QUANTILE AND PROBABILITY CURVES WITHOUT CROSSING

VICTOR CHERNOZHUKOV[†] IVÁN FERNÁNDEZ-VAL[§] ALFRED GALICHON[‡]

Abstract. The most common approach to estimating conditional quantile curves is to fit a curve, typically linear, pointwise for each quantile. Linear functional forms, coupled with pointwise fitting, are used for a number of reasons including parsimony of the resulting approximations and good computational properties. The resulting fits, however, may not respect a logical monotonicity requirement – that the quantile curve be increasing as a function of probability. This paper studies the natural monotonization of these empirical curves induced by sampling from the estimated non-monotone model, and then taking the resulting conditional quantile curves that by construction are monotone in the probability. This construction of monotone quantile curves may be seen as a bootstrap and also as a monotonic rearrangement of the original non-monotone function. It is shown that the monotonized curves are *closer* to the true curves in finite samples, for *any* sample size. Under correct specification, the rearranged conditional quantile curves have the same asymptotic distribution as the original non-monotone curves. Under misspecification, however, the asymptotics of the rearranged curves may partially differ from the asymptotics of the original non-monotone curves. An analogous procedure is developed to monotonize the estimates of conditional distribution functions. The results are derived by establishing the compact (Hadamard) differentiability of the monotonized quantile and probability curves with respect to the original curves in discontinuous directions, tangentially to a set of continuous functions. In doing so, the compact differentiability of the rearrangement-related operators is established.

Keywords: Quantile regression, Monotonicity, Rearrangement, Approximation, Functional Delta Method, Hadamard Differentiability of Rearrangement Operators.

AMS 2000 subject classification: Primary 62J02; Secondary 62E20, 62P20

Date: First version is of April 6, 2005. The last modification was done August 13, 2007. The title of this paper is (partially) borrowed from the work of Xuming He (1997), to whom we are grateful for the inspiration and formulation of the problem. We would like to thank Josh Angrist, Andrew Chesher, Phil Cross, Raymond Guiteras, Xuming He, Roger Koenker, Vadim Marmer, Ilya Molchanov, Francesca Molinari, Whitney Newey, Steve Portnoy, Shinichi Sakata, Art Shneyerov, Alp Simsek, and seminar participants at BU, Columbia, Cornell, Georgetown, Harvard-MIT, MIT, Northwestern, UBC, UCL, and UIUC for very useful comments that helped improve the paper.

1. Introduction and Discussion

The problem studied in this paper can be best described using linear quantile regression as the prime example (Koenker, 2005). Suppose that $x'\beta(u)$ is a linear approximation to the u-quantile $Q_0(u|x)$ of a real response variable Y, given a vector of regressors $X = x$. The typical estimation methods fit the conditional curve $x'\beta(u)$ pointwise in $u \in (0,1)$ producing an estimate $x'\widehat{\beta}(u)$. Linear functional forms, coupled with pointwise fitting, are used for a number of reasons including parsimony of the resulting approximations and good computational properties (Portnoy and Koenker, 1997). However, a problem that might occur is that the map

$$u \mapsto x'\widehat{\beta}(u)$$

may not be increasing in u, which violates the logical monotonicity requirement. Another manifestation of this issue, known as the "quantile crossing problem" (He, 1997), is that the conditional quantile curves $x \mapsto x'\beta(u)$ may cross for different values of u.

In the analysis we shall distinguish the following two cases, each leading to the lack of monotonicity or the crossing problem:

(1) Monotonically correct case: The population curve $u \mapsto x'\beta(u)$ is increasing in u, and thus satisfies the monotonicity requirement. However, the empirical curve $u \mapsto x'\widehat{\beta}(u)$ may be non-monotone due to estimation error.
(2) Monotonically incorrect case: The population curve $u \mapsto x'\beta(u)$ is non-monotone due to imperfect approximation to the true conditional quantile function. Accordingly, the resulting empirical curve $u \mapsto x'\widehat{\beta}(u)$ is also non-monotone due to both non-monotonicity of the population curve and estimation error.

Consider the random variable

$$Y_x := x'\widehat{\beta}(U) \text{ where } U \sim U(0,1).$$

This variable can be seen as a bootstrap draw from the estimated quantile regression model, as in Koenker (1994), and has the distribution function

$$\widehat{F}(y|x) = \int_0^1 1\{x'\widehat{\beta}(u) \leq y\}du. \tag{1.1}$$

Moreover, inverting the distribution function, one obtains a proper quantile function

$$\widehat{F}^{-1}(u|x) = \inf\{y : \widehat{F}(y|x) \geq u\}, \tag{1.2}$$

which is monotone in u. The rearranged quantile function $\widehat{F}^{-1}(u|x)$ coincides with the original curve $x'\widehat{\beta}(u)$ if the original curve is increasing in u, but differs from the original curve otherwise. Thus, starting with a possibly non-monotone original curve $u \mapsto x'\widehat{\beta}(u)$, the rearrangement (1.1)-(1.2) produces a monotone quantile curve $u \mapsto \widehat{F}^{-1}(u|x)$. In what follows, we focus our attention on the interval $(0, 1)$ without loss of generality. Indeed, any closed subinterval of $(0, 1)$ can also be considered isomorphically to the treatment of the unit interval case, as commented further in Section 2.

As mentioned above, this rearrangement mechanism has a direct relation to the quantile regression bootstrap (Koenker, 1994), since the rearranged quantile curve is produced by sampling from the estimated original quantile model. Moreover, both the mechanism and its name have a direct relation to rearrangement maps in variational analysis and operations research (e.g., Hardy, Littlewood, and Polya, 1952, and Villani, 2003). Further important references on the rearrangement method are discussed below.

The purpose of this paper is to establish the empirical properties of the rearranged quantile curves and their distribution counterparts:

$$u \mapsto \widehat{F}^{-1}(u|x) \ \text{ and } \ y \mapsto \widehat{F}(y|x),$$

under scenarios (1) and (2). The paper also characterizes certain analytical and approximation properties of the corresponding population curves:

$$u \mapsto F^{-1}(u|x) = \inf\{y : F(y|x) \geq u\}, \ \text{ and } y \mapsto F(y|x) = \int_0^1 1\{x'\beta(u) \leq y\}du.$$

The first main result of the paper establishes the improved estimation properties of the rearranged curves. We show that the rearranged curve $\widehat{F}^{-1}(u|x)$ is closer to the true conditional quantile curve $Q_0(u|x)$ than the original curve. Formally, for each x, we have that for all $p \in [1, \infty]$

$$\left(\int_{\mathcal{U}} |Q_0(u|x) - \widehat{F}^{-1}(u|x)|^p du\right)^{1/p} \leq \left(\int_{\mathcal{U}} |Q_0(u|x) - x'\widehat{\beta}(u)|^p du\right)^{1/p},$$

where the inequality is strict for $p \in (1, \infty)$ whenever $u \mapsto x'\widehat{\beta}(u)$ is decreasing on a subset of $\mathcal{U} := (0,1)$ of positive Lebesgue measure, while $u \mapsto Q_0(u|x)$ is strictly increasing. This property is independent of the sample size, and thus continues to hold in the population, and also regardless of whether the linear quantile estimator $x'\widehat{\beta}(u)$ estimates $Q_0(u|x)$ consistently or not, i.e. whether $Q_0(u|x) = x'\beta(u)$ or $Q_0(u|x) \neq x'\beta(u)$. In other words, the rearranged quantile curves have smaller estimation error than the original curves whenever the latter are not monotone. This is a very important property that does not depend on the way the quantile model is estimated. It also does not rely on any other specifics of the current context and is therefore applicable quite generally.

Towards describing the essence of the rest of results, let us fix the value of the regressor X to x. Suppose that $\widehat{\beta}(u)$ is an estimator for $\beta(u)$ that converges weakly to a Gaussian process $G(u)$, so that

$$\sqrt{n}x'(\widehat{\beta}(u) - \beta(u)) \Rightarrow x'G(u), \tag{1.3}$$

as a stochastic process indexed by u in the metric space of bounded functions $\ell^\infty(0,1)$. For sufficient conditions, see, for example, Gutenbrunner and Jurečková (1992), Portnoy (1991), and Angrist, Chernozhukov, and Fernandez-Val (2006).

The second main result of the paper is that in the monotonically correct case (1),

$$\sqrt{n}(\widehat{F}(y|x) - F(y|x)) \Rightarrow F'(y|x)[x'G(F(y|x))], \tag{1.4}$$

as a stochastic process indexed by y in the metric space $\ell^\infty(\mathcal{Y})$, where $\mathcal{Y}$ is the support of Y_x; and

$$F'(y|x) = \frac{1}{x'\beta'(u)}\Big|_{u=F(y|x)}, \quad \text{with} \quad \beta'(u) := \frac{\partial \beta(u)}{\partial u}.$$

Moreover, we show that

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow x'G(u), \tag{1.5}$$

as a stochastic process indexed by u, in $\ell^\infty(0,1)$; which, remarkably, coincides with the first order asymptotics (1.3) of the original curve. This result has a convenient practical implication: if the population curve is monotone, then the empirical non-monotone curve can be re-arranged to be monotonic *without* affecting its (first order) asymptotic properties. To derive the above results we find the functional Hadamard derivatives of $F(y|x)$ and $F^{-1}(u|x)$ with respect to perturbations of the underlying curve $x'\beta(u)$ in discontinuous directions, tangentially to the set of continuous functions, and then use the functional delta method. Establishing the Hadamard differentiability of the rearranged distribution and quantile curves in discontinuous directions is the second main theoretical result of the paper.

