QUANTILE AND PROBABILITY CURVES WITHOUT CROSSING

VICTOR CHERNOZHUKOV   IVAN FERNANDEZ-VAL   ALFRED GALICHON

Abstract. The commonly used approach in estimating conditional quantile curves is to fit typically a linear curve pointwise for each quantile. This is done for a number of reasons: linear models enjoy good approximation properties as well as have excellent computational properties. The resulting fits may not respect a logical monotonicity requirement – that the quantile curve, as a function of probability, should be monotone in that 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 do not cross. This construction of monotone quantile curves may be seen as a semi-parametric bootstrap and also as a monotonic rearrangement of the original non-monotone function. These conditional quantile curves are monotone in the probability and we show that, under correct specification, these curves have the same asymptotic distribution as the original non-monotone curves. Thus, the empirical non-monotone curves can be rearranged to be monotone without changing their (first order) asymptotic distribution. However, this property does not hold under misspecification and the asymptotics of these curves partially differs from the asymptotics of the original non-monotone curves. Towards establishing the result, we establish the compact (Hadamard) differentiability of the monotonized quantile and probability curves with respect to the original curves. In doing so, we establish the results on the compact differentiability of functions related to rearrangement operators. These results therefore generalize earlier results on the compact differentiability of the inverse (quantile) operators.

Date: September 18, 2006.

1. Introduction

We can best describe the problem studied in this paper using linear quantile regression models as the prime example. Suppose that $x'\beta(u)$ is a linear approximation to the u-quantile of real response variable Y given a vector of repressors $X = x$, for a given index u. The typical approximations as well as estimation algorithms fit the conditional curve $x'\beta(u)$ pointwise in $u \in (0, 1)$, producing an estimate $\hat{\beta}(u)$. The linear functional forms as well as pointwise fitting are used for a number of reasons, including parsimony of resulting approximation coupled with good computational properties. However, a problem that might occur is that the map

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

may not be monotone in u, which violates the obvious logical requirement. In fact, the non-monotonicity can occur due to either of the following reasons:

(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 $u \mapsto x'\hat{\beta}(u)$ is not monotone due to an estimation error.

(2) (Monotonically incorrect case). The population curve $u \mapsto x'\beta(u)$ is not monotone due to imperfect approximation, and thus does not satisfy the monotonicity requirement. Therefore, the resulting empirical curve $u \mapsto x'\hat{\beta}(u)$ is also not monotone due to both estimation error and the non-monotonicity of the population curve.

Consider the function

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

This function is monotone in y. Moreover, this is a proper distribution function of the random variable

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

Hence

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

is monotone in u, hence this "rearranged" quantile curve is monotone in u, for each x. Thus, starting with an original non-monotone curve $u \mapsto x'\hat{\beta}(u)$, the rearrangement produces a monotone quantile curve $\widehat{F}^{-1}(u|x)$. This rearrangement mechanism has a direct relation to the bootstrap, since the "rearranged" quantile curve is produced by sampling from the estimated original quantile model (cf. Koenker, 1994). This mechanism (and the name we adopt for it) also has the direct relation to the "rearrangement mappings" in variational analysis and operations research (e.g. Hardy et. al. (1953) and Villani (2003)).

The purpose of this paper is to establish the empirical properties of the corrected quantile curve and its distribution curve:

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

under scenarios (1) and (2). We will also look closely at the analytical properties of the population curves: $u \mapsto F^{-1}(u|x)$ and $u \mapsto F(u|x)$.

Towards describing the essence of the result, let us fix an x, and suppose that an estimate of $\hat{\beta}(u)$ of $\beta(u)$ is available that converges weakly to a Gaussian process:

$$\sqrt{n}x'(\hat{\beta}(u) - \beta(u)) \Rightarrow x'G(u) \tag{1.1}$$

in the metric space of bounded functions $\ell^\infty(0,1)$, e.g. under the conditions of Jureckova (1992). Let $F(y|x) := \int_0^1 1\{x'\beta(u) \leq y\}du$. The main result is that in the monotonically correct case,

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

in metric space $\ell^\infty(\mathcal{Y})$, where $\mathcal{Y}$ is the support of (Y_x), and

$$F'(y|x) = \frac{1}{x'\beta'(u)}\Big|_{u=F(y|x)}.$$

This result follows by finding the Hadamard derivative, which is the main effort of the paper, and then appealing to a delta method. Further, we show that

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow x'[G(u)]$$

in $\ell^\infty((0,1) \times \mathcal{X})$, which, remarkably, coincides with the asymptotics of the original curve $u \mapsto x'\beta(u)$. This 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. This result follows by finding the Hadamard derivative, which is again the main effort of the paper, and then appealing to a delta method.

