least one point $x \in \mathcal{X}$.[11]

[11]This test is conservative, but it is generic and very inexpensive. To build nonconservative tests, we need to derive the limit laws for $\|\widehat{Q} - \widehat{Q}^*\|$ for suitable norms $\| \cdot \|$. These laws will depend on higher order functional limit laws for quantile estimators, which appear to be nongeneric and have to be dealt with on a case by case basis.

Let us now discuss inference for the case with a nonmonotonic population curve $Q$. In this case, the large sample properties of the rearranged quantile estimators $\widehat{Q}^*$ substantially differ from those of the initial quantile estimators $\widehat{Q}$. Proposition 5 still enables us to perform uniform inference on the rearranged population curve $Q^*$ based on the rearranged estimator $\widehat{Q}^*$, but only after excluding certain nonregular neighborhoods (for the distribution estimators, the neighborhoods of the critical values of the map $u \mapsto Q(u|x)$; for the rearranged quantile estimators, the image of these neighborhoods under $F$). These neighborhoods can be excluded by locating the points $(u, x)$ where a consistent estimate of $|\partial_u Q(u|x)|$ is close to zero; see Hendricks and Koenker (1991) for a consistent estimator of $|\partial_u Q(u|x)|$.

Next we consider the linear functionals of the rearranged quantile and distribution estimates,

$$(y', x) \mapsto \int_{\mathcal{Y}} g(y|x, y')\widehat{F}(y|x)\,dy,$$

$$(u', x) \mapsto \int_{\mathcal{U}} g(u|x, u')\widehat{Q}^*(u|x)\,du.$$

The following proposition derives functional limit laws for these functionals.[12] Here the convergence results hold without excluding any nonregular neighborhoods, which is convenient for practice in the nonmonotonic case.

PROPOSITION 6—Functional Limit Laws for Linear Functionals of $\widehat{Q}^*$ and $\widehat{F}$: *Under the same restrictions on the function $g$ as in Proposition 3, the following results hold with the limits being continuous on the specified domains*:

(i) *As a stochastic process indexed by $(y', x) \in \mathcal{YX}$, in $\ell^\infty(\mathcal{YX})$,*

$$a_n \int_{\mathcal{Y}} g(y|x, y')(\widehat{F}(y|x) - F(y|x))\,dy \Rightarrow \int_{\mathcal{Y}} g(y|x, y') D_G(y|x)\,dy. \tag{2.14}$$

(ii) *As a stochastic process indexed by $(u', x) \in \mathcal{UX}$, in $\ell^\infty(\mathcal{UX})$,*

$$a_n \int_{\mathcal{U}} g(u|x, u')(\widehat{Q}^*(u|x) - Q^*(u|x))\,du \Rightarrow \int_{\mathcal{U}} g(u|x, u')\widetilde{D}_G(u|x)\,du. \tag{2.15}$$

[12] Working with these functionals is equivalent to placing our empirical processes into the space $L^p$ ($p = 1$ for rearranged distributions and $p = \infty$ for quantiles), equipped with weak* topology, instead of strong topology. Convergence in law of the integral functionals, shown in Proposition 6, is equivalent to the convergence in law of the rearranged estimated processes in such a metric space.

The linear functionals defined above are useful building blocks for various statistics, such as partial means, various moments, and Lorenz curves. For example, the conditional Lorenz curve based on rearranged quantile estimators is

$$\widehat{L}(u'|x) := \left(\int_{\mathcal{U}} 1\{u \le u'\}\widehat{Q}^*(u|x)\,du\right)\Big/\left(\int_{\mathcal{U}} \widehat{Q}^*(u|x)\,du\right), \tag{2.16}$$

which is a ratio of partial and overall conditional means. Hadamard differentiability of the mapping

$$Q \mapsto L(u'|x) := \left(\int_{\mathcal{U}} 1\{u \le u'\}Q^*(u|x)\,du\right)\Big/\left(\int_{\mathcal{U}} Q^*(u|x)\,du\right), \tag{2.17}$$

with respect to $Q$ immediately follows from (a) the differentiability of a ratio $\beta/\gamma$ with respect to its numerator $\beta$ and denominator $\gamma$ at $\gamma \neq 0$, (b) Hadamard differentiability of the numerator and denominator in (2.17) with respect to $Q$ established in Proposition 3, and (c) the chain rule for the Hadamard derivative. Hence, provided that $Q > 0$ so that $Q^* > 0$, we have that in the metric space $\ell^\infty(\mathcal{U}\mathcal{X})$,

$$\begin{aligned} &a_n(\widehat{L}(u'|x) - L(u'|x)) \\ &\quad\Rightarrow\quad L(u'|x)\cdot\left(\frac{\int_{\mathcal{U}} 1\{u \le u'\}\widetilde{D}_G(u|x)\,du}{\int_{\mathcal{U}} 1\{u \le u'\}Q^*(u|x)\,du} - \frac{\int_{\mathcal{U}} \widetilde{D}_G(u|x)\,du}{\int_{\mathcal{U}} Q^*(u|x)\,du}\right) \end{aligned} \tag{2.18}$$

as an empirical process indexed by $(u', x) \in \mathcal{U}\mathcal{X}$. In particular, validity of the bootstrap for estimating this functional limit law in (2.18) holds by the functional delta method for the bootstrap.

We next consider the empirical properties of the smoothed curves obtained by applying the linear smoothing operator $S$ defined in (2.10) to $\widehat{F}$ and $\widehat{Q}^*$:

$$\begin{aligned} S\widehat{F}(y'|x) &:= \int k_\delta(y' - y)\widehat{F}(y|x)\,dy, \\ S\widehat{Q}^*(u'|x) &:= \int k_\delta(u' - u)\widehat{Q}^*(u|x)\,du. \end{aligned}$$

The following corollary immediately follows from Corollary 2 and the functional delta method.

COROLLARY 4—Functional Limit Laws for Smoothed $\widehat{Q}^*$ and $\widehat{F}$: *In* $\ell^\infty(\mathcal{Y}\mathcal{X})$,

$$a_n(S\widehat{F}(y'|x) - SF(y'|x)) \quad\Rightarrow\quad SD_G(y'|x), \tag{2.19}$$

*as a stochastic process indexed by* $(y', x) \in \mathcal{Y}\mathcal{X}$; *and in* $\ell^\infty(\mathcal{U}\mathcal{X})$,

$$a_n(S\widehat{Q}^*(u'|x) - SQ^*(u'|x)) \quad \Rightarrow \quad S\widetilde{D}_G(u'|x), \tag{2.20}$$

*as a stochastic process indexed by* $(u', x) \in \mathcal{U}\mathcal{X}$. *The results hold uniformly in the smoothing parameter* $\delta \in [\delta_1, \delta_2]$, *where* $\delta_1$ *and* $\delta_2$ *are positive constants.*

Thus, as in the case of linear functionals, we can perform inference on $SQ^*$ based on the smoothed rearranged estimates without excluding nonregular neighborhoods, which is convenient for practice in the nonmonotonic case. Furthermore, validity of the bootstrap for the smoothed curves follows by the functional delta method for the bootstrap. Last, we note that it is not possible to simultaneously allow $\delta \to 0$ and preserve the uniform convergence stated in the corollary.