The third main result is that in the monotonically incorrect case

$$\sqrt{n}(\widehat{F}(y|x) - F(y|x)) \Rightarrow \sum_{k=1}^{K(y|x)} \frac{x'G(u_k(y|x))}{|x'\beta'(u_k(y|x))|}, \tag{1.6}$$

as a stochastic process indexed by $y \in K$, in $\ell^\infty(K)$, where K is an appropriate set defined in the next section. Here $u_1(y|x) < ... < u_{K(y|x)}(y|x)$ are solutions to the equation $y = x'\beta(u)$, assuming that $K(y|x)$ is bounded. Similarly, for the rearranged quantile curve,

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow \left(\sum_{k=1}^{K(y|x)} \frac{1}{|x'\beta'(u_k(y|x))|} \right)^{-1} \sum_{k=1}^{K(y|x)} \frac{x'G(u_k(y|x))}{|x'\beta'(u_k(y|x))|} \Bigg|_{y=F^{-1}(u|x)}, \tag{1.7}$$

as a stochastic process indexed by $u \in K'$, in $\ell^\infty(K')$, where K' is an appropriate set defined in the next section.

Analogously to quantiles, most estimation methods for conditional distribution functions do not impose monotonicity, and therefore can give rise to non-monotonic empirical conditional distribution curves; see, for example, Hall, Wolff, and Yao (1999). A similar monotone rearrangement can be applied to these distribution curves by exchanging the roles played by the quantile and the probability spaces. Thus, suppose that $\widehat{P}(y|x)$ is a candidate estimate of a conditional distribution function, which is not monotone in y. The rearranged monotone quantile curve associated with $\widehat{P}(y|x)$ is

$$\widehat{Q}\,(u|x) = \int_0^\infty 1\{\widehat{P}\,(y|x) \leq u\}dy - \int_{-\infty}^0 1\{\widehat{P}\,(y|x) > u\}dy.$$

The rearranged probability curve can then be obtained as the inverse of this quantile curve, i.e.,

$$\widehat{F}(y|x) = \inf\left\{u : \widehat{Q}\,(u|x) \geq y\right\},$$

which is monotone by construction. Section 3 shows in more detail that similar improved estimation properties and an asymptotic distribution theory goes through for $\widehat{Q}$ and $\widehat{F}$.

The distributional results in the paper do not rely on the sampling properties of the particular estimation method used, because they are expressed in terms of the differentiability of the operator with respect to the basic estimated process. Moreover, the results that follow are derived without imposing linearity of the functional forms. The only conditions required are that (1) a central limit theorem like (1.3) applies to the estimator of the curve, and (2) the population curves have some smoothness properties. The exact nature of these population curves does not affect the validity of the results. For example, the results hold regardless of whether the underlying model is an ordinary or an instrumental quantile regression model.

There exist other methods to obtain monotonic fits based on quantile regression. He (1997), for example, proposes to impose a location-scale regression model, which naturally satisfies monotonicity. This approach is fruitful for location-scale situations, but in numerous cases data do not satisfy the location-scale model, as discussed, for example, in Lehmann (1974), Doksum (1974), and Koenker (2005). Koenker and Ng (2005) develop a computational method for quantile regression that imposes the non-crossing constraints in simultaneous fitting of quantile curves. This approach may be fruitful in many situations, but the statistical properties of the method remain unknown. Clearly, Koenker and Ng's proposal is different from the rearrangement method.

The distributional results obtained in the paper can also be viewed as a functional delta method for the rearrangement-related operators (1.1) and (1.2) that include the inverse (quantile) operators as a special case. In this sense, they extend the previous results by Gill and Johansen (1990), Doss and Gill (1992), and Dudley and Norvaisa (1999) on compact differentiability of the quantile operator. The main technical difficulty here, as well as in the quantile case, is that differentiability needs to be established in discontinuous directions (that converge to continuous directions, i.e., tangentially to the set of continuous functions), because the empirical perturbations of the quantile processes are typically step functions.

Both the statistical and mathematical results of this paper complement the important work of Dette, Neumeyer, and Pilz (2006), which applies the rearrangement operators to kernel mean regressions, in order to obtain mean regression functions that are monotonic in the regressors. Our results on Hadamard differentiability in discontinuous directions are new. They complement the local expansions in smooth directions subsumed in the proofs in Dette *et. al.* for the case called here the monotonically correct case. In addition, our results cover monotonically incorrect cases. The statistical problem also differs quite substantially. The mathematical results of this paper also complement the results on directional differentiability of L_1- functionals of rearranged functions like (1.2) by Mossino and Temam (1981). The results for L_1-functionals do not imply the main results of this paper, such as (1.4)-(1.7), but the converse is shown to be true. (See discussion after Proposition 4 in Section 2 for more details.)

There are many potential applications of the estimation and differentiability results of the paper to objects other than probability or quantile curves. For example, in a companion work we present applications to economic demand and production functions and to biometric growth curves, where monotonization is used to impose useful theoretical or logical restrictions (Chernozhukov, Fernández-Val, and Galichon, 2006a).

We organize the rest of the paper as follows. In Section 2.1 we describe some basic analytical properties of the rearranged population curves. In Section 2.2 we derive the functional differentiability results. In Section 2.3 we present estimation properties of the rearranged curves and establish their limit distributions. In Section 3 we extend the previous results to monotonize estimates of distribution curves. In Section 4.1 we illustrate the rearrangement procedure with an empirical application, and in Section 4.2 we provide a Monte-Carlo example. In Section 5 we conclude with a summary of the main results.

2. Rearranged Quantile Curves: Analytical and Empirical Properties

In this section the treatment of the problem is somewhat more general than in the introduction. In particular, we replace the linear functional form $x'\beta(u)$ by $Q(u|x)$. Define $Y_x := Q(U|x)$, where $U \sim \text{Uniform}(\mathcal{U})$ with $\mathcal{U} = (0,1)$. Let $F(y|x) := \int_0^1 1\{Q(u|x) \leq y\}du$ be the distribution function of Y_x, and $F^{-1}(u|x) := \inf\{y : F(y|x) \geq u\}$ be the quantile function of Y_x.

Remark. We consider the interval $(0,1)$ without loss of generality. Indeed, suppose we are interested in a particular subinterval $(a, a+b)$ of $(0,1)$. For example, we may wish to focus estimation on a particular range of quantiles or to avoid estimation of tail quantiles. For this purpose, we define all objects conditionally on the event $U \in (a, a+b)$: $\tilde{Y}_x := \tilde{Q}(\tilde{U}|x) = Q(a + b\tilde{U}|x)$, where $\tilde{U} \sim U(0,1)$, $\tilde{F}(\tilde{y}|x) := \int_0^1 1\{\tilde{Q}(\tilde{u}|x) \leq \tilde{y}\}d\tilde{u}$, and $\tilde{F}^{-1}(\tilde{u}|x) := \inf\{y : F(y|x) \geq \tilde{u}\}$ for $\tilde{u} \in (0,1)$. The analysis of the paper applies to the functions $\tilde{Q}$ and $\tilde{F}^{-1}$. In order to go back to the unconditional quantities, we can use the transformations $Q(u|x) = \tilde{Q}((u-a)/b|x)$ for $u \in (a, a+b)$ and $F(y|x) = a + b\tilde{F}(y|x)$ for $y \in \{Q(u|x) : u \in (a, a+b)\}$.

2.1. **Basic Analytical Properties.** We start by developing some basic properties for $F(y|x)$ and $F^{-1}(u|x)$, the population counterparts of the rearranged distribution curve and its inverse. We need these properties to derive various empirical properties stated in the next section.

Recall first the following definitions from Milnor (1965): Let $g : \mathcal{U} \subset \mathbb{R} \to \mathbb{R}$ 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, i.e., $g'(u) \neq 0$. 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, i.e., if $\forall u \in g^{-1}(\{y\})$, $g'(u) \neq 0$. A value y which is not a regular value is called a critical value.

Denote by $\mathcal{Y}_x$ the support of Y_x, $\mathcal{YX} := \{(y,x) : y \in \mathcal{Y}_x, x \in \mathcal{X}\}$, and $\mathcal{UX} := \mathcal{U} \times \mathcal{X}$. We assume throughout that $\mathcal{Y}_x \subset \mathcal{Y}$, which is 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, the set $\mathcal{X}$ is taken to be a singleton $\mathcal{X} = \{x\}$. We make the following assumptions about $Q(u|x)$:

(a) $Q(u|x) : \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ is a continuously differentiable function in both arguments,
(b) For each $x \in \mathcal{X}$, the number of elements of $\{u \in \mathcal{U} \mid Q'(u|x) = 0\}$ is finite and uniformly bounded on $x \in \mathcal{X}$.

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

Proposition 1 (Basic Properties of $F(y|x)$ and $F^{-1}(u|x)$). *Under assumptions (a) - (b), the functions $F(y|x)$ and $F^{-1}(u|x)$ satisfy the following properties:*

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

2. For any $y \in \mathcal{Y}_x^*$

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

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

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

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

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

4. The quantile function $F^{-1}(u|x)$ partially coincides with $Q(u|x)$; namely

$$F^{-1}(u|x) = Q(u|x),$$

provided that $Q(u|x)$ is increasing at u, and the equation $Q(u|x) = y$ has unique solution for $y = F^{-1}(u|x)$.

5. The quantile function $F^{-1}(u|x)$ is equivariant to location and scale transformations of $Q(u|x)$.

6. The quantile function $F^{-1}(u|x)$ has an ordinary continuous derivative

$$1/f(F^{-1}(u|x)|x),$$

when $F^{-1}(u|x) \in \mathcal{Y}_x^*$. This function is also a Radon-Nikodym derivative with respect to the the Lebesgue measure.

7. The map $(y, x) \mapsto F(y|x)$ is continuous on $\mathcal{YX}$ and the map $(u, x) \mapsto F^{-1}(u|x)$ is continuous on $\mathcal{UX}$.