Further, the second result is that in the monotonically incorrect case,

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

where $u_1(x,y) < ... < u_K(x,y)$ are solutions of equation $y = x'\beta(u)$, assuming there is only a finite number of them. This holds uniformly in (y,x) over appropriate sets defined in the forthcoming section. Further,

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

This result holds also holds uniformly in (u,x) over appropriate sets defined in the forthcoming section.

Note that the same method also applies, with adaptation, to rearranging probability distributions curves which are not monotonic. As discussed in detail in Section 3, the monotone rearrangement on the quantile curve can also be done on a probability curve (by exchanging the roles played by the quantile and the probability spaces). Suppose $\hat{P}(y|x)$ is a candidate estimate of a probability distribution curve which is not monotone. Consider a rearranged monotone quantile curve associated with $\hat{P}(y|x)$:

$$\hat{Q}\left(y|x\right) = \int_0^\infty 1\{\hat{P}\left(y|x\right) < t\}dy - \int_{-\infty}^0 1\{\hat{P}\left(y|x\right) > t\}dy$$

Then take the inverse of this quantile curve as the rearranged probability curve:

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

which is monotone. As it will be argued with more detail, a similar asymptotic distribution theory goes through for $\hat{Q}$ and $\hat{F}$. These results are useful in a variety of applications where conditional distribution curves are estimated, but monotonicity does not hold due to the pointwise nature of the approximations used, see for example Hall et. al. (1999).

Our results can be viewed as a general functional delta method for rearrangement operators that include the inverse (quantile) operators as the special case. In this way our results extend the previous results on compact differentiability of the quantile operators by Gill and Johansen (1990) and Doss and Gill (1992) (see also the book by Dudley and Norvaisa (1999) for an excellent treatment). Since the results on quantile operators have seen numerous applications, we expect as many applications of the results on rearrangement operators, especially in the regression context. There are many more potential applications to the objects that are not probability or quantile curves, for example, demand curves and production curves in economics, or growth curves in biometrics. In these applications, the pointwise nature of fitting often introduces an error, which can be removed by the operation of the rearrangement.

It can also be emphasized that since the results are given in terms of the differentiability of the operator with respect to the basic estimated process, our results do not rely on particular sampling properties or on a particular estimator being used. We also do not rely on functional forms being linear (the previous discussion used them only as an example). The only condition required is that a central limit theorem like (1.1) applies to the estimate of a curve and that the population curves have some smoothness properties. Further, the results do not rely on the exact nature of the population curve, for example, the results do not depend on whether the underlying model is an ordinary or a structural quantile regression model. Our results apply in either case without modification.

2. Quantile Curves without Crossing: Analytical and Empirical Properties

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

2.1. **Basic Analytical Properties.**

In this section, we develop the basic properties of the population objects $F(y|x)$, its density $f(y|x)$, and its inverse $F^{-1}(u|x)$.

We first recall the following basic definitions from Millnor (1965): Let $g : \mathcal{U} \subset \mathbb{R} \to \mathbb{R}$ a continuously differentiable function. A point $x \in \mathcal{U}$ is called a regular point of g if the derivative of g at this point does not vanish, ie. $g'(x) \neq 0$. A point x 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, ie. if $\forall x \in g^{-1}(\{y\})$, $g'(x) \neq 0$. A value y which is not a regular value is called a critical value.

Denote $\mathcal{Y}_x$ the support of Y_x, and $\mathcal{Y}_x^*$ be the subset of regular values of $u \mapsto Q(u,x)$ in $\mathcal{Y}_x$. Denote $\mathcal{XY} = \{(x,y) : x \in \mathcal{X}, y \in \mathcal{Y}_x\}$, and $\mathcal{XY}^* = \{(x,y) : x \in \mathcal{X}, y \in \mathcal{Y}_x^*\}$. 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 $\mathbb{R}^d$.