Our final corollary asserts validity of the bootstrap for inference on rearranged estimators and their functionals. This corollary follows from the functional delta method for the bootstrap (e.g., Theorem 13.9 in Van der Vaart (1998)).

COROLLARY 5—Validity of the Bootstrap for Estimating Laws of Rearranged Estimators: *If the bootstrap consistently estimates the functional limit law* (2.11) *of the empirical process* $\{a_n(\widehat{Q}(u|x) - Q(u|x), (u, x) \in \mathcal{U}\mathcal{X}\}$, *then it also consistently estimates the functional limit laws* (2.12), (2.13), (2.14), (2.15), (2.18), (2.19), *and* (2.20).

## 3. EXAMPLES

In this section we apply rearrangement to the estimation of structural quantile and distribution functions. We show how rearrangement monotonizes instrumental quantile and distribution function estimates, and demonstrate how to perform inference on the target functions using the results developed in this paper. Using a supporting numerical example, we show that rearranged estimators noticeably improve upon original estimators and also outperform isotonized estimators. Thus, rearrangement is necessarily preferable to the standard approach of simply ignoring nonmonotonicity. Moreover, in quantile estimation problems, rearrangement is also preferable to the standard approach of isotonization used primarily in mean estimation problems.

### 3.1. *Empirical Example*

We consider estimation of the causal/structural effects of Vietnam veteran status $X \in \{0, 1\}$ in the quantiles and distribution of civilian earnings $Y$. Since veteran status is likely to be endogenous relative to potential civilian earnings, we employ an instrumental variables approach, using the U.S. draft lottery as an instrument for the Vietnam status (Angrist (1990)). We use the same data subset from the Current Population Survey as in Abadie (2002).[13] We then estimate structural quantile and distribution functions with the instrumental quantile regression estimator of Chernozhukov and Hansen (2005, 2006) and the instrumental distribution regression estimator of Abadie (2002). Under some assumptions, these procedures consistently estimate the structural quantile and distribution functions of interest.[14] However, like most estimation methods mentioned in the Introduction, neither of these procedures explicitly imposes monotonicity of the distribution and quantile functions. Accordingly, they can produce estimates in finite samples that are nonmonotonic due to either sampling variation or violations of instrument independence or other modeling assumptions. We monotonize these estimates using rearrangement and perform inference on the target structural functions using uniform confidence bands constructed via bootstrap. We use the programming language R to implement the procedures (R Development Core Team (2007)). We present our estimation and inference results in Figures 3–5.

In Figure 3, we show Abadie's estimates of the structural distribution of earnings for veterans and nonveterans (left panel) as well as their rearrangements (right panel). For both veterans and nonveterans, the original estimates of the distributions exhibit clear local nonmonotonicity. The rearrangement fixes this problem, producing increasing estimated distribution functions. In Figure 4, we show Chernozhukov and Hansen's estimates of the structural quantile functions of earnings for veterans (left panel) as well as their rearrangements (right panel). For both veterans and nonveterans, the estimates of the quantile functions exhibit pronounced local nonmonotonicity. The rearrangement fixes this problem, producing increasing estimated quantile functions. In the case of quantile functions, the nonmonotonicity problem is specially acute for the small sample of veterans.

In Figure 5, we plot uniform 90% confidence bands for the structural quantile functions of earnings for veterans and nonveterans, together with uniform 90% confidence bands for the effect of Vietnam veteran status on the quantile functions for earnings, which measures the difference between the structural quantile functions for veterans and nonveterans. We construct the uniform confidence bands using both the original estimators and the rearranged estimators based on 500 bootstrap repetitions and a fine net of quantile indices

[13]These data consist of a sample of white men, born in 1950–1953, from the March Current Population Surveys of 1979 and 1981–1985. The data include annual labor earnings, Vietnam veteran status, and an indicator on the Vietnam era lottery. There are 11,637 men in the sample, with 2461 Vietnam veterans and 3234 eligible for U.S. military service according to the draft lottery indicator. Abadie (2002) gave additional information on the data and the construction of the variables.

[14]More specifically, Abadie's (2002) procedure consistently estimates these functions for the subpopulation of compliers under instrument independence and monotonicity. Chernozhukov and Hansen's (2005, 2006) approach consistently estimates these functions for the entire population under instrument independence and rank similarity.

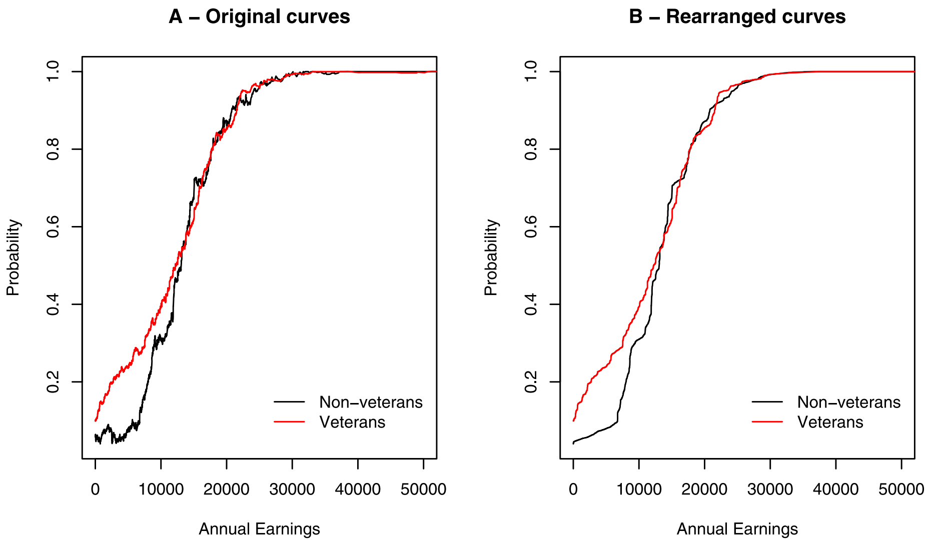


FIGURE 3.—Abadie's estimates of the structural distributions of earnings for veteran and non-veterans (left panel), and their rearrangements (right panel).

$\{0.01, 0.02, \ldots, 0.99\}$. We obtain the bands for the rearranged functions assuming that the population structural quantile regression functions are monotonic,

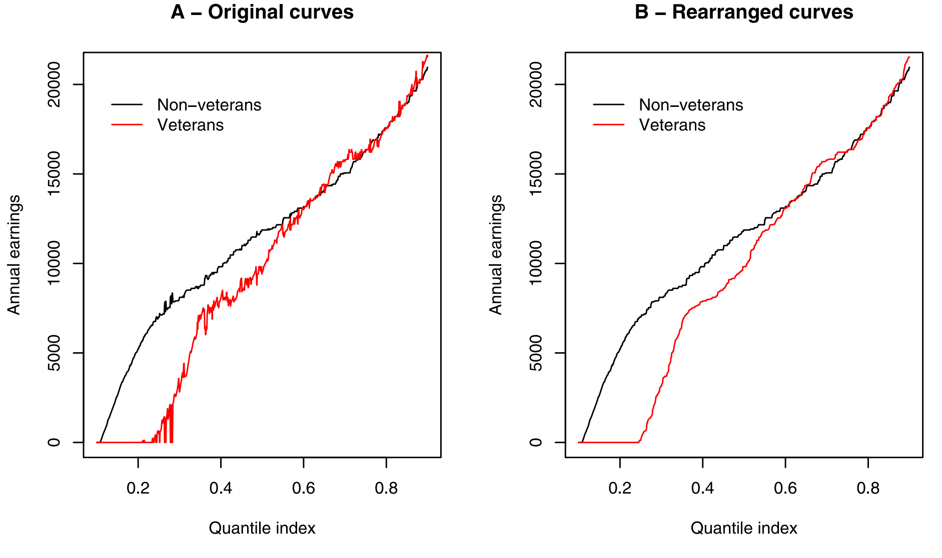


FIGURE 4.—Chernozhukov and Hansen's estimates of the structural quantile functions of earnings for veterans (left panel), and their rearrangements (right panel).