The following simple example illustrates some of these basic properties in a situation where the initial population pseudo-quantile curve is highly non-monotone. Consider the following pseudo-quantile function:

$$Q(u) = 5\left(u + \frac{1}{\pi}\sin(2\pi u)\right). \tag{2.1}$$

The left panel of Figure 1 shows that this function is non-monotone in $[0, 1]$. In particular, the slope of $Q(u)$ changes sign twice at $1/3$ and $2/3$. The rearranged quantile curve $F^{-1}(u)$, also plotted in this panel, is continuous and monotonically increasing. The results 1, 2, 4 and 7 of the proposition are illustrated in the right panel of Figure 1, which plots the original and rearranged distribution curves. Here we can see that the rearranged distribution function is continuous, does not have mass points, and coincide with the original curve for values of y where the original curve is one to one and increasing.

Figure 2 illustrates the third and sixth results of the proposition by plotting the sparsity function for $F^{-1}(u)$ and the density function of $F(y)$. The derivative of $F(y)$ in the right panel is continuous at the regular values of $Q(u)$. Similarly, the sparsity function for $F^{-1}(u)$ in the left panel is continuous at the corresponding image values (under $F(y)$).

2.2. **Functional Derivatives.** Next, we establish the main results of the paper on Hadamard differentiability of $F(y|x)$ and $F^{-1}(u|x)$ with respect to $Q(u|x)$, tangentially to the space of continuous functions on $\mathcal{UX}$. This differentiability property is important for deriving the asymptotic distributions of the rearranged estimates. In particular, the property allows us to establish generic convergence results for rearranged curves based on any initial quantile estimator, provided the initial estimator satisfies a functional central limit theorem. The property also implies that the bootstrap is valid for performing inference on the rearranged estimates, provided the bootstrap is valid for the initial estimates. This result follows from the functional delta method for the bootstrap (e.g., Theorem 13.9 in van der Vaart, 1998).

In what follows, $\ell^\infty(\mathcal{UX})$ denotes the set of bounded and measurable functions $h : \mathcal{UX} \to \mathbb{R}$, $C(\mathcal{UX})$ denotes the set of continuous functions mapping $h : \mathcal{UX} \to \mathbb{R}$, and

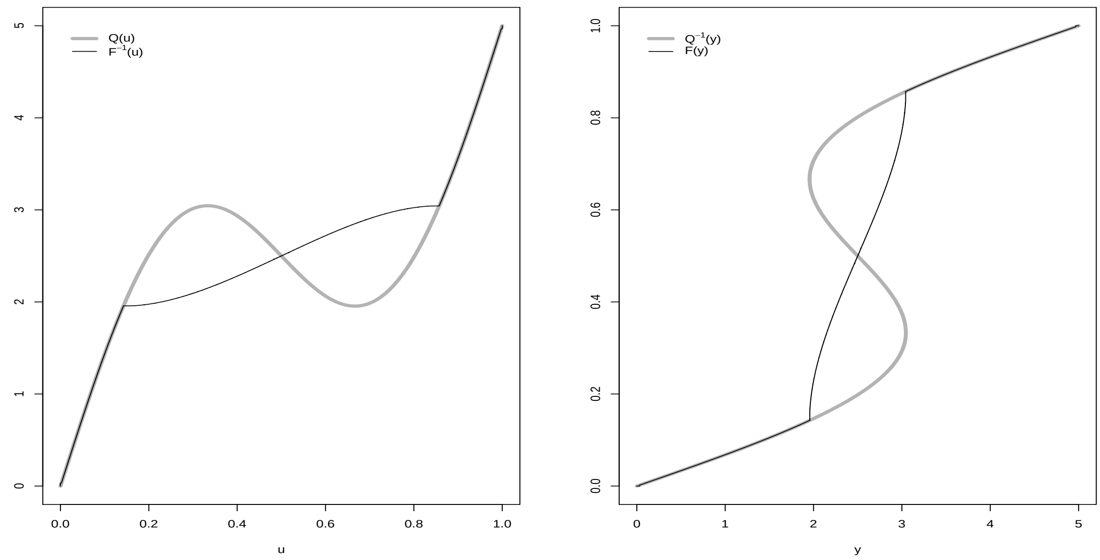

FIGURE 1. Left: The pseudo-quantile function $Q(u)$ and the rearranged quantile function $F^{-1}(u)$. Right: The pseudo-distribution function $Q^{-1}(y)$ and the rearranged distribution function $F(y)$.

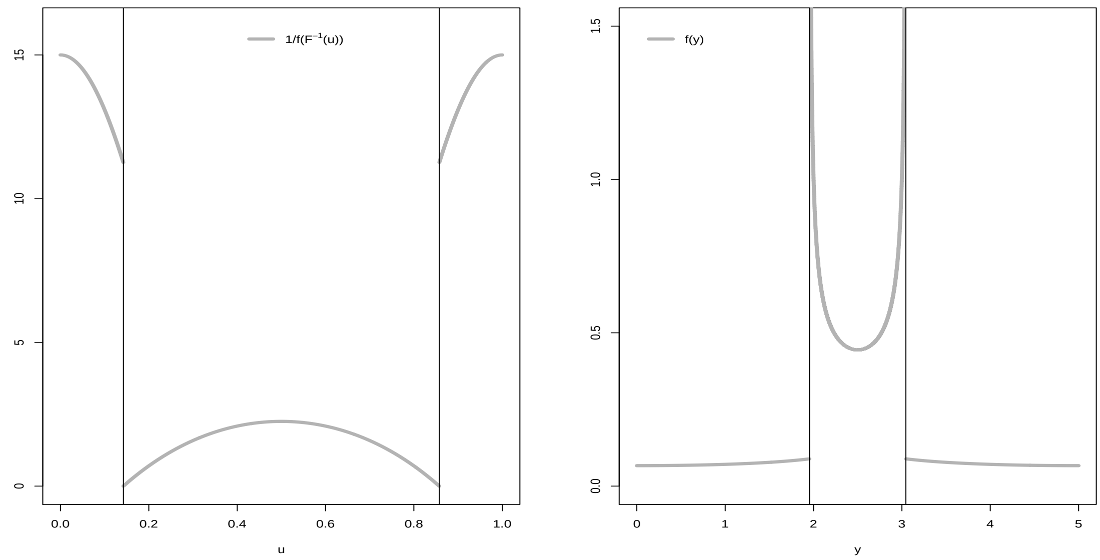

FIGURE 2. Left: The density (sparsity) function of the rearranged quantile function $F^{-1}(u)$. Right: The density function of the rearranged distribution function $F(y)$.

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

Proposition 2 (Hadamard Derivative of $F(y|x)$ with respect to $Q(u|x)$)**.** *Define* $F(y|x, h_t) := \int_0^1 1\{Q(u|x) + th_t(u|x) \le y\}du$. *Under assumptions (a)-(b), as* $t \to 0$,

$$D_{h_t}(y|x,t) = \frac{F(y|x,h_t) - F(y|x)}{t} \to D_h(y|x), \tag{2.2}$$

$$D_h(y|x) := -\sum_{k=1}^{K(y|x)} \frac{h(u_k\,(y|x)\,|x)}{|Q'(u_k(y|x)|x)|}. \tag{2.3}$$

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* $h \in C(\mathcal{UX})$.

Proposition 3 (Hadamard Derivative of $F^{-1}(u|x)$ with respect to $Q(u|x)$)**.** *Under assumptions (a)-(b), as* $t \to 0$,

$$\tilde{D}_{h_t}(u|x,t) := \frac{F^{-1}(u|x,h_t) - F^{-1}(u|x)}{t} \to \tilde{D}_h(u|x), \tag{2.4}$$

$$\tilde{D}_h(u|x) := -\frac{1}{f(F^{-1}(u|x)|x)} \cdot D_h(F^{-1}(u|x)|x). \tag{2.5}$$

The convergence holds uniformly in any compact subset of $\mathcal{UX}^* = \{(u,x) : (F^{-1}(u|x), x) \in \mathcal{YX}^*\}$, *for every* $|h_t - h|_\infty \to 0$, *where* $h_t \in \ell^\infty(\mathcal{UX})$, *and* $h \in C(\mathcal{UX})$.

The convergence results hold uniformly on regions that exclude the critical values of the mapping $u \mapsto Q(u|x)$. At the critical values, $Q(u|x)$ possibly changes from increasing to decreasing. Moreover, in the monotonically correct case (1), the following result is worth emphasizing:

Corollary 1 (Monotonically correct case)**.** *Suppose* $u \mapsto Q(u|x)$ *has* $Q'(u|x) > 0$, *for each* $(u,x) \in \mathcal{UX}$, *then* $\mathcal{YX}^* = \mathcal{YX}$ *and* $\mathcal{UX}^* = \mathcal{UX}$. *Therefore, the convergence in Propositions 2 and 3 holds uniformly over the entire* $\mathcal{YX}$ *and* $\mathcal{UX}$, *respectively. Moreover,* $\tilde{D}_h(u|x) = h$, *i.e., the Hadamard derivative of the rearranged quantile with respect to the original curve is the identity operator.*

The convergence is uniform over the entire domain in the monotonically correct case. This result raises naturally the question of whether uniform convergence can be achieved by some operation of smoothing in the monotonically incorrect case – namely integrating either over y (or over u). The answer is indeed yes.

The following proposition calculates the Hadamard derivative of the following functionals obtained by integration:

$$(y',x) \mapsto \int_{\mathcal{Y}} 1\{y \leq y'\} g(y|x) F(y|x) dy, \quad (u',x) \mapsto \int_{\mathcal{U}} 1\{u \leq u'\} g(u|x) F^{-1}(u|x) du,$$

with the restrictions on g specified below. These elementary functionals are useful building blocks for various statistics, as briefly mentioned in the next section.