Introduce the following assumptions:

(a) $Q(u,x) : \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ is a continuously differentiable function in both arguments,
(b) $Q'(x,u) := \partial Q(u,x)/\partial u$ does not vanish almost everywhere on $\mathcal{U}$, for each $x \in \mathcal{X}$,
(c) the number of times $\partial Q(u,x)/\partial u$ changes its sign is bounded, uniformly in $x \in \mathcal{X}$.

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 $\mathcal{X} = \{x\}$.

Proposition 1 (Basic Properties of $F(y|x)$, $f(y|x)$, and $F^{-1}(y|x)$)**.** *Under assumpions (a) and (b), Let $\mathcal{Y}_x^*$ be the subset of regular values of $Q(u,x)$ in $\mathcal{Y}_x$, the support of Y_x. The basic properties of function $F(y|x)$ are the following,*

1. *The set of non-regular 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 $y \in \mathcal{Y}_x^*$*

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

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 $y \in \mathcal{Y}^*$, the ordinary derivative $f(y|x) = \partial F(y|x)/\partial y$ exists, and its value is*

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

and it is continuous at each $y \in \mathcal{Y}_x^*$. $F(y|x)$ is absolutely continuous in $y \in \mathcal{Y}$ and is strictly increasing in $y \in \mathcal{Y}$. 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 $Q(u,x)$ on regions where the latter function is increasing, namely

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

provided $u^* = u$ is the unique solution of the equation $F^{-1}(u|x) = Q(u^*, x)$.

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

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

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

when $F^{-1}(u|x) \in \mathcal{Y}_x^*$, and 0 when $F^{-1}(u|x) \in \mathcal{Y}_x \setminus \mathcal{Y}_x^*$, and the derivative is continuous. This 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 synthetic simple example illustrates some of these basic properties in a situation where the original quantile curve is chosen to be highly non monotone. Consider the following hypothetical quantile function:

$$Q(u) = (4u-2)^3 - (4u-2). \tag{2.1}$$

Figure 1 shows that this function is non monotone. In particular, the slope of $Q(u)$ changes sign twice at .36 and .64 approximately. Figures 1 and 2 illustrate the second and fourth results of the proposition by plotting together the original and rearranged quantile and distribution curves. Here we can see that the rearranged functions are monotonically increasing and coincide with the original ones for points where the initial distribution is uniquely defined. For the rest of the points y, the rearranged distribution is a linear combination of the solutions to the equation $Q(u) = y$. The coefficients in this linear combination alternate between 1 and -1 depending on whether the original quantile curve is increasing or decreasing at the corresponding solution point.

Figures 2 illustrates the third and sixth results of the proposition by plotting the rearranged density and sparsity curves. The density function is continuous at the regular values of $Q(u)$, and grows asymptotically to infinity at the critical values. Similarly, the sparsity function is continuous at the regular points of $Q(u)$, and takes the value zero at the critical points.

2.2. **Functional Derivatives.** In this section, establish the Hadamard differentiability of $F(y|x)$ and $F^{-1}(y|x)$ with respect to $Q(u,x)$, tangentially to the space of continuous functions on $\mathcal{U} \times \mathcal{X}$. This differentiability property is important for statistical applications.

In what follows, $\ell^{\infty}(\mathcal{U} \times \mathcal{X})$ denotes the set of bounded and measurable functions $h : \mathcal{U} \times \mathcal{X} \to \mathbb{R}$, $C(\mathcal{U} \times \mathcal{X})$ denotes the set of continuous functions mapping $h : \mathcal{U} \times \mathcal{X} \to \mathbb{R}$, and $\ell^{1}(\mathcal{U} \times \mathcal{X})$ denotes the set of measurable functions $h : \mathcal{U} \times \mathcal{X} \to \mathbb{R}$ 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 Derivative of $F(y|x)$ with respect to $Q(u,x)$) Define $F(y|x,h_t) := \int_0^1 \{Q(x,u) + th_t(u,x) \leq y\} du$. Under assumptions (a)-(c), 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 $h_t \in \ell^\infty[\mathcal{U} \times \mathcal{X}]$, and $h \in C(\mathcal{U} \times \mathcal{X})$.