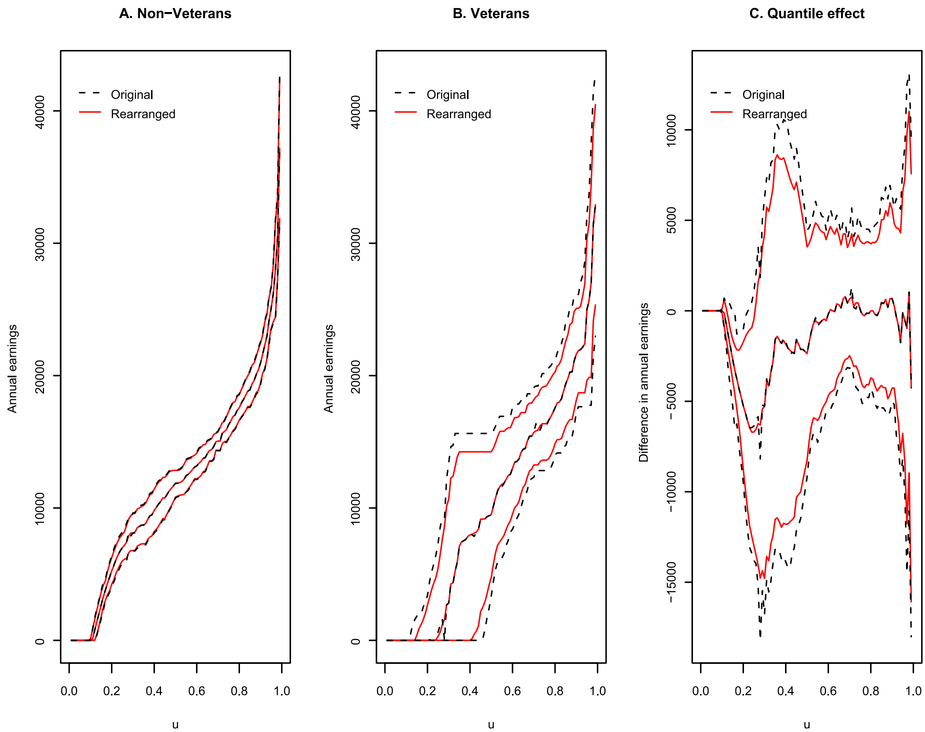


FIGURE 5.—Original and rearranged point estimates and corresponding simultaneous 90% confidence bands for structural quantile functions of earnings (panels A and B) and structural quantile effect of Vietnam veteran status on earnings (panel C). The bands for the quantile functions (panels A and B) are intersected with the class of monotone functions.

so that the first order behavior of the rearranged estimators coincides with the behavior of the original estimators. The figure shows that even for the large sample of nonveterans, the rearranged estimates lie within the original bands, thus passing our automatic test of monotonicity specified in Remark 2. Thus, the lack of monotonicity of the estimated quantile functions in this case is likely caused by sampling error. From the figure, we conclude that veteran status has a statistically significant negative effect in the lower tail, with the bands for the rearranged estimates showing a wider range of quantile indices for which this holds.

### 3.2. *Monte Carlo*

We design a Monte Carlo experiment to closely match the previous empirical example. In particular, we consider a location model, where the outcome is $Y = [1, X]\alpha + \varepsilon$, the endogenous regressor is $X = 1\{[1; Z]\pi + v \geq 0\}$, the instrument $Z$ is a binary random variable, and the disturbances $(\varepsilon, v)$ are jointly normal and independent of $Z$. The true structural quantile functions are $Q_0(u|x) = [1; x]\alpha + Q_\varepsilon(u)$, $x \in \{0, 1\}$, where $Q_\varepsilon$ is the quantile function of the normal variable $\varepsilon$. The corresponding structural distribution functions are the inverse of the quantile functions with respect to $u$. We select the value of the parameters by estimating this location model parametrically by maximum likelihood, and then generate samples from the estimated model, holding the values of the instruments $Z$ equal to those in the data set.[15] We use the estimators for the structural distribution and quantile functions described in the previous section. We monotonize the estimates using either rearrangement or isotonization. We use isotonization as a benchmark since it is the standard approach in mean regression problems (Mammen (1991)); it amounts to projecting the estimated function on the set of monotone functions.

Table I reports ratios of estimation errors of the rearranged and isotonized estimates to those of the original estimates, recorded in percentage terms. The target functions are the structural distribution and quantile functions. We measure estimation errors using the average $L^p$ norms $\|\cdot\|_p$ with $p = 1, 2$, and $\infty$, and we compute them as Monte Carlo averages of $\|f_0 - \tilde{f}\|_p$, where $f_0$ is the target function and $\tilde{f}$ is either the original or rearranged or isotonized estimate of this function.

We find that the rearranged estimators noticeably outperform the original estimators, achieving a reduction in estimation error up to 14%, depending on the target function and the norm. Moreover, in this case the better approximation of the rearranged estimates to the structural functions also produces more

TABLE I

RATIOS OF ESTIMATION ERROR OF REARRANGED AND ISOTONIC ESTIMATORS TO THOSE OF ORIGINAL ESTIMATORS, IN PERCENTAGE TERMS

| | Veterans | | Nonveterans | | Effect | |
|---|---|---|---|---|---|---|
| | Rearranged | Isotonized | Rearranged | Isotonized | Rearranged | Isotonized |
| | | | Structural Distribution Function | | | |
| $L^1$ | 99 | 99 | 97 | 98 | 97 | 98 |
| $L^2$ | 99 | 99 | 97 | 98 | 97 | 99 |
| $L^\infty$ | 96 | 98 | 90 | 94 | 91 | 95 |
| | | | Structural Quantile Function | | | |
| $L^1$ | 97 | 98 | 100 | 100 | 97 | 98 |
| $L^2$ | 96 | 97 | 100 | 100 | 96 | 98 |
| $L^\infty$ | 86 | 87 | 98 | 99 | 86 | 88 |

[15]More specifically, after normalizing the standard deviation of $v$ to 1, we set $\pi = [-0.92; 0.40]^\top$, $\alpha = [11{,}753; -911]^\top$, the standard deviation of $\varepsilon$ to 8100, and the covariance between $\varepsilon$ and $v$ to 379. We draw 5000 Monte Carlo samples of size $n = 11{,}627$. We generate the values of $Y$ and $X$ by drawing disturbances $(\varepsilon, v)$ from a bivariate normal distribution with zero mean and the estimated covariance matrix.

accurate estimates of the distribution and quantile effects, achieving a 3% to 9% reduction in estimation error for the distribution estimator and a 3% to 14% reduction in estimation error for the quantile estimator, depending on the norm.