Proposition 4. *The following results are true with the limits being continuous on the specified domains:*

$$1. \quad \int_{\mathcal{Y}} 1\{y \leq y'\} g(y|x) D_{h_t}(y|x,t) dy \to \int_{\mathcal{Y}} 1\{y \leq y'\} g(y|x) D_h(y|x) dy$$

uniformly in $(y',x) \in \mathcal{Y}\mathcal{X}$, *for any* $g \in \ell^\infty(\mathcal{Y}\mathcal{X})$ *such that* $x \mapsto g(y|x)$ *is continuous for a.e.* y.

$$2. \quad \int_{\mathcal{U}} 1\{u \leq u'\} g(u|x) \tilde{D}_{h_t}(u|x,t) du \to \int_{\mathcal{U}} 1\{u \leq u'\} g(u|x) \tilde{D}_h(u|x) du$$

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

This proposition essentially is a corollary of Propositions 2 and 3. Indeed, the results (1)-(2) follow from the fact that the pointwise convergence of Propositions 2 and 3, coupled with the uniform integrability shown in Lemma 3 in the Appendix, permits the interchange of limits and integrals. An alternative way of proving result (2), but *not* any other result in the paper, can be based on exploiting the convexity of the functional in (2) 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. However, details of this approach are described in Chernozhukov, Fernandez-Val, and Galichon (2006b) with an application to some nonparametric estimation problems.

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 as

$$Sf(y) := \int k_\delta(y - y')f(y')dy', \tag{2.6}$$

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

$$SF(y|x) := \int k_\delta(y - y')F(y'|x)dy', \quad SF^{-1}(u|x) := \int k_\delta(u - u')F^{-1}(u'|x)du'.$$

Since these curves are merely formed as differences of the elementary functionals in Proposition 4, followed by a division by δ, the following corollary is immediate.

Corollary 2. *We have that $SD_{h_t}(y|x,t) \to SD_h(y|x)$ uniformly in $(y, x) \in \mathcal{YX}$, and $S\tilde{D}_{h_t}(u|x,t) \to S\tilde{D}_h(u|x)$ uniformly in $(u, x) \in \mathcal{UX}$.*

Note that smoothing accomplishes uniform convergence over the entire domain, which is a good property to have from the perspective of data analysis.

2.3. **Empirical Properties of $\widehat{F}(y|x)$ and $\widehat{F}^{-1}(u|x)$.** We are now ready to state the main results for this section.

Proposition 5 (Improvement in Estimation Property Provided by Rearrangement)**.** *Suppose that $\widehat{Q}(\cdot|\cdot)$ is an estimator (not necessarily consistent) for some true quantile curve $Q_0(\cdot|\cdot)$. Then, the rearranged curve $\widehat{F}^{-1}(u|x)$ is closer to the true curve than $\widehat{Q}(u|x)$ in the sense that, for each $x \in \mathcal{X}$,*

$$\left(\int_{\mathcal{U}} |Q_0(u|x) - \widehat{F}^{-1}(u|x)|^p du\right)^{1/p} \leq \left(\int_{\mathcal{U}} |Q_0(u|x) - \widehat{Q}(u|x)|^p du\right)^{1/p}, \ p \in [1, \infty],$$

where the inequality is strict for $p \in (1, \infty)$ whenever $\widehat{Q}(u|x)$ is decreasing on a subset of $\mathcal{U}$ of positive Lebesgue measure, while $Q_0(u|x)$ is 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 proposition establishes that the rearranged quantile curves have smaller estimation error than the original curves whenever the latter are not monotone. This is a very important property that does not depend on the way the quantile model is estimated. It also does not rely on any other specifics and is thus applicable quite generally.

The following proposition investigates the asymptotic distributions of the rearranged curves.

Proposition 6 (Empirical Properties of $(y,x) \mapsto \widehat{F}(y|x)$ and $(u,x) \mapsto \widehat{F}^{-1}(u|x)$)**.** *Suppose that $\widehat{Q}(\cdot|\cdot)$ is an estimator for $Q(\cdot|\cdot)$ that takes its values in the space of bounded measurable functions defined on $\mathcal{UX}$, and that, in $\ell^\infty(\mathcal{UX})$,*

$$\sqrt{n}(\widehat{Q}(u|x) - Q(u|x)) \Rightarrow G(u|x),$$

as a stochastic process indexed by $(u,x) \in \mathcal{UX}$, where $(u,x) \mapsto G(u|x)$ is a Gaussian process with continuous paths. Assume also that $Q(u|x)$ satisfies the basic conditions (a) and (b). Then in $\ell^\infty(K)$, where K is any compact subset of $\mathcal{YX}^*$,

$$\sqrt{n}(\widehat{F}(y|x) - F(y|x)) \Rightarrow D_G(y|x)$$

as a stochastic process indexed by $(y,x) \in \mathcal{YX}^*$; and in $\ell^\infty(\mathcal{UX}_K)$, with $\mathcal{UX}_K = \{(u,x) : (F^{-1}(u|x), x) \in K\}$,

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow \widetilde{D}_G(u|x),$$

as a stochastic process indexed by $(u,x) \in \mathcal{UX}_K$.

Corollary 3 (Monotonically correct case)**.** *Suppose $u \mapsto Q(u|x)$ has $Q'(u|x) > 0$ 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, $\widetilde{D}_G(u|x) = G(u|x)$, i.e., the rearranged quantile curves have the same first order asymptotic distribution as the original quantile curves.*

Thus, in the monotonically correct case, the first order properties of the rearranged and initial quantile estimates coincide. Hence, all the inference tools that apply to original quantile estimates also apply to the rearranged quantile estimates. In particular, if the bootstrap is valid for the original estimate, it is also valid for the rearranged estimate, by the functional delta method for the bootstrap. In the empirical example of Section 4, we exploit this useful property to construct uniform confidence bands for the conditional quantile functions based on the rearranged quantile function estimates.

In addition to the results on quantile function estimates, Proposition 6 provides the asymptotic properties of the distribution function estimates. The preceding remark about the validity of bootstrap applies also to these estimates.

In the monotonically incorrect case, the large sample properties of the rearranged quantile estimates differ from those of the initial quantile estimates. Proposition 6 enables us to perform inferences for rearranged curves in this case, including by the bootstrap, but only after excluding certain nonregular neighborhoods (for the distribution estimates, the neighborhood of the critical values of the map $u \mapsto Q(u|x)$, and, for the rearranged quantile estimates, the image of the latter neighborhood under $F(y|x)$). However, if we consider the following linear functionals of the rearranged quantile and distribution estimates:

$$(y', x) \mapsto \int_{\mathcal{Y}} 1\{y \leq y'\} g(y|x) \widehat{F}(y|x) dy, \quad (u', x) \mapsto \int_{\mathcal{U}} 1\{u \leq u'\} g(u|x) \widehat{F}^{-1}(u|x) du,$$

then we no longer need to exclude the nonregular neighborhoods. The following proposition describes the empirical properties of these functionals in large samples.

Proposition 7 (Empirical Properties of Integrated Curves)**.** *Under the conditions of Proposition 6, the following results are true with the limits being continuous on the specified domains:*

1. $\sqrt{n} \int_{\mathcal{Y}} 1\{y \leq y'\} g(y|x)(\widehat{F}(y|x) - F(y|x)) dy \Rightarrow \int_{\mathcal{Y}} 1\{y \leq y'\} g(y|x) D_G(y|x) dy,$

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

2. $\sqrt{n} \int_{\mathcal{U}} 1\{u \leq u'\} g(u|x)(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) du \Rightarrow \int_{\mathcal{U}} 1\{u \leq u'\} g(u|x) \tilde{D}_G(u|x) du,$

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

The restrictions on the function g *are the same as in Proposition 4.*

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 is

$$\hat{L}(u|x) = \Big(\int_{\mathcal{U}} 1\{t \leq u\} \hat{F}^{-1}(t|x) dt \Big) / \Big(\int_{\mathcal{U}} \hat{F}^{-1}(t|x) dt \Big),$$

which is a ratio of a partial mean to the mean. Hadamard differentiability of these statistics with respect to the underlying $Q(u|x)$ immediately follows from the Hadamard differentiability of the elementary functionals of Proposition 7 by means of the chain rule. Therefore, the asymptotic distribution of these statistics can be determined from the asymptotic distribution of the linear functionals, by the functional delta method. In particular, the validity of the bootstrap for these functionals is preserved 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 (2.6) to $\widehat{F}(y'|x)$ and $\widehat{F}^{-1}(u|x)$:

$$S\widehat{F}(y|x) := \int k_\delta(y-y')\widehat{F}(y'|x)dy', \quad S\widehat{F}^{-1}(u|x) := \int k_\delta(u-u')\widehat{F}^{-1}(u'|x)du'.$$

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

Corollary 4 (Large Sample Properties of Smoothed Curves)**.** *Under the conditions of Proposition 6, in* $\ell^\infty(\mathcal{YX})$*,*

$$\sqrt{n}(S\widehat{F}(y|x) - SF(y|x)) \Rightarrow S[D_G(y|x)],$$

as a stochastic process indexed by $(y,x) \in \mathcal{YX}$*, and in* $\ell^\infty(\mathcal{UX})$*,*

$$\sqrt{n}(S\widehat{F}^{-1}(u|x) - SF^{-1}(u|x)) \Rightarrow S[\widetilde{D}_G(u|x)],$$

as a stochastic process indexed by $(u,x) \in \mathcal{UX}$*.*

Thus, inference on the smoothed rearranged estimates can be performed without excluding nonregular neighborhoods, which is convenient for practice. Furthermore, validity of the bootstrap for the smoothed curves follows by the functional delta method for the bootstrap.

3. Theory of Rearranged Distribution Curves

The rearrangement method can also be applied to rearrange cumulative distribution curves monotonically by exchanging the roles of the quantile and probability spaces.

There are several situations where one might be faced with the problem of non-increasing empirical distribution curves. In an option pricing context, for example, Ait-Sahalia and Duarte (2003) use market data to estimate a risk-neutral distribution. Estimation error may cause the resulting distribution function to be non-monotonic. In other cases the distribution curve is obtained by some inverse transformation and local non-monotonicity comes as an artefact of the regularization technique. In other situations the particular estimation technique may not respect monotonicity (see, e.g., Hall, Wolff, and Yao, 1999). We present an alternative solution to this problem that uses the rearrangement method.