Proposition 3 (Hadamard Derivative of $F^{-1}(u|x)$ with respect to $Q(u,x)$)**.** *Under assumptions (a) and (b) of Lemma 1, 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 subsets of $\mathcal{UX}^* = \{(u,x) : (F^{-1}(u|x),x) \in \mathcal{YX}^*\}$, for every $|h_t - h|_\infty \to 0$, where $h_t \in \ell^\infty[\mathcal{U} \times \mathcal{X}]$, and $h \in C(\mathcal{U} \times \mathcal{X})$.

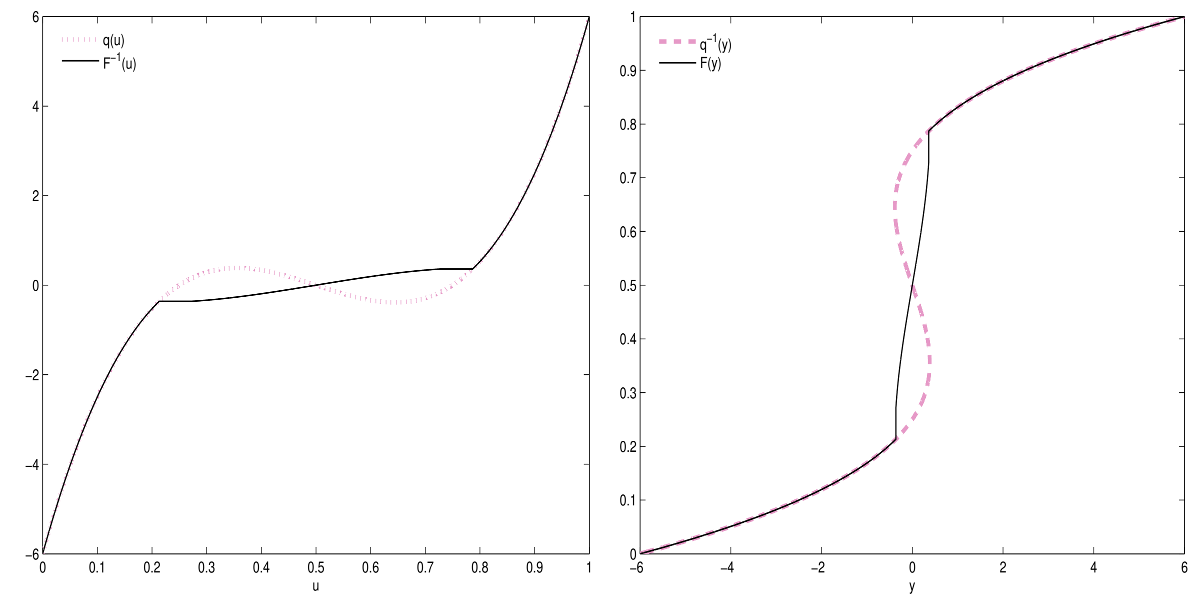

FIGURE 1. Left: "Approximate" quantile function $Q(u)$ and the rearranged quantile function. Right: "Approximate" distribution function $Q^{-1}(y)$ and the rearranged distribution function $F(y)$.