We also find that the rearranged estimators outperform the isotonized estimators, achieving up to a further 4% reduction in estimation error, depending on the target function and the norm. The reason is that isotonization projects the original fitted function on the set of monotone functions, converting nonmonotone segments into flat segments. In contrast, rearrangement sorts the original fitted function, converting nonmonotone segments into steep, increasing segments that preserve measure. In the context of estimating quantile and distribution functions, the target functions tend to be nonflat, suggesting that rearrangement should be typically preferred over isotonization.[16]

## 4. CONCLUSION

This paper develops a monotonization procedure for estimation of conditional and structural quantile and distribution functions based on rearrangement-related operations. Starting from a possibly nonmonotone empirical curve, the procedure produces a rearranged curve that not only satisfies the natural monotonicity requirement, but also has smaller estimation error than the original curve. We derive asymptotic distribution theory for the rearranged curves, and illustrate the usefulness of the approach with an empirical application and a simulation example. There are many potential applications of the results given in this paper and companion work (Chernozhukov, Fernandez-Val, and Galichon (2009)) to other econometric problems with shape restrictions (see, e.g., Matzkin (1994)).

## APPENDIX A: PROOFS

PROOF OF PROPOSITION 1: First, note that the distribution of $Y_x$ has no atoms, that is,

$$\begin{aligned}\Pr[Y_x = y] &= \Pr[Q(U|x) = y] \\ &= \Pr\big[U \in \{u \in \mathcal{U} : u \text{ is a root of } Q(u|x) = y\}\big] = 0,\end{aligned}$$

[16]To give some intuition about this point, it is instructive to consider a simple example with a two-point domain $\{0, 1\}$. Suppose that the target function $f_0 : \{0, 1\} \to \mathbb{R}$ is increasing and steep, namely $f_0(0) > f_0(1)$, and the fitted function $\widehat{f} : \{0, 1\} \to \mathbb{R}$ is decreasing, with $\widehat{f}(0) > \widehat{f}(1)$. In this case, isotonization produces a nondecreasing function $\bar{f} : \{0, 1\} \to \mathbb{R}$, which is flat, with $\bar{f}(0) = \bar{f}(1) = [\widehat{f}(0) + \widehat{f}(1)]/2$, which is somewhat unsatisfactory. In such cases rearrangement can significantly outperform isotonization, since it produces the steepest fit, namely it produces $\widehat{f}^* : \{0, 1\} \to \mathbb{R}$ with $\widehat{f}^*(0) = \widehat{f}(1) < \widehat{f}^*(1) = \widehat{f}(0)$. This observation provides a simple theoretical underpinning for the estimation results we see in Table I.

since the number of roots of $Q(u|x) = y$ is finite under Assumption 1 and $U \sim \text{Uniform}(\mathcal{U})$. Next, by Assumption 1, the number of critical values of $Q(u|x)$ is finite, hence claim (i) follows.

Next, for any regular $y$, we can write $F(y|x)$ as

$$\int_0^1 1\{Q(u|x) \le y\}\, du = \sum_{k=0}^{K(y|x)-1} \int_{u_k(y|x)}^{u_{k+1}(y|x)} 1\{Q(u|x) \le y\}\, du + \int_{u_{K(y|x)}(y|x)}^{1} 1\{Q(u|x) \le y\}\, du,$$

where $u_0(y|x) := 0$ and $\{u_k(y|x),\ \text{for } k = 1, 2, \dots, K(y|x) < \infty\}$ are the roots of $Q(u|x) = y$ in increasing order. Note that the sign of $\partial_u Q(u|x)$ alternates over consecutive $u_k(y|x)$, determining whether $1\{Q(y|x) \le y\} = 1$ on the interval $[u_{k-1}(y|x), u_k(y|x)]$. Hence the first term in the previous expression simplifies to $\sum_{k=0}^{K(y|x)-1} 1\{\partial_u Q(u_{k+1}(y|x)|x) \ge 0\}(u_{k+1}(y|x) - u_k(y|x))$, while the last term simplifies to $1\{\partial_u Q(u_{K(y|x)}(y|x)|x) \le 0\}(1 - u_{K(y|x)}(y|x))$. An additional simplification yields the expression given in claim (ii) of the proposition.

The proof of claim (iii) follows by taking the derivative of the expression in claim (ii), noting that at any regular value $y$, the number of solutions $K(y|x)$ and $\text{sign}(\partial_u Q(u_k(y|x)|x))$ are locally constant; moreover,

$$\partial_y u_k(y|x) = \frac{\text{sign}(\partial_u Q(u_k(y|x)|x))}{|\partial_u Q(u_k(y|x)|x)|}.$$

Combining these facts, we get the expression for the derivative given in claim (iii).

To show the absolute continuity of $F$ with $f$ being the Radon–Nikodym derivative, it suffices to show that for each $y' \in \mathcal{Y}_x$, $\int_{-\infty}^{y'} f(y|x)\, dy = \int_{-\infty}^{y'} dF(y|x)$ (cf. Theorem 31.8 in Billingsley (1995)). Let $V_t^x$ be the union of closed balls of radius $t$ centered on the critical points $\mathcal{Y}_x \setminus \mathcal{Y}_x^*$, and define $\mathcal{Y}_x^t = \mathcal{Y}_x \setminus V_t^x$. Then $\int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\} f(y|x)\, dy = \int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\}\, dF(y|x)$. Since the set of critical points $\mathcal{Y}_x \setminus \mathcal{Y}_x^*$ is finite and has mass zero under $F$, $\int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\}\, dF(y|x) \uparrow \int_{-\infty}^{y'} dF(y|x)$ as $t \to 0$. Therefore, $\int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\} f(y|x)\, dy \uparrow \int_{-\infty}^{y'} f(y|x)\, dy = \int_{-\infty}^{y'} dF(y|x)$.

Claim (iv) follows by noting that at the regions where $s \mapsto Q(s|x)$ is increasing and one-to-one, we have that $F(y|x) = \int_{Q(s|x) \le y} ds = \int_{s \le Q^{-1}(y|x)} ds = Q^{-1}(y|x)$. Inverting the equation $u = F(Q^*(u|x)|x) = Q^{-1}(Q^*(u|x)|x)$ yields $Q^*(u|x) = Q(u|x)$.

From claim (v), note that $Y_x = Q(U|x)$ has quantile function $Q^*$. A quantile function is known to be equivariant to monotone increasing transformations, including location–scale transformations. Thus, this is true in particular for $Q^*$.

Claim (vi) is immediate from claim (iii).

(vii) The proof of continuity of $F$ is subsumed in the Step 1 of the proof of Proposition 2 (see below). Therefore, for any sequence $x_t \to x$, we have that $F(y|x_t) \to F(y|x)$ uniformly in $y$, and $F$ is continuous. Let $u_t \to u$ and $x_t \to x$. Since $F(y|x) = u$ has a unique root $y = Q^*(u|x)$, the root of $F(y|x_t) = u_t$, that is, $y_t = Q^*(u_t|x_t)$, converges to $y$ by a standard argument; see, for example, Van der Vaart and Wellner (1996). *Q.E.D.*

In the proofs of Propositions 2 and 3 that follow we will repeatedly use Lemma 1, which establishes the equivalence of continuous convergence and uniform convergence.