Here we do not present the conditional case for notational convenience. All derivations for conditional distributions, however, are exactly parallel to those presented in this section. Suppose we have $y \mapsto \widehat{P}(y)$ as a candidate empirical probability distribution curve, which does not necessarily satisfy monotonicity, with population counterpart $P(y)$. Define the following quantile curve

$$\widehat{Q}(u) = \int_0^\infty 1\{\widehat{P}\,(y) < u\}dy - \int_{-\infty}^0 1\{\widehat{P}\,(y) > u\}dy,$$

which is monotone. In what follows, we further assume that the support of $P(y)$ is $\mathcal{Y} \subset [0, +\infty)$, so that the second term drops out (otherwise it can be treated analogously to the first).

The inverse of the quantile curve is the rearranged probability curve

$$\widehat{F}(y) = \inf\left\{y : \widehat{Q}\,(u) \geq y\right\},$$

which is also monotone by construction. It should be clear at this point that the quantities $\widehat{Q}$ and $\widehat{F}$ are exactly symmetric to $\widehat{F}$ and $\widehat{Q}$ in the quantile case.

The following improved approximation property is true for $\widehat{F}$: Let $F_0(y)$ be the true distribution function, then for all $p \in [1, \infty]$,

$$\left(\int_{\mathbb{R}} |F_0(y) - \widehat{F}(y)|^p dy\right)^{1/p} \leq \left(\int_{\mathbb{R}} |F_0(y) - \widehat{P}(y)|^p dy\right)^{1/p},$$

where the inequality is strict for $p \in (1, \infty)$ whenever the integral on the right is finite and $y \mapsto \widehat{P}(y)$ is decreasing on a subset of positive Lebesgue measure, while $F_0(u)$ is strictly increasing. This property is independent of the sample size, and thus continues to hold in the population.

In the monotonically correct case, that is when $P'(Q(u)) > 0$ for all $u \in [0,1]$, if the empirical distribution curve $P(y)$ satisfies

$$\sqrt{n}\left(\widehat{P}(y) - P(y)\right) \Rightarrow G(y)$$

in $\ell^\infty(\mathcal{Y})$, where G is a Gaussian process, then

$$\sqrt{n}\left(\widehat{Q}(u) - Q(u)\right) \Rightarrow \left(\frac{1}{P'(Q(u))}\right) G\left(Q(u)\right)$$

in $\ell^\infty([0,1])$, and, in $\ell^\infty(\mathcal{Y})$,

$$\sqrt{n}\left(\widehat{F}\left(y\right) - F\left(y\right)\right) \Rightarrow G\left(y\right). \tag{3.1}$$

Results paralleling those of the previous section also follow for the monotonically incorrect case. In particular, we have

$$\sqrt{n}\left(\widehat{Q}(u) - Q(u)\right) \Rightarrow \sum_{k=1}^{K(u)} \frac{G\left(y_k\left(u\right)\right)}{|P'(y_k(u))|}$$

in $\ell^\infty([0,1])$, and

$$\sqrt{n}\left(\widehat{F}\left(y\right) - F\left(y\right)\right) \Rightarrow \left(\sum_{k=1}^{K(u)} \frac{1}{|P'(y_k\left(u\right))|}\right)^{-1} \sum_{k=1}^{K(u)} \frac{G\left(y_k\left(u\right)\right)}{|P'(y_k\left(u\right))|}\Bigg|_{u=P(y)} \tag{3.2}$$

in $\ell^\infty(\mathcal{Y})$, where $\{y_k(u),\ \text{for } k = 1, ..., K(u)\}$ are the roots of $P(y) = u$, assuming $K(u)$ is bounded uniformly in u.

4. Illustrative Examples

4.1. **Empirical Example.** To illustrate the practical applicability of the rearrangement method, we consider the estimation of expenditure curves. We use the original Engel (1857) data, from 235 budget surveys of 19th century working-class Belgium households, to estimate the relationship between food expenditure and annual household income (see Koenker, 2005). Ernst Engel originally presented these data to support the hypothesis that food expenditure constitutes a declining share of household income (Engel's Law).

In Figure 3, we show a scatterplot of the Engel data on food expenditure versus household income, along with quantile regression curves with the quantile indices $\{05, 0.1, ..., 0.95\}$. We see that the quantile regression lines become closer and cross at low values of income. This crossing problem of the Engel curves is also evident in Figure 4, in which we plot the quantile regression process of food expenditure as a function of the quantile index. For low values of income, the quantile regression process is clearly non-monotone. The rearrangement procedure fixes the non-monotonicity producing increasing quantile functions. Moreover, the rearranged curves coincide with their quantile regression counterparts for the middle values of income where there is no quantile-crossing problem.

In Figure 5, we plot simultaneous 90% confidence intervals for the conditional quantile function of food expenditure for different values of income (at the sample median, and the 5% percentile of income). We construct the bands using both original quantile regression curves and rearranged quantile curves based on 500 bootstrap repetitions and a grid of quantile indices $\{0.10, 0.11, ..., 0.90\}$. We obtain the bands for the rearranged curves assuming that the population quantile regression curves are monotonically correct, so that the first order behavior of the rearranged curves coincides with the behavior of the original curves. The figure shows that even for the low value of income the rearranged bands lie within the quantile regression bands. This observation points towards the maintained assumption of the monotonically correct case. The lack of monotonicity of the estimated quantile regression process in this case is likely to by caused by sampling error.

We find more evidence consistent with the monotonically correct case in Figure 6, in which we plot the simultaneous confidence bands for the smoothed quantile regression and rearranged curves. We construct the band by bootstrapping the smoothed curves (with bandwidth equal to .05). The bootstrap bands are valid for the smoothed rearranged curves even in the monotonically incorrect case. The almost perfect overlapping between the confidence bands points towards the monotonically correct case. Interestingly, smoothing reduces the width of the confidence bands, but does not completely monotonize the quantile regression curves.

4.2. **Monte Carlo.** We use the following Monte Carlo experiment, matching closely the previous empirical application, to illustrate the estimation properties of the rearranged curves in finite samples. In particular, we consider two designs based on the location-scale shift model: $Y = Z(X)'\alpha + (Z(X)'\gamma)\epsilon$, where ϵ is independent of X, with the true

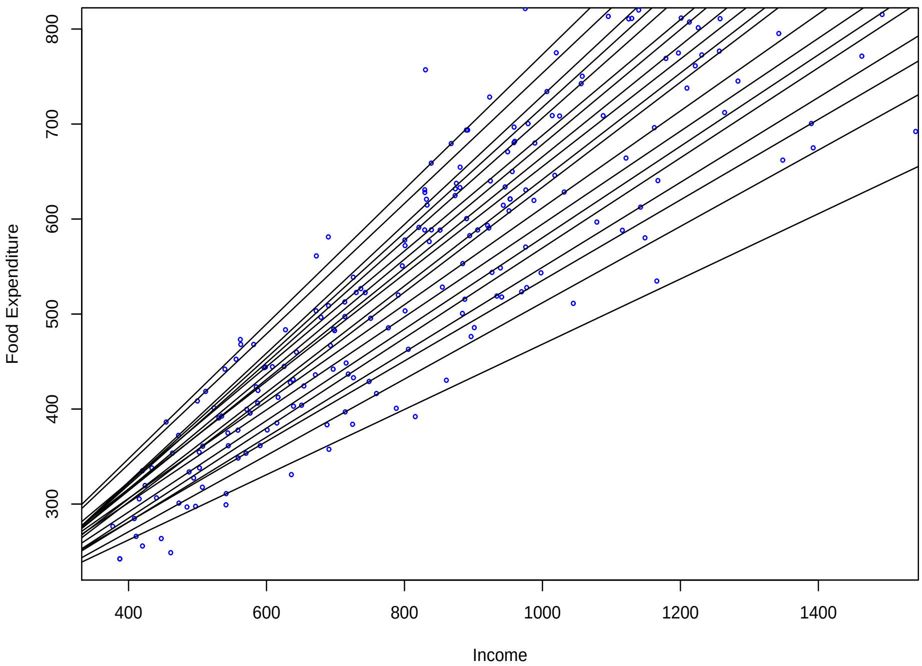

Figure 3. The scatterplot and quantile regression fits of the Engel food expenditure data. The plot shows a scatterplot of the Engel data on food expenditure vs. household income for a sample of 235 19th century working-class Belgium households. Superimposed on the plot are the $\{0.05, 0.10, ..., 0.95\}$ quantile regression curves. The range displayed corresponds to values of income lower than 1500 and values of food expenditure lower than 800.

conditional quantile function

$$Q_0(u|X) = Z(X)'\alpha + (Z(X)'\gamma)Q_\epsilon(u).$$

Design 1 includes a constant and a regressor, namely $Z(X) = (1, X)$; and design 2 has an additional nonlinear regressor, namely, $Z(X) = (1, X, 1\{X > a\} \cdot X)$, where $a = \text{median}(X)$. We select the parameters for designs 1 and 2 to match the Engel empirical example, employing the estimation method of Koenker and Xiao (2002). For design 1 we set $\alpha = (624.15, 0.55)$ and $\gamma = (1, 0.0013)$; and for design 2 we set $\alpha =$