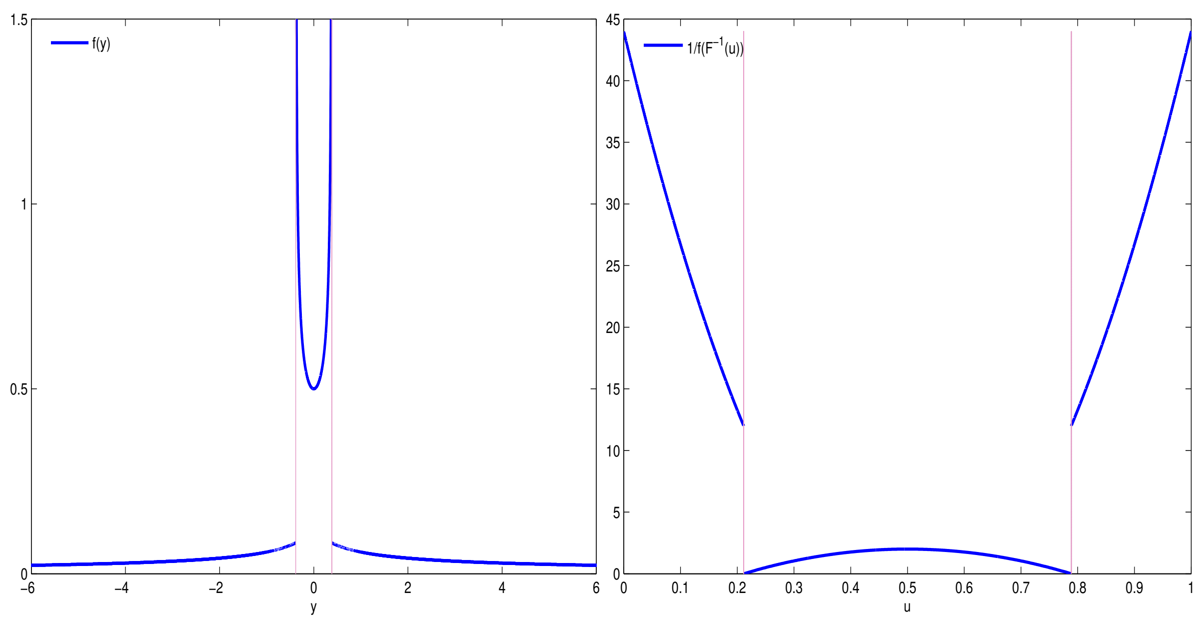

FIGURE 2. Left: Density function $f(y)$ of rearranged distribution function $F(y)$. Right: Density (sparsity) function the rearranged quantile function $F^{-1}(u))$.

The convergence holds uniformly on the 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, the following result is worth emphasizing:

Corollary 1 (Monotonically correct case)**.** *Suppose* $u \mapsto Q(u,x)$ *has* $Q'(u,x) > 0$ *for each* $x \in \mathcal{X}$ *then*

$$D_{h_t}(y,x,t) \to D_h(x,y)$$

uniformly in $(y,x) \in \mathcal{Y}\mathcal{X}$*, and*

$$\tilde{D}_{h_t}(u,x,t) \to \tilde{D}_h(u,x)$$

uniformly in $(u,x) \in \mathcal{U}\mathcal{X}$*.*

Thus, the convergence is uniform over the entire domain in the monotonically correct case. Regarding the monotonically incorrect case, a natural question that arises next is whether by some operation of smoothing – namely integrating in y or x – we can achieve the uniform convergence over the entire space. The answer appears positive in the case of integrating in y, while some exclusions are needed in the case of integrating $F(y|x)$ in x.

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

$$(y',x) \mapsto \int_{\mathcal{Y}} 1\{y \leq y'\} g(y,x) F(y|x) dy, \quad (y,x') \mapsto \int_{\mathcal{X}_0} 1\{x \leq x'\} g(y,x) F(y|x) dx,$$
$$(u',x) \mapsto \int_{\mathcal{U}} 1\{u \leq u'\} g(u,x) F^{-1}(u|x) du, \quad (u,x) \mapsto \int_{\mathcal{X}} 1\{x \leq x'\} g(u,x) F^{-1}(u|x) dx,$$

with restrictions on g and the set and the set $\mathcal{X}_0$ specified below in the statement of the theorem. These elementary functionals are useful building blocks for various statistics, such as Lorentz curves, means and other moments, as well as smoothed versions of curves $F(y|x)$ and $F^{-1}(y|x)$. For example, the conditional Lorentz curve is

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

which is a ratio of a partial mean to the mean. Hadamard differentiability of this statistic with respect to the underlying $Q(u,x)$ immediately follows from the Hadamard differentiability of the elementary statistics above.