LEMMA 1: *Let $\mathbb{D}$ and $\mathbb{D}'$ be complete separable metric spaces, with $\mathbb{D}$ compact. Suppose $f : \mathbb{D} \mapsto \mathbb{D}'$ is continuous. Then a sequence of functions $f_n : \mathbb{D} \mapsto \mathbb{D}'$ converges to $f$ uniformly on $\mathbb{D}$ if and only if for any convergent sequence $x_n \to x$ in $\mathbb{D}$, we have that $f_n(x_n) \to f(x)$.*

For the proof of Lemma 1, see, for example, Resnick (1987, p. 2).

PROOF OF PROPOSITION 2: (i) We have that for any $\delta > 0$, there exists $\varepsilon > 0$ such that for $u \in B_\varepsilon(u_k(y|x))$ and for small enough $t \geq 0$,

$$1\{Q(u|x) + th_t(u|x) \leq y\} \leq 1\big\{Q(u|x) + t\big(h(u_k(y|x)|x) - \delta\big) \leq y\big\}$$

for all $k \in \{1, 2, \ldots, K(y|x)\}$, whereas for all $u \notin \bigcup_k B_\varepsilon(u_k(y|x))$, as $t \to 0$,

$$1\{Q(u|x) + th_t(u|x) \leq y\} = 1\{Q(u|x) \leq y\}.$$

Therefore,

$$\frac{\int_0^1 1\{Q(u|x) + th_t(u|x) \leq y\}\, du - \int_0^1 1\{Q(u|x) \leq y\}\, du}{t} \tag{A.1}$$
$$\leq \sum_{k=1}^{K(y|x)} \int_{B_\varepsilon(u_k(y|x))} \big(1\big\{Q(u|x) + t\big(h(u_k(y|x)|x) - \delta\big) \leq y\big\}$$
$$- 1\{Q(u|x) \leq y\}\big)/t\, du,$$

which by the change of variable $y' = Q(u|x)$ is equal to

$$\frac{1}{t} \sum_{k=1}^{K(y|x)} \int_{J_k \cap [y, y - t(h(u_k(y|x)|x) - \delta)]} \frac{1}{|\partial_u Q(Q^{-1}(y'|x)|x)|}\, dy',$$

where $J_k$ is the image of $B_\varepsilon(u_k(y|x))$ under $u \mapsto Q(\cdot|x)$. The change of variable is possible because for $\varepsilon$ small enough, $Q(\cdot|x)$ is one-to-one between $B_\varepsilon(u_k(y|x))$ and $J_k$.

Fixing $\varepsilon > 0$, for $t \to 0$, we have that $J_k \cap [y, y - t(h(u_k(y|x)|x) - \delta)] = [y, y - t(h(u_k(y|x)|x) - \delta)]$ and $|\partial_u Q(Q^{-1}(y'|x)|x)| \to |\partial_u Q(u_k(y|x)|x)|$ as $Q^{-1}(y'|x) \to u_k(y|x)$. Therefore, the right hand term in (A.1) is no greater than

$$\sum_{k=1}^{K(y|x)} \frac{-h(u_k(y|x)|x) + \delta}{|\partial_u Q(u_k(y|x)|x)|} + o(1).$$

Similarly

$$\sum_{k=1}^{K(y|x)} \frac{-h(u_k(y|x)|x) - \delta}{|\partial_u Q(u_k(y|x)|x)|} + o(1)$$

bounds (A.1) from below. Since $\delta > 0$ can be made arbitrarily small, the result follows.

To show that the result holds uniformly in $(y, x) \in K$, a compact subset of $\mathcal{YX}^*$, we use Lemma 1. Take a sequence of $(y_t, x_t)$ in $K$ that converges to $(y, x) \in K$. Then the preceding argument applies to this sequence, since (a) the function $(y, x) \mapsto -h(u_k(y|x)|x)/|\partial_u Q(u_k(y|x)|x)|$ is uniformly continuous on $K$ and (b) the function $(y, x) \mapsto K(y|x)$ is uniformly continuous on $K$. To see (b), note that $K$ excludes a neighborhood of critical points $(\mathcal{Y} \setminus \mathcal{Y}_x^*, x \in \mathcal{X})$, and therefore can be expressed as the union of a finite number of compact sets $(K_1, \dots, K_M)$ such that the function $K(y|x)$ is constant over each of these sets, that is, $K(y|x) = k_j$ for some integer $k_j > 0$, for all $(y, x) \in K_j$ and $j \in \{1, \dots, M\}$. Likewise, (a) follows by noting that the limit expression for the derivative is continuous on each of the sets $(K_1, \dots, K_M)$ by the assumed continuity of $h(u|x)$ in both arguments, continuity of $u_k(y|x)$ (implied by the implicit function theorem), and the assumed continuity of $\partial_u Q(u|x)$ in both arguments.

(ii) For a fixed $x$, the result follows by part (i) of this proposition, by Step 1 below, and by an application of the Hadamard differentiability of the quantile operator shown by Doss and Gill (1992). Step 2 establishes uniformity over $x \in \mathcal{X}$.

*Step* 1. Let $K$ be a compact subset of $\mathcal{YX}^*$. Let $(y_t, x_t)$ be a sequence in $K$, convergent to a point, say $(y, x)$. Then, for every such sequence, $\varepsilon_t := t\|h_t\|_\infty + \|Q(\cdot|x_t) - Q(\cdot|x)\|_\infty + |y_t - y| \to 0$, and

$$\text{(A.2)} \quad |F(y_t|x_t, h_t) - F(y|x)| \leq \left| \int_0^1 \left[1\{Q(u|x_t) + t h_t(u|x) \leq y_t\} - 1\{Q(u|x) \leq y\}\right] du \right|$$

$$\leq \left| \int_0^1 1\{|Q(u|x) - y| \leq \varepsilon_t\}\, du \right| \to 0,$$

where the last step follows from the absolute continuity of $y \mapsto F(y|x)$, the distribution function of $Q(U|x)$. By setting $h_t = 0$, the above argument also verifies that $F(y|x)$ is continuous in $(y, x)$. Lemma 1 implies uniform convergence of $F(y|x, h_t)$ to $F(y|x)$, which in turn implies by a standard argument[17] the uniform convergence of quantiles $Q^*(u|x, h_t) \to Q^*(u|x)$, uniformly over $K^*$, where $K^*$ is any compact subset of $\mathcal{U}\mathcal{X}^*$.

*Step* 2. We have that uniformly over $K^*$,

$$\text{(A.3)} \qquad \frac{F(Q^*(u|x, h_t)|x, h_t) - F(Q^*(u|x, h_t)|x)}{t} = D_h(Q^*(u|x, h_t)|x) + o(1) = D_h(Q^*(u|x)|x) + o(1),$$

using part (i) of Proposition 2, Step 1, and the continuity properties of $D_h(y|x)$. Furthermore, uniformly over $K^*$, by Taylor expansion and Proposition 1, as $t \to 0$,

$$\text{(A.4)} \qquad \frac{F(Q^*(u|x, h_t)|x) - F(Q^*(u|x)|x)}{t} = f(Q^*(u|x)|x)\frac{Q^*(u|x, h_t) - Q^*(u|x)}{t} + o(1)$$

and (as will be shown below)

$$\text{(A.5)} \qquad \frac{F(Q^*(u|x, h_t)|x, h_t) - F(Q^*(u|x)|x)}{t} = o(1)$$

as $t \to 0$. Observe that the left-hand side of (A.5) equals that of (A.4) plus that of (A.3). The result then follows.