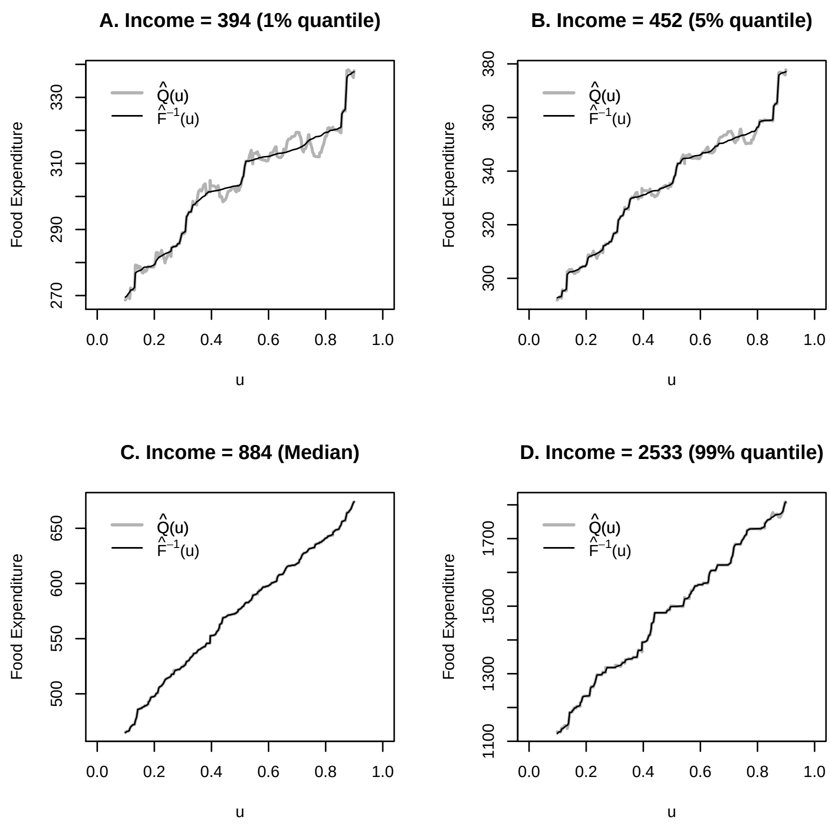

FIGURE 4. Quantile regression processes and rearranged quantile processes for the Engel food expenditure data. Quantile regression estimates are plotted with a thick gray line, whereas the rearranged estimates are plotted in black.

$(624.15, 0.55, -0.003)$ and $\gamma = (1, 0.0017, -0.0003)$. For each design, we draw 1,000 Monte Carlo samples of size $n = 235$. To generate the values of the dependent variable, we draw observations from a normal distribution with the same mean and variance as the residuals $\epsilon = (Y - Z(X)'\alpha)/(Z(X)'\gamma)$ of the Engel data set; and we fix the regressor X in all the replications to the observations of income in the Engel data set.

We use designs 1 and 2 to assess the estimation properties of the original and rearranged quantile regressions under the correct and incorrect specification of the functional

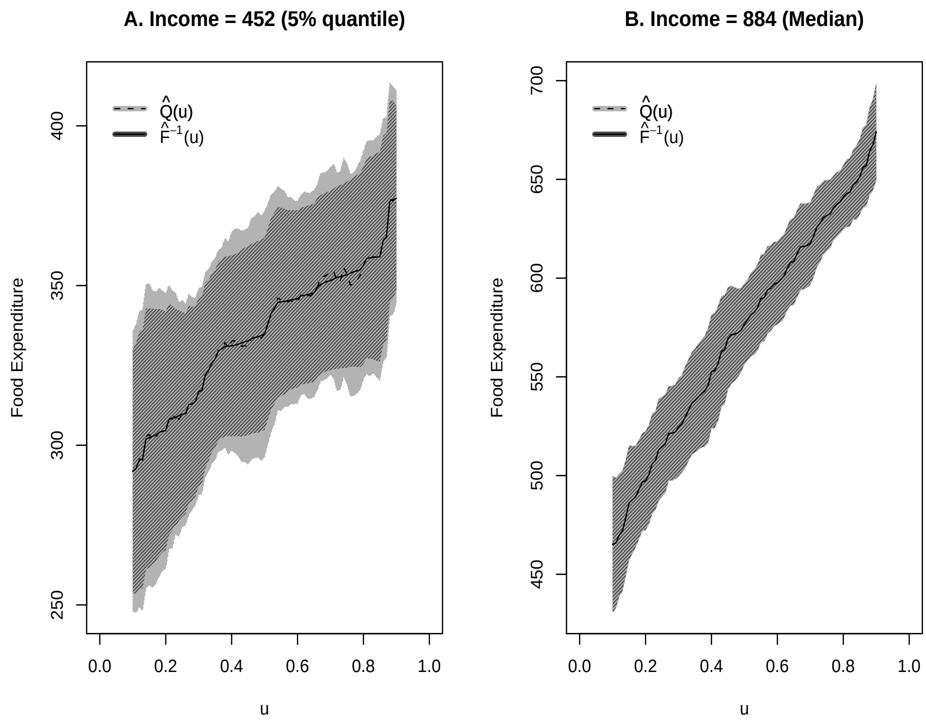

Figure 5. Simultaneous 90% confidence bands for quantile regression processes and rearranged quantile processes for the Engel food expenditure data. Two different values of the income regressor are considered. The bands for quantile regression are plotted in light gray, whereas the bands for rearranged quantile regression are plotted in dark gray.

form. Thus, in each replication, we estimate the model

$$Q(u|X) = Z(X)'\beta(u), \quad Z(X) = (1, X).$$

This gives the correct functional form for design 1, that is, $Q(u|X) \equiv Q_0(u|X)$, and an incorrect functional form for design 2, that is $Q(u|X) \not\equiv Q_0(u|X)$ (due to the omission of a nonlinear regressor). Accordingly, estimation error for design 1 arises entirely due to sampling error, while the estimation error for design 2 arises due to both sampling error and specification error. Regardless of the nature of the estimation error, Proposition 5 establishes that the rearranged quantile curves should be closer to the true conditional quantiles than the original curves.

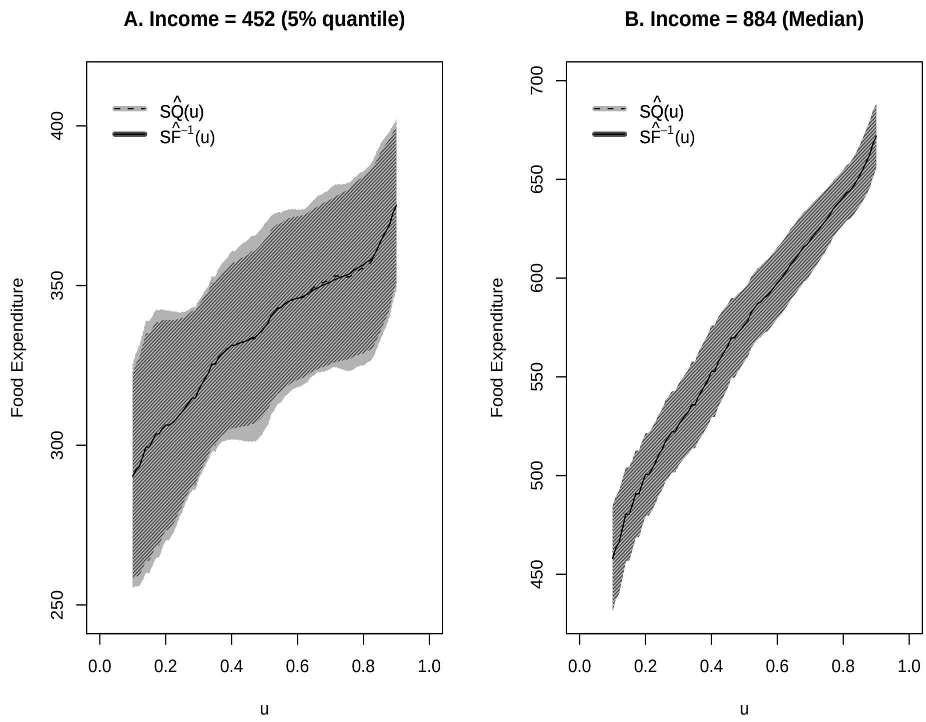

FIGURE 6. Simultaneous 90% confidence bands for the smoothed quantile regression processes and the smoothed rearranged quantile processes for the Engel food expenditure data. The bands for the smoothed original curves are plotted in light gray, whereas the bands for the smoothed rearranged curves are plotted in dark gray. The smoothed curves are obtained using a bandwidth equal to 0.05.

In each replication, we fit a linear quantile regression curve $\widehat{Q}(u|X) = X'\widehat{\beta}(u)$ and monotonize this curve to get $\widehat{F}^{-1}(u|X)$ using the rearrangement method. Table 1 reports measures of the estimation error of the original and rearranged estimated conditional quantile curves using different norms ($p = 1, 2, 3, 4$, and ∞), with the regressor fixed at a value, $X = x_0$, that corresponds to the 5% quantile of the regressor X ($X = 452$). We select this value motivated by the empirical example. Each entry of the table gives a Monte Carlo average of

$$L^p := \left(\int_{\mathcal{U}} |Q_0(u|x_0) - \widetilde{Q}(u|x_0)|^p du \right)^{1/p},$$

for $\widetilde{Q}(u|x_0) = x_0'\widehat{\beta}(u)$ and $\widetilde{Q}(u|x_0) = \widehat{F}^{-1}(u|x_0)$. We evaluate the integral using a net of indices u of size .01.

Both in the correctly specified case and in the misspecified case, we find that the rearranged curves estimate the true quantile curves $Q_0(u|X)$ more accurately than the original curves, providing a 4% to 15% reduction in the estimation/approximation error, depending on the norm.

Table 1. Estimation Error of Original and Rearranged Curves.

	Design 1: Correct Specification			Design 2: Incorrect Specification		
	Original	Rearranged	Ratio	Original	Rearranged	Ratio
L^1	6.79	6.61	0.96	7.33	7.02	0.95
L^2	7.99	7.69	0.95	8.72	8.20	0.93
L^3	8.93	8.51	0.95	9.85	9.12	0.92
L^4	9.70	9.17	0.94	10.78	9.86	0.91
L^∞	17.14	15.32	0.90	19.44	16.44	0.85

5. Conclusion

This paper analyzes a simple regularization procedure for estimation of conditional quantile and distribution functions based on rearrangement operators. Starting from a possibly non-monotone 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. Asymptotic distribution theory is derived for the rearranged curves, and the usefulness of the approach is illustrated with an empirical example and a simulation experiment.