We need the following definitions for some of the results presented below. Let $\mathcal{X}^o$ be a region such that when $X' \sim \text{Uniform}(\mathcal{X}_0)$, $Q(X',u)$ has a density denoted $m(q|x)$ bounded above uniformly in u. Note that this is not an assumption on the sampling behavior of regressors, it is merely a definition of the region $\mathcal{X}_0$. In typical linear applications, $\mathcal{X}^o$ coincides with $\mathcal{X}$ if regressors are not redundant.[1] Further let $\mathcal{Y}_0$ be a subset of $\mathcal{Y}$ such that $\int_{\mathcal{X}_0} f(y|x)dx$ is bounded above for all $y \in \mathcal{Y}$. It should be noted that the integral is finite for almost every value of y, and thus $\mathcal{Y}_0$ may be smaller than $\mathcal{Y}$.

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

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

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

2. $$\int_{\mathcal{X}_0} 1\{x \le x'\} g(y,x) D_{h_t}(y,x,t)dx \to \int_{\mathcal{X}_0} 1\{x \le x'\} g(y,x) D_h(y,x)dx$$

uniformly in $(x',y) \in \mathcal{X}_0 \times \mathcal{Y}_0$, *for any* $g \in \ell^\infty(\mathcal{Y}\mathcal{X})$ *such that such that* $y \mapsto g(y,x)$ *is continuous for a.e.* x.

3. $$\int_{\mathcal{U}} 1\{u \le u'\} g(u,x) \tilde{D}_{h_t}(u,x,t)du \to \int_{\mathcal{U}} 1\{u \le u'\} g(u,x) \tilde{D}_h(u,x)dy$$

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

4. $$\int_{\mathcal{X}} 1\{x \le x'\} g(u,x) \tilde{D}_{h_t}(u,x,t)dx \to \int_{\mathcal{X}} 1\{x \le x'\} g(u,x) \tilde{D}_h(u,x)dx$$

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

---

[1] In the linear case $\partial Q(u,x)/\partial x = \beta_{-1}(u)$.

This proposition essentially is a corollary of Proposition 4. Indeed, results (1)-(4) follow from the fact that pointwise convergence of Proposition 3 coupled with uniform integrability, shown in the proof of Proposition 4, allows to exchange limits and integration. Another way of proving result (3) but *not* any other result in the paper can be based on exploiting the convexity of the functional in (3) with respect to the underlying curve, following the approach of Massino and Temam (1982) and Alvino et. al. (1990). Due to this limitation, we did not follow this approach in this paper. However, details of this approach are described in Chernozhukov, Fernandez-Val, and Galichon (2006) with an application to some nonparametric estimation problems.

It is worthwhile emphasizing also the properties of the following functionals obtained by applying a smoothing operator, denoted by S:

$$SF(y|x) = \int k_\delta(y-y')F(y'|x)dy', \quad k_\delta(v) = 1\{|v| \le \delta\}/2\delta,$$
$$SF^{-1}(u|x) = \int k_\delta(u-u')F(u'|x)du', \quad k_\delta(v) = 1\{|v| \le \delta\}/2\delta.$$

Since these curves are merely formed as differences of the elementary functionals in Proposition 4, parts (1) and (3), followed by a division by δ, the following corollary is immediate.

Corollary 2. *We have that*

$$SD_{h_t}(y,x,t) \to SD_h(x,y)$$

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 prospective 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 result for this section:

Proposition 5 (Empirical Properties of $(y,x) \mapsto \widehat{F}(y|x)$ and $(y,x) \mapsto \widehat{F}^{-1}(u,x)$)**.** *Suppose an estimate available such that $\hat{Q}(\cdot,\cdot)$ takes its values in the space of bounded measurable functions defined on $\mathcal{U} \times \mathcal{X}$, and that in $\ell^\infty(\mathcal{U} \times \mathcal{X})$*

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

where $G(x,u)$ is a Gaussian process with continuous paths. Suppose also that $Q(x,u)$ satisfies basic conditions (a),(b) and (c). Then in $\ell^\infty(K)$, where K is any subset of $\mathcal{Y}\mathcal{X}^*$ that is compact,

$$\sqrt{n}(\widehat{F}(y|x) - F(y|x)) \Rightarrow D_{G(\cdot,x)}(y,x)$$

and in $\ell^\infty(\mathcal{U}\mathcal{X}_K) = \{u : F^{-1}(u|x) \in K, x \in \mathcal{X}\}$,