It only remains to show that equation (A.5) holds uniformly in $K^*$. Note that for any right-continuous cumulative distribution function (c.d.f.) $F$, we have that $u \leq F(Q^*(u)) \leq u + F(Q^*(u)) - F(Q^*(u)-)$, where $F(\cdot-)$ denotes the left limit of $F$, that is, $F(x_0-) = \lim_{x \uparrow x_0} F(x)$. For any continuous, strictly increasing c.d.f. $F$, we have that $F(Q^*(u)) = u$. Therefore, write

$$0 \leq \frac{F(Q^*(u|x, h_t)|x, h_t) - F(Q^*(u|x)|x)}{t} \leq \frac{u + F(Q^*(u|x, h_t)|x, h_t) - F(Q^*(u|x, h_t) - |x, h_t) - u}{t}$$

[17]See, for example, Lemma 1 in Chernozhukov and Fernandez-Val (2005).

$$\leq \frac{F(Q^*(u|x,h_t)|x,h_t) - F(Q^*(u|x,h_t)-|x,h_t)}{t}$$

$$\overset{(1)}{=} \frac{[F(Q^*(u|x,h_t)|x,h_t) - F(Q^*(u|x,h_t)|x)]}{t}$$

$$- \frac{[F(Q^*(u|x,h_t)-|x,h_t) - F(Q^*(u|x,h_t)-|x)]}{t}$$

$$\overset{(2)}{=} D_h(Q^*(u|x,h_t)|x) - D_h(Q^*(u|x,h_t)-|x) + o(1) = o(1),$$

as $t \to 0$, where in (1) we use that $F(Q^*(u|x,h_t)|x) = F(Q^*(u|x,h_t)-|x)$ since $F(y|x)$ is continuous and strictly increasing in $y$, and in (2) we use part (i) of Proposition 2. *Q.E.D.*

The following lemma, due to Pratt (1960), will be very useful to prove Lemma 3 and Proposition 3.

LEMMA 2: *Let $|f_n| \leq G_n$, and suppose that $f_n \to f$ and $G_n \to G$ almost everywhere. Then if $\int G_n \to \int G$ and $\int G$ is finite, then $\int f_n \to \int f$.*

LEMMA 3—Boundedness and Integrability Properties: *Under the hypotheses of Proposition 2, we have that, for all $(u,x) \in \mathcal{UX}$ and $h'_t \in \ell^\infty(\mathcal{UX})$,*

$$|\widetilde{D}_{h'_t}(u|x,t)| \leq \|h'_t\|_\infty, \tag{A.6}$$

*and, for all $(y,x) \in \mathcal{YX}$,*

$$|D_{h'_t}(y|x,t)| \leq \Delta(y|x,t) = \int_0^1 \frac{1\{|Q(u|x) - y| \leq t\|h'_t\|_\infty\}}{t}\,du, \tag{A.7}$$

*where for any $x_t \to x \in \mathcal{X}$ and $\|h'_t - h'\|_\infty \to 0$ with $h' \in C(\mathcal{UX})$, as $t \to 0$,*

$$\Delta(y|x_t,t) \to 2\|h'\|_\infty f(y|x) \quad \textit{for a.e. } y \in \mathcal{Y} \quad \textit{and}$$

$$\int_{\mathcal{Y}} \Delta(y|x_t,t)\,dy \to \int_{\mathcal{Y}} 2\|h'\|_\infty f(y|x)\,dy.$$

PROOF: To show (A.6), note that

$$\sup_{(u,x)\in\mathcal{UX}} |\widetilde{D}_{h'_t}(u|x,t)| \leq \|h'_t\|_\infty \tag{A.8}$$

immediately follows from the equivariance property noted in claim (v) of Proposition 1.

The inequality (A.7) is trivial. That for any $x_t \to x \in \mathcal{X}$, $\Delta(y|x_t,t) \to 2\|h'\|_\infty f(y|x)$ for a.e. $y \in \mathcal{Y}$ follows by applying Proposition 2, with functions $h_t(u|x) = \|h'_t\|_\infty$ and $h_t(u|x) = -\|h'_t\|_\infty$ (for the case when $f(y|x) > 0$, and trivially otherwise). Similarly, that for any $y_t \to y \in \mathcal{Y}$, $\Delta(y_t|x,t) \to 2\|h'\|_\infty f(y|x)$ for a.e. $x \in \mathcal{X}$ follows by Proposition 2 (for the case when $f(y|x) > 0$, and trivially otherwise).

Furthermore, by Fubini's theorem,

$$\int_{\mathcal{Y}} \Delta(y|x_t,t)\,dy = \int_0^1 \underbrace{\left(\int_{\mathcal{Y}} \frac{1\{|Q(u|x_t) - y| \le t\|h'_t\|_\infty\}}{t}\,dy\right)}_{=:\,f_t(u)} du. \tag{A.9}$$

Note that $f_t(u) \le 2\|h'_t\|_\infty$. Moreover, for almost every $u$, $f_t(u) = 2\|h'_t\|_\infty$ for small enough $t$, and $2\|h'_t\|_\infty$ converges to $2\|h'\|_\infty$ as $t \to 0$. Then, trivially, $2\int_0^1 \|h'_t\|_\infty\,du \to 2\|h'\|_\infty$. By Lemma 2, the right-hand side of (A.9) converges to $2\|h'\|_\infty$. *Q.E.D.*

PROOF OF PROPOSITION 3: Define $m_t(y|x,y') := g(y|x,y')D_{h_t}(y|x,t)$ and $m(y|x,y') := g(y|x,y')D_h(y|x)$. To show claim (i), we need to demonstrate that for any $y'_t \to y'$ and $x_t \to x$,

$$\int_{\mathcal{Y}} m_t(y|x_t,y'_t)\,dy \to \int_{\mathcal{Y}} m(y|x,y')\,dy, \tag{A.10}$$

and that the limit is continuous in $(x,y')$. We have that $|m_t(y|x_t,y_t)|$ is bounded, for some constant $C$, by $C\Delta(y|x_t,t)$ which converges a.e. and the integral of which converges to a finite number by Lemma 3. Moreover, by Proposition 2, for almost every $y$, we have $m_t(y|x_t,y'_t) \to m(y|x,y')$. We conclude that (A.10) holds by Lemma 2.

To check continuity, we need to show that for any $y'_t \to y'$ and $x_t \to x$,

$$\int_{\mathcal{Y}} m(y|x_t,y'_t)\,dy \to \int_{\mathcal{Y}} m(y|x,y')\,dy. \tag{A.11}$$

We have that $m(y|x_t,y'_t) \to m(y|x,y')$ for almost every $y$. Moreover, $m(y|x_t,y_t)$ is dominated by $2\|g\|_\infty\|h\|_\infty f(y|x_t)$, which converges to $2\|g\|_\infty\|h\|_\infty f(y|x)$ for almost every $y$, and, moreover, $\int_{\mathcal{Y}} \|g\|_\infty\|h\|_\infty f(y|x)\,dy$ converges to $\|g\|_\infty\|h\|_\infty$. We conclude that (A.11) holds by Lemma 2.