† *Massachusetts Institute of Technology, Department of Economics and Operations Research Center, University College London, CEMMAP, and The University of Chicago. E-mail: vchern@mit.edu.* *Research support from the Castle Krob Chair, National Science Foundation, the Sloan Foundation, and CEMMAP is gratefully acknowledged.*

§ *Boston University, Department of Economics. E-mail: ivanf@bu.edu.*

‡ *Harvard University, Department of Economics. E-mail: galichon@fas.harvard.edu.* *Research support from the Conseil Général des Mines and the National Science Foundation is gratefully acknowledged.*

Appendix A. Proofs

A.1. **Proof of Proposition 1.** First, note that the distribution of Y_x has no atoms, i.e.,

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

since the number of roots of $Q(u|x) = y$ is finite under (a) - (b), and $U \sim \text{Uniform}(\mathcal{U})$. Next, by assumptions (a)-(b) the number of critical values of $Q(u|x)$ is finite, hence claim (1) 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 $\{u_k(y|x),\ \text{for } k = 1, ..., K(y|x) < \infty\}$ are the roots of $Q(u|x) = y$ in increasing order. Note that the sign of $Q'(u|x)$ 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 $\sum_{k=0}^{K(y|x)-1} 1\{Q'(u_{k+1}(y|x)|x) \ge 0\}(u_{k+1}(y|x) - u_k(y|x))$; while the last term simplifies to $1\{Q'(u_{K(y|x)}(y|x)|x) \le 0\}(1 - u_{K(y|x)}(y|x))$. An additional simplification yields the expression given in claim (2) of the proposition.

The proof of claim (3) follows by taking the derivative of expression in claim (2), noting that at any regular value y the number of solutions $K(y|x)$ and $\text{sign}(Q'(u_k(y|x)|x))$ are locally constant; moreover,

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

Combining these facts we get the expression for the derivative given in claim (3).

To show the absolute continuity of $F(y|x)$ with $f(y|x)$ being the Radon-Nykodym 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 $F(y|x)$, $\int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\} dF(y|x) \uparrow \int_{-\infty}^{y'} dF(y|x)$ 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 (4) follows by noting that at the regions where $s \to Q(s|x)$ is increasing and one-to-one, we have that $F(y|x) = \int_{Q(s|x) \leq y} ds = \int_{s \leq Q^{-1}(y|x)} ds = Q^{-1}(y|x)$. Inverting the equation $u = F(F^{-1}(u|x)|x) = Q^{-1}(F^{-1}(u|x)|x)$ yields $F^{-1}(u|x) = Q(u|x)$.

Claim (5). We have $Y_x = Q(U|x)$ has quantile function $F^{-1}(u|x)$. The quantile function of $\alpha + \beta Q(U|x) = \alpha + \beta Y_x$, for $\beta > 0$, is therefore $\inf\{y : \Pr(\alpha + \beta Y_x \leq y) \geq u\} = \alpha + \beta F^{-1}(u|x)$.

Claim (6) is immediate from claim (3).

Claim (7). The proof of continuity of $F(y|x)$ is subsumed in the step 1 of the proof of Proposition 3 (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 $F(y|x)$ is continuous. Let $u_t \to u$ and $x_t \to x$. Since $F(y|x) = u$ has a unique root $y = F^{-1}(u|x)$, the root of $F(y|x_t) = u_t$, i.e., $y_t = F^{-1}(u_t|x_t)$, converges to y by a standard argument, see, e.g., van der Vaart and Wellner (1997). □

A.2. **Proof of Propositions 2-7.** In the proofs that follow we will repeatedly use Lemma 1, which establishes the equivalence of continuous convergence and uniform convergence:

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

Proof of Lemma 1: See, for example, Resnick (1987), page 2. □

Proof of Proposition 2. We have that for any $\delta > 0$, there exists $\epsilon > 0$ such that for $u \in B_\epsilon(u_k(y|x))$ and for small enough $t \geq 0$

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

for all $k \in 1, ..., K(y|x)$; whereas for all $u \notin \cup_k B_\epsilon(u_k(y|x))$, as $t \to 0$,

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

Therefore,

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

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

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

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

Fixing $\epsilon > 0$, 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 $|Q'(Q^{-1}(y'|x)|x)| \to |Q'(u_k(y|x)|x)|$ as $Q^{-1}(y'|x) \to u_k(y|x)$. Therefore, the right hand term in (A.1) is no greater than

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

Similarly $\sum_{k=1}^{K(y|x)} \frac{-h(u_k(y|x)|x)-\delta}{|Q'(u_k(y|x)|x)|} + o(1)$ 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 (1) the function $(y,x) \mapsto \frac{-h(u_k(y|x)|x)}{|Q'(u_k(y|x)|x)|}$ is uniformly continuous on K, and (2) the function $(y,x) \mapsto K(y|x)$ is uniformly continuous on K. To see (2), 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 $(K_1, ..., K_M)$ such that the function $K(y|x)$ is constant over each of these sets, i.e., $K(y|x) = k_j$ for some integer $k_j > 0$, for all $(y,x) \in K_j$ and $j \in \{1, ..., M\}$. Likewise, (1) follows by noting that the limit expression for the derivative is continuous on each of the sets $(K_1, ..., K_M)$ 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 $Q'(u|x)$ in both arguments. $\square$

Proof of Proposition 3. For a fixed x the result follows by Proposition 2, by step 1 of the proof 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, $\epsilon_t := t\|h_t\|_\infty + \|Q(\cdot|x_t) - Q(\cdot|x)\|_\infty + |y_t - y| \to 0$, and

$$
\begin{aligned}
|F(y_t|x_t, h_t) - F(y|x)| &\leq \left| \int_0^1 [1\{Q(u|x_t) + th_t(u|x) \leq y_t\} - 1\{Q(u|x) \leq y\}]du \right| \\
&\leq \left| \int_0^1 1\{|Q(u|x) - y| \leq \epsilon_t\}du \right| \to 0, \qquad \text{(A.2)}
\end{aligned}
$$

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[1] the uniform convergence of quantiles $F^{-1}(u|x, h_t) \to F^{-1}(u|x)$, uniformly over K^*, where K^* is any compact subset of $\mathcal{UX}^*$.

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

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

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

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

and (as will be shown below)

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

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.

[1]See, e.g., Lemma 1 in Chernozhukov and Fernandez-Val (2005).

It only remains to show that equation (A.5) holds uniformly in K^*. Note that for any right-continuous cdf F, we have that $u \leq F(F^{-1}(u)) \leq u + F(F^{-1}(u)) - F(F^{-1}(u)-)$, where $F(\cdot -)$ denotes the left limit of F, i.e., $F(x_0-) = \lim_{x \uparrow x_0} F(x)$. For any continuous, strictly increasing cdf F, we have that $F(F^{-1}(u)) = u$. Therefore, write

$$\begin{aligned}
0 &\leq \frac{F(F^{-1}(u|x,h_t)|x,h_t) - F(F^{-1}(u|x)|x)}{t} \\
&\leq \frac{u + F(F^{-1}(u|x,h_t)|x,h_t) - F(F^{-1}(u|x,h_t) - |x,h_t) - u}{t} \\
&\leq \frac{F(F^{-1}(u|x,h_t)|x,h_t) - F(F^{-1}(u|x,h_t) - |x,h_t)}{t} \\
&\overset{(1)}{=} \frac{[F(F^{-1}(u|x,h_t)|x,h_t) - F(F^{-1}(u|x,h_t)|x)]}{t} \\
&- \frac{[F(F^{-1}(u|x,h_t) - |x,h_t) - F(F^{-1}(u|x,h_t) - |x)]}{t} \\
&\overset{(2)}{=} D_h(F^{-1}(u|x,h_t)|x) - D_h(F^{-1}(u|x,h_t) - |x) + o(1) = o(1),
\end{aligned}$$

as $t \to 0$, where in (1) we use that $F(F^{-1}(u|x,h_t)|x) = F(F^{-1}(u|x,h_t) - |x)$ since $F(y|x)$ is continuous and strictly increasing in y, and in (2) we use Proposition 2. □

The following lemma, due to Pratt (1960), will be very useful to prove Proposition 4.

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

Proof of Lemma 2. See Pratt (1960). □

Lemma 3 (Boundedness and Integrability Properties)**.** *Under the hypotheses of Proposition 2 and 3, we have that for all $(y,x) \in \mathcal{Y}\mathcal{X}$:*

$$|\widetilde{D}_{h_t}(u|x,t)| \leq \|h_t\|_\infty, \tag{A.6}$$

and

$$|D_{h_t}(y|x,t)| \leq \Delta(y|x,t) = \int_0^1 \frac{1\{|Q(u|x) - y| \leq t\|h_t\|_\infty\}}{t} du, \tag{A.7}$$

where for any $x_t \to x \in \mathcal{X}$, as $t \to 0$,

$$\Delta(y|x_t,t) \to 2\|h\|_\infty f(y|x) \text{ for a.e } y \in \mathcal{Y} \text{ and } \int_{\mathcal{Y}} \Delta(y|x_t,t)dy \to \int_{\mathcal{Y}} 2\|h\|_\infty f(y|x)dy.$$

Proof of Lemma 3. To show (A.6) note that

$$\sup_{x\in\mathcal{X},y\in\mathcal{Y}} |\widetilde{D}_{h_t}(y|x,t)| \le \|h_t\|_\infty \tag{A.8}$$

immediately follows from the equivariance property noted in Claim (5) 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 respectively 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) .