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow f(F^{-1}(u|x)|x) \cdot D_{G(\cdot,x)}(F^{-1}(u|x),x).$$

Proposition 6 (Empirical Properties of Integrated Curves)**.** *Under the hypotheses of Proposition 5, 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 stochastic processes indexed by (x,y'), in $\ell^\infty(\mathcal{X} \times \mathcal{Y})$

2. $$\sqrt{n}\int_{\mathcal{X}_0} 1\{x \leq x'\} g(y,x)(\widehat{F}(y|x) - F(y|x))dx \Rightarrow \int_{\mathcal{X}_0} 1\{x \leq x'\} g(y,x) D_G(y,x)dx$$

as stochastic processes indexed by (x',y), in $\ell^\infty(\mathcal{X}_0 \times \mathcal{Y}_0)$

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

as stochastic processes indexed by (x,u'), in $\ell^\infty(\mathcal{X} \times \mathcal{U})$

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

as stochastic processes indexed by (x',u) in $\ell^\infty(\mathcal{X} \times \mathcal{U})$. The restriction on function g are as those specified in Proposition 4.

Corollary 3 (Large Sample Properties of Smoothed Curves)**.** *Under the hypotheses of Proposition 5, in* $\ell^\infty(\mathcal{Y}\mathcal{X})$,

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

and in $\ell^\infty(\mathcal{U}\mathcal{X})$,

$$\sqrt{n}(S\widehat{F}^{-1}(u|x) - SF^{-1}(u|x)) \Rightarrow S[f(F^{-1}(u|x)|x) \cdot D_{G(\cdot,x)}(F^{-1}(u|x),x)],$$

3. Theory of Probability Curves without Crossing

The same method used here to monotonically rearrange quantile curves also applies symmetrically to monotonically rearrange probability distribution curves, by a permutation of the roles of the quantile and probability spaces.

There are several situations where one might be faced with the problem of non-increasing probability curves. In an option pricing context, one could face situations similar to those considered in Ait-Sahalia and Duarte (2003) where a "risk-neutral distribution" is inferred from some market data and due to some noise it is not completely monotonic. In some situation, the probability distribution curve may have been obtained by some inverse transformation and local non-monotonicity came as an artefact of the regularization technique. In other situations, the particular smoothing technique may not respect monotonicity (see Hall et. al. (1999) for a statement of this problem and various proposed solutions). Here we propose to use the same method to rearrange the probability curves as we did for quantile curves.

Suppose we have $y \mapsto \hat{P}(y)$ as a candidate empirical probability distribution curve, which does not necessarily satisfy monotonicity. For the notational convenience we do not expose the conditional case, which however goes through with an exact parallel as above. Define the following quantile curve

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

In what follows further assume support of $P(y)$ is on $\mathcal{Y} \subset [0, +\infty)$, so that the second term drops down (otherwise it can be treated exactly in parallel to the first).

Consider also the inverse of the quantile curve is the rearranged probability curve

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

which is monotone. It should be clear that the quantities $\hat{Q}$ and $\hat{F}$ play a role which is exactly symmetric as the role played respectively by $\hat{F}$ and $\hat{Q}$ in the quantile case.

Therefore, if distribution curve (non necessarily monotonic) $P(y)$ can be estimated using an estimator $\hat{P}$ such that

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

in $\ell^{\infty}(\mathcal{Y})$, where G is a Gaussian process, one has in the monotonically correct case, that is when $P'\left(Q(u)\right) > 0$ for all $u \in [0,1]$, that

$$\sqrt{n}\left(\hat{Q}(u) - Q(u)\right) \Rightarrow \left(\frac{1}{P'(Q(u))}\right) G\left(Q(u)\right) \tag{3.1}$$

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

$$\sqrt{n}\left(\hat{F}\left(y\right) - F\left(y\right)\right) \Rightarrow G\left(y\right) \tag{3.2}$$

in $\ell^{\infty}(\mathcal{Y})$.