To show claim (ii), define $m_t(u|x,u') = g(u|x,u')\widetilde{D}_{h_t}(u|x)$ and $m(u|x,u') = g(u|x,u')\widetilde{D}_h(u|x)$. Here we need to show that for any $u'_t \to u'$ and $x_t \to x$,

$$\int_{\mathcal{U}} m_t(u|x_t,u'_t)\,du \to \int_{\mathcal{U}} m(u|x,u')\,du, \tag{A.12}$$

and that the limit is continuous in $(u',x)$. We have that $m_t(u|x_t,u'_t)$ is bounded by $g(u|x_t)\|h_t\|_\infty$, which converges to $g(u|x)\|h\|_\infty$ for a.e. $u$. Furthermore, the integral of $g(u|x_t)\|h_t\|_\infty$ converges to the integral of $g(u|x)\|h\|_\infty$ by the dominated convergence theorem. Moreover, by Proposition 2, we have that $m_t(u|x_t, u_t') \to m(u|x, u')$ for almost every $u$. We conclude that (A.12) holds by Lemma 2.

To check the continuity of the limit, we need to show that for any $u_t' \to u'$ and $x_t \to x$,

$$\int_{\mathcal{U}} m(u|x_t, u_t')\, du \to \int_{\mathcal{U}} m(u|x, u')\, du. \tag{A.13}$$

We have that $m(u|x_t, u_t') \to m(u|x, u')$ for almost every $u$. Moreover, for small enough $t$, $m(u|x_t, u_t')$ is dominated by $|g(u|x_t, u_t')|\|h\|_\infty$, which converges for almost every value of $u$ to $|g(u|x, u')|\|h\|_\infty$ as $t \to 0$. Furthermore, the integral of $|g(u|x_t, u_t')|\|h\|_\infty$ converges to the integral of $|g(u|x, u')|\|h\|_\infty$ by the dominated convergence theorem. We conclude that (A.13) holds by Lemma 2. *Q.E.D.*

PROOF OF PROPOSITION 5: This proposition simply follows by the functional delta method (e.g., Van der Vaart (1998)). Instead of restating this method, it takes less space to simply recall the proof in the current context.

To show the first part, consider the map $g_n(y, x|h) = a_n(F(y|x, h/a_n) - F(y|x))$. The sequence of maps satisfies $g_{n'}(y, x|h_{n'}) \to D_h(y|x)$ in $\ell^\infty(K)$ for every subsequence $h_{n'} \to h$ in $\ell^\infty(\mathcal{U}\mathcal{X})$, where $h$ is continuous. It follows by the extended continuous mapping theorem that, in $\ell^\infty(K)$, $g_n(y, x|a_n(\widehat{Q}(u|x) - Q(u|x))) \Rightarrow D_G(y|x)$ as a stochastic process indexed by $(y, x)$, since $a_n(\widehat{Q}(u|x) - Q(u|x)) \Rightarrow G(u|x)$ in $\ell^\infty(\mathcal{U}\mathcal{X})$.

We conclude similarly for the second part. *Q.E.D.*

PROOF OF PROPOSITION 6: This follows by the functional delta method, similarly to the proof of Proposition 5. *Q.E.D.*

## REFERENCES


ABADIE, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Models," *Journal of the American Statistical Association*, 97, 284–292. [1094,1097,1111]

ABADIE, A., J. ANGRIST, AND G. IMBENS (2002): "Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings," *Econometrica*, 70, 91–117. [1097]

ALVINO, A., P. L. LIONS, AND G. TROMBETTI (1989): "On Optimization Problems With Prescribed Rearrangements," *Nonlinear Analysis*, 13, 185–220. [1104]

ANGRIST, J. D. (1990): "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence From Social Security Administrative Records," *American Economic Review*, 80, 313–336. [1110]

ANGRIST, J. D., V. CHERNOZHUKOV, AND I. FERNANDEZ-VAL (2006): "Quantile Regression Under Misspecification, With an Application to the U.S. Wage Structure," *Econometrica*, 74, 539–563. [1106]

BASSETT, G., JR., AND R. KOENKER (1982): "An Empirical Quantile Function for Linear Models With iid Errors," *Journal of the American Statistical Association*, 77, 407–415. [1093,1096]

BELLONI, A., AND V. CHERNOZHUKOV (2007): "Conditional Quantile and Probability Processes Under Increasing Dimension," Preprint, MIT. [1106]

BILLINGSLEY, P. (1995): *Probability and Measure* (Third Ed.). New York: Wiley. [1116]
BLUNDELL, R., AND J. POWELL (2003): "Endogeneity in Nonparametric and Semiparametric Models," in *Advances in Econometrics, Eight World Congress*, Vol. II, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky. Cambridge: Cambridge University Press. [1094]
BRONSHTEIN, I. N., K. A. SEMENDYAYEV, G. MUSIOL, AND H. MUEHLIG (2003): *Handbook of Mathematics* (Fourth Ed.). Berlin: Springer-Verlag. [1095]
BUCHINSKY, M. (1994): "Changes in the US Wage Structure 1963–1987: Application of Quantile Regression," *Econometrica*, 62, 405–458. [1097]
BUCHINSKY, M., AND J. HAHN (1998): "An Alternative Estimator for the Censored Quantile Regression Model," *Econometrica*, 66, 653–671. [1097]
CARLIER, G., AND R.-A. DANA (2005): "Rearrangement Inequalities in Non-Convex Insurance Problems," *Journal of Mathematical Economics*, 41, 483–503. [1095]
CHAMBERLAIN, G. (1994): "Quantile Regression, Censoring, and the Structure of Wages," in *Advances in Econometrics, Sixth World Congress*, Vol. 1, ed. by C. A. Sims. Cambridge: Cambridge University Press. [1097]
CHAUDHURI, S. (1991): "Nonparametric Estimates of Regression Quantiles and Their Local Bahadur Representation," *The Annals of Statistics*, 19, 760–777. [1097]
CHERNOZHUKOV, V., AND I. FERNÁNDEZ-VAL (2005): "Subsampling Inference on Quantile Regression Processes," *Sankhyā*, 67, 253–276. [1119]
CHERNOZHUKOV, V., AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245–261. [1094,1111]
——— (2006): "Instrumental Quantile Regression Inference for Structural and Treatment Effect Models," *Journal of Econometrics*, 132, 491–525. [1094,1097,1106,1111]
CHERNOZHUKOV, V., I. FERNANDEZ-VAL, AND A. GALICHON (2009): "Improving Point and Interval Estimators of Monotone Functions by Rearrangement," *Biometrika*, 96, 559–575. [1094,1095,1105,1106,1115]
——— (2010): "Supplement to 'Quantile and Probability Curves Without Crossing'," *Econometrica Supplemental Material*, 78, http://www.econometricsociety.org/ecta/Supmat/7880_data and programs.zip. [1096]
CHESHER, A. (2003): "Identification in Nonseparable Models," *Econometrica*, 71, 1405–1441. [1094]
DAVYDOV, Y., AND R. ZITIKIS (2005): "An Index of Monotonicity and Its Estimation: A Step Beyond Econometric Applications of the Gini Index," *Metron—International Journal of Statistics*, 63, 351–372. [1095]
DETTE, H., AND R. SCHEDER (2006): "Strictly Monotone and Smooth Nonparametric Regression for Two or More Variables," *The Canadian Journal of Statistics*, 34, 535–561. [1095]
DETTE, H., AND S. VOLGUSHEV (2008): "Non-Crossing Nonparametric Estimates of Quantile Curves," *Journal of the Royal Statistical Society, Ser. B*, 70, 609–627. [1095]
DETTE, H., N. NEUMEYER, AND K. PILZ (2006): "A Simple Nonparametric Estimator of a Strictly Monotone Regression Function," *Bernoulli*, 12, 469–490. [1095,1097]
DOKSUM, K. (1974): "Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case," *The Annals of Statistics*, 2, 267–277. [1094]
DOSS, H., AND R. D. GILL (1992): "An Elementary Approach to Weak Convergence for Quantile Processes, With Applications to Censored Survival Data," *Journal of the American Statistical Association*, 87, 869–877. [1103,1118]
DUDLEY, R. M., AND R. NORVAISA (1999): *Differentiability of Six Operators on Nonsmooth Functions and p-Variation*. Berlin: Springer-Verlag. [1103]
EGGERMONT, P. P. B., AND V. N. LARICCIA (2000): "Maximum Likelihood Estimation of Smooth Monotone and Unimodal Densities," *The Annals of Statistics*, 28, 922–947. [1106]
FOUGÈRES, A.-L. (1997): "Estimation de Densités Unimodales," *Canadian Journal of Statistics*, 25, 375–387. [1095]
GILL, R. D., AND S. JOHANSEN (1990): "A Survey of Product-Integration With a View Toward Application in Survival Analysis," *The Annals of Statistics*, 18, 1501–1555. [1103]