Further, 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 $2\|h\|_\infty$. □

A.3. **Proof of Proposition 4.** Define $m_t(y|x,y') := 1\{y \le y'\}g(y|x)D_{h_t}(y|x,t)$ and $m(y|x,y') := 1\{y \le y'\}g(y|x)D_h(y|x)$. To show claim (1), we need to demostrate 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.

In order 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 $\|g\|_\infty\|h\|_\infty f(y|x_t)$, which converges to $\|g\|_\infty\|h\|_\infty f(y|x)$ for almost every

y, and, moreover, $\int_{\mathcal{Y}} \|g\|_\infty \|h\|_\infty f(y|x)dy$ converges to $\|g\|_\infty \|h\|_\infty$. Conclude that (A.11) holds by Lemma 2.

To show claim (2), define $m_t(u|x,u') = 1\{u \leq u'\}g(u|x)\tilde{D}_{h_t}(u|x)$ and $m(u|x,u') = 1\{u \leq u'\}g(u|x)\tilde{D}_h(u|x)$. 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.

In order 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 $|g(u|x_t)|\|h\|_\infty$, which converges for almost every value of u to $|g(u|x)|\|h\|_\infty$ as $t \to 0$. Furthermore, the integral of $|g(u|x_t)|\|h\|_\infty$ converges to the integral of $|g(u|x)|\|h\|_\infty$ by dominated convergence theorem. We conclude that (A.13) holds by Lemma 2. □

The following lemma will be used to prove Proposition 5:

Lemma 4. *Assume that $Q(u)$ is a measurable function mapping $\mathcal{U} := (0,1)$ to K, a bounded subset of $\mathbb{R}$, and that $Q_0(u)$ is a non-decreasing "true" function mapping $\mathcal{U}$ to K, that is also measurable. Think of $Q(u)$ as an approximation to $Q_0(u)$. Let $F_Q(y) = \int_{\mathcal{U}} 1\{Q(u) \leq y\}du$ denote the distribution function of $Q(U)$ when $U \sim U(0,1)$. Let $Q^*(u) = F_Q^{-1}(u) = \inf\{y \in \mathbb{R} : F_Q(y) \geq u\}$. Then, for any $p \in [1, \infty]$,*

$$\left[\int_{\mathcal{U}} |Q_0(u) - Q^*(u)|^p \, du\right]^{1/p} \leq \left[\int_{\mathcal{U}} |Q_0(u) - Q(u)|^p \, du\right]^{1/p}.$$

Moreover, this inequality is strict provided (1) $p \in (1,\infty)$, (2) $Q(u)$ is decreasing on a subset of $\mathcal{U}$ that has positive Lebesgue measure, and (3) the true function $Q_0(u)$ is increasing on $\mathcal{U}$.

Proof of Lemma 4. A direct proof of this lemma is given in Proposition 1 of Chernozhukov, Fernandez-Val, Galichon (2006a). It is helpful to give a quick indirect proof of the weak inequality contained in the lemma using the following inequality due to Lorentz (1953): Let Q and G be two functions mapping $\mathcal{U}$ to K, a bounded subset of $\mathbb{R}$. Let Q^* and G^* denote their corresponding increasing rearrangements. Then, we have

$$\int_{\mathcal{U}} L(Q^*(u), G^*(u))du \leq \int_{\mathcal{U}} L(Q(u), G(u))du,$$

for any submodular discrepancy function $L : \mathbb{R}^2 \mapsto \mathbb{R}_+$ In our case, $G(u) = Q_0(u) = G^*(u) = Q_0^*(u)$ almost everywhere. Thus, the true function is its own rearrangement. Moreover, $L(v, w) = |w - v|^p$ is submodular for $p \in [1, \infty)$. For the proof of the strict inequality, please refer to Chernozhukov, Fernandez-Val, Galichon (2006a), Proposition 1. For $p = \infty$, the inequalities follows by taking limit as $p \to \infty$. □

A.4. **Proof of Proposition 5.** This proposition is an immediate consequence of Lemma 4. □

A.5. **Proof of Proposition 6.** This Proposition simply follows by the functional delta method (e.g. van der Vaart, 1998). Instead of restating what this method is, it takes less space to simply recall the proof in the current context.

To show the first part, consider the map $g_n(y, x|h) = \sqrt{n}(F(y|x, n^{-1/2}h) - F(y|x))$. 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 $\ell^\infty(K)$, $g_n(y, x|\sqrt{n}(\widehat{Q}(u|x) - Q(u|x))) \Rightarrow D_G(y|x)$ as a stochastic process indexed by (y, x), since $\sqrt{n}(\widehat{Q}(u|x) - Q(u|x)) \Rightarrow G(u|x)$ in $\ell^\infty(K)$.

Conclude similarly for the second part. □

A.6. **Proof of Proposition 7.** This follows by the functional delta method, similarly to the proof of Proposition 6. □

References

[1] Ait-Sahalia, Y. and Duarte, J. (2003),"Nonparametric option pricing under shape restrictions," *Journal of Econometrics* 116, pp. 9–47.

[2] Alvino, A., Lions, P. L. and Trombetti, G. (1989), "On Optimization Problems with Prescribed Rearrangements," *Nonlinear Analysis* 13 (2), pp. 185–220.

[3] Angrist, J., Chernozhukov, V., and I. Fernandez-Val (2006): "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure," *Econometrica* 74, pp. 539–563.

[4] Billingsley, P. (1995), *Probability and measure*. Third edition. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York.

[5] Chernozhukov, V., and I. Fernández-Val (2005): "Subsampling Inference on Quantile Regression Processes. " *Sankhya* 67, pp. 253–276.

[6] Chernozhukov, V., Fernandez-Val, I., and A. Galichon (2006a): "Improving Estimates of Monotone Functions by Rearrangement," Preprint, available at arxiv.org and ssrn.com.

[7] Chernozhukov, V., Fernandez-Val, I., and A. Galichon (2006b): "An Addendum for Quantile and Probability Curves Without Crossing (Alternative Proof Directions and Explorations) " Preprint.

[8] Dette, H., Neumeyer, N., and K. Pilz (2006): "A simple Nonparametric Estimator of a Strictly Monotone Regression Function," *Bernoulli*, 12, no. 3, pp 469-490.

[9] Doksum, K. (1974): "Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case," *Annals of Statistics* 2, pp. 267–277.

[10] Doss, Hani; Gill, Richard D. (1992), "An elementary approach to weak convergence for quantile processes, with applications to censored survival data." *Journal of the American Statistical Association* 87, no. 419, 869–877.

[11] Dudley, R. M., and R. Norvaisa (1999), *Differentiability of six operators on nonsmooth functions and p-variation*. With the collaboration of Jinghua Qian. Lecture Notes in Mathematics, 1703. Springer-Verlag, Berlin.

[12] Engel, E. (1857), "Die Produktions und Konsumptionsverhältnisse des Königreichs Sachsen," *Zeitschrift des Statistischen Bureaus des Königlich Sächsischen Misisteriums des Innerm,* 8, pp. 1-54.

[13] Gill, R. D., and S. Johansen (1990), "A survey of product-integration with a view toward application in survival analysis." *Annals of Statistics* 18, no. 4, 1501–1555.

[14] Gutenbrunner, C., and J. Jurečková (1992): "Regression Quantile and Regression Rank Score Process in the Linear Model and Derived Statistics," *Annals of Statistics* 20, pp. 305-330.

[15] Hall, P., Wolff, R., and Yao, Q. (1999), "Methods for estimating a conditional distribution function," *Journal of the American Statistical Association* 94, pp. 154–163.

[16] Hardy, G., Littlewood, J., and G. Polya (1952), *Inequalities*. Cambridge: Cambridge University Press.

[17] He, X. (1997), "Quantile Curves Without Crossing," *American Statistician*, 51, pp. 186–192.

[18] Koenker, R. (1994): "Confidence Intervals for Regression Quantiles," in M.P. and M. Hušková (eds.), *Asymptotic Statistics: Proceeding of the 5th Prague Symposium on Asymptotic Statistics.* Physica-Verlag.

[19] Koenker, R. (2005), *Quantile Regression.* Econometric Society Monograph Series 38, Cambridge University Press.

[20] Koenker, R., P. Ng (2005), "Inequality constrained quantile regression." *Sankhyā* 67, no. 2, 418–440.

[21] Koenker, R., and Z. Xiao (2002): "Inference on the Quantile Regression Process," *Econometrica* 70, no. 4, pp. 1583–1612.

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

[23] Lorentz, G. G. (1953): "An Inequality for Rearrangements," *The American Mathematical Monthly* 60, pp. 176–179.

[24] Milnor, J. (1965), *Topology from the differential viewpoint*, Princeton University Press.

[25] Mossino J. and R. Temam (1981), "Directional derivative of the increasing rearrangement mapping and application to a queer differential equation in plasma physics," *Duke Math. J.* 48 (3), 475–495.

[26] Portnoy, S. (1991), "Asymptotic behavior of regression quantiles in nonstationary, dependent cases," *Journal of Multivariate Analysis* 38 , no. 1, 100–113.

[27] Portnoy, S. , and R. Koenker (1997), "The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators," *Statist. Sci.* 12, no. 4, 279–300.

[28] Pratt, J.W. (1960), "On interchanging limits and integrals." *Annals of Mathematical Statistics* 31, 74–77.

[29] Resnick, S. I. (1987), *Extreme values, regular variation, and point processes*, Applied Probability. A Series of the Applied Probability Trust, 4. Springer-Verlag, New York.

[30] Vaart, A. van der (1998). *Asymptotic statistics.* Cambridge Series in Statistical and Probabilistic Mathematics, 3.

[31] Vaart, A. van der, and J. Wellner (1996), *Weak convergence and empirical processes: with applications to statistics*, New York: Springer.

[32] Villani, C. (2003), *Topics in Optimal Transportation*, Providence: American Mathematical Society.