GLAD, I. K., N. L. HJORT, AND N. G. USHAKOV (2003): "Corrections of Density Estimators That Are Not Densities," *Scandinavian Journal of Statistics*, 30, 415–427. [1106]

GUTENBRUNNER, C., AND J. JUREČKOVÁ (1992): "Regression Quantile and Regression Rank Score Process in the Linear Model and Derived Statistics," *The Annals of Statistics*, 20, 305–330. [1106]

HALL, P., R. WOLFF, AND Q. YAO (1999): "Methods for Estimating a Conditional Distribution Function," *Journal of the American Statistical Association*, 94, 154–163. [1097]

HARDY, G., J. LITTLEWOOD, AND G. POLYA (1952): *Inequalities*. Cambridge: Cambridge University Press. [1094,1095,1097]

HE, X. (1997): "Quantile Curves Without Crossing," *The American Statistician*, 51, 186–192. [1093,1094,1097]

HE, X., AND Q.-M. SHAO (2000): "On Parameters of Increasing Dimensions," *Journal of Multivariate Analysis*, 73, 120–135. [1106]

HENDRICKS, W., AND R. KOENKER (1991): "Hierarchical Spline Models for Conditional Quantiles and the Demand for Electricity," *Journal of the American Statistical Association*, 87, 58–68. [1108]

HONORÉ, B., S. KHAN, AND J. POWELL (2002): "Quantile Regression Under Random Censoring," *Journal of Econometrics*, 109, 67–105. [1097]

IMBENS, G. W., AND W. K. NEWEY (2009): "Identification and Estimation of Nonseparable Triangular Simultaneous Equations Models Without Additivity," *Econometrica*, 77, 1481–1512. [1094]

KOENKER, R. (1994): "Confidence Intervals for Regression Quantiles," in *Asymptotic Statistics: Proceeding of the 5th Prague Symposium on Asymptotic Statistics*, ed. by P. Mandl and M. Hušková. Heidelberg, Germany: Physica-Verlag. [1095,1097]

——— (2005): *Quantile Regression*. Econometric Society Monograph Series, Vol. 38. Cambridge: Cambridge University Press. [1094,1098]

——— (2007): "quantreg: Quantile Regression," R package version 4.10. Available at http://www.r-project.org. [1093]

KOENKER, R., AND G. BASSETT (1978): "Regression Quantiles," *Econometrica*, 46, 33–50. [1094, 1097]

KOENKER, R., AND P. NG (2005): "Inequality Constrained Quantile Regression," *Sankhyā*, 67, 418–440. [1094]

LEHMANN, E. (1974): *Nonparametrics: Statistical Methods Based on Ranks*. San Francisco: Holden-Day. [1094]

LORENTZ, G. G. (1953): "An Inequality for Rearrangements," *The American Mathematical Monthly*, 60, 176–179. [1095,1105]

MA, L., AND R. KOENKER (2006): "Quantile Regression Methods for Recursive Structural Equation Models," *Journal of Econometrics*, 134, 471–506. [1094]

MAMMEN, E. (1991): "Nonparametric Regression Under Qualitative Smoothness Assumptions," *The Annals of Statistics*, 19, 741–759. [1094,1095,1114]

MATZKIN, R. L. (1994): "Restrictions of Economic Theory in Nonparametric Methods," in *Handbook of Econometrics*, Vol. IV. Amsterdam: North-Holland. [1094,1115]

MILNOR, J. (1965): *Topology From the Differential Viewpoint*. Princeton, NJ: Princeton University Press. [1099]

MOSSINO, J., AND R. TEMAM (1981): "Directional Derivative of the Increasing Rearrangement Mapping and Application to a Queer Differential Equation in Plasma Physics," *Duke Mathematical Journal*, 48, 475–495. [1104]

PORTNOY, S. (1991): "Asymptotic Behavior of Regression Quantiles in Nonstationary, Dependent Cases," *Journal of Multivariate Analysis*, 38, 100–113. [1106]

POWELL, J. L. (1986): "Censored Regression Quantiles," *Journal of Econometrics*, 32, 143–155. [1094,1097]

PRATT, J. W. (1960): "On Interchanging Limits and Integrals," *The Annals of Mathematical Statistics*, 31, 74–77. [1120]

R DEVELOPMENT CORE TEAM (2007): *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at http://www.R-project.org. [1111]

RESNICK, S. I. (1987): *Extreme Values, Regular Variation, and Point Processes*. Applied Probability: A Series of the Applied Probability Trust, Vol. 4. New York: Springer-Verlag. [1117]

ROBERTSON, T., F. T. WRIGHT, AND R. L. DYKSTRA (1988): *Order Restricted Statistical Inference*. Chichester: Wiley. [1105]

VAN DER VAART, A. (1998): *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics, Vol. 3. Cambridge. [1110,1122]

VAN DER VAART, A., AND J. WELLNER (1996): *Weak Convergence and Empirical Processes: With Applications to Statistics*. New York: Springer. [1117]

VILLANI, C. (2003): *Topics in Optimal Transportation*. Providence: American Mathematical Society. [1095,1096]

YU, K., AND M. C. JONES (1998): "Local Linear Quantile Regression," *Journal of the American Statistical Association*, 93, 228–237. [1097]

*Dept. of Economics, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, U.S.A.; vchern@mit.edu,*

*Dept. of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, U.S.A.; ivanf@bu.edu,*

*and*

*Dept. of Economics, Ecole Polytechnique, 91128 Palaiseau, France; alfred.galichon@polytechnique.edu.*

*Manuscript received April, 2008; final revision received November, 2009.*