Results paralleling those of the previous section are also obtainable for the monotonically incorrect case:

$$\sqrt{n}\left(\hat{Q}(u) - Q(u)\right) \Rightarrow \sum_{k=1}^{K(u)} \frac{G\left(y_k\left(u\right)\right)}{|P'(y_k(u))|} \tag{3.3}$$

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

$$\sqrt{n}\left(\hat{F}\left(y\right) - F\left(y\right)\right) \Rightarrow \left(\sum_{k=1}^{K(u)} \frac{1}{|P'(y_k\left(u\right))|}\right)^{-1} \sum_{k=1}^{K(u)} \frac{G\left(y_k\left(t\right)\right)}{|P'(y_k\left(u\right))|}\Big|_{u=P(y)} \tag{3.4}$$

in $\ell^{\infty}(\mathcal{Y})$.

4. Illustrative Example: Engel Curves

To illustrate the application of our approach, we consider the estimation of expenditure curves. We use the original Engel (1857) data to estimate the relationship between food expenditure and annual household income. This is a classical data set in economics and is based on 235 budget surveys of 19th century working-class Belgium households (see also

Koenker, 2005). The data were originally presented by Ernst Engel to support his hypothesis that food expenditure constitutes a declining share of household income (Engel's Law).

Figure 3 reproduces Figure A.1 in Koenker (2005). It shows a scatterplot of the Engel data on food expenditure vs. household income. Superimposed on the plot are the $\{0.05, 0.1, 0.25, 0.75, 0.90, 0.95\}$ quantile regression curves as dashed lines, together with the median fit as a solid line. Here we can see that the quantile regression lines become closer as we approach to the origin of the graph, indicating a potential problem of quantile crossing for low values of income. This crossing problem of the Engel curves is more evident in Figure 4, which plots the quantile regression process of food expenditure as a function of the quantile index, instead of against the regressor. For low and very high quantiles of the income, the quantile regression process, plotted as a thick gray line, is clearly non monotone. Our rearrangement procedure fixes this undesirable feature of the quantile regression estimates producing monotonically increasing quantile functions, which are plotted with black lines in the figure. Moreover, the rearranged curves coincide with their quantile regression counterparts for the middle and upper values of income where there is no quantile-crossing problem.

Figure 5 illustrates how to perform quantile-uniform inference using the rearranged quantile curves. It plots simultaneous 90% confidence intervals for the conditional quantile process of food expenditure for two different values of income, the sample mean and the 1 percent sample percentile. The bands are constructed from both quantile regression, plotted in light gray, and rearranged quantile curves, plotted in dark gray. A grid of quantiles $\{0.10, 0.11, ..., 0.90\}$ and 500 bootstrap repetitions are used to obtain the quantile regression bands. For the rearranged curves, we assume that the estimand of the quantile regression process is monotone and derive the bands using Lemma 1. As a result of this lemma, bootstrap critical values for quantile regression are also valid for the rearranged curves, what greatly reduces the computation time since we do not need to rearrange the curves in each bootstrap iteration. The figure shows that even for the lower value of income there is important overlapping between the bands for all the quantiles considered. This observation points towards the maintained assumption that in this case the lack of monotonicity of the quantile regression process is caused by sampling error due to the small sample size.

Appendix A. Proofs

A.1. **Proof of Proposition 1.** At first note that the distribution of $\mathcal{Y}_x$ has no atoms.

$$Pr[Y_x = y] = Pr[Q(U, x) = y] = Pr[U \in \{ \text{ roots of } Q(u, x) = y\}] = 0,$$

since the number of roots of $Q(u, x) = y$ is finite under (a) and (b). Further, by assumptions (a) and (b) the number of critical values of $Q(x, u)$ is clearly finite, hence the claim (1) follows.

For any regular y, letting $u_0(y, x) := 0$

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

Note that the signs of of $Q'(u, x)$ alternate over consecutive $u_k(y, x)$'s, determining whether $1\{Q(y, x) \leq y\} = 1$ on the interval $[u_{k-1}(y, x), u_k(y, x)]$ (with $u_0(y, x) := 0$), thus the first term in the expression above simplifies to $\sum_{k=0}^{K(y,x)-1} 1(Q'(u_{k+1}, x) \geq 0)(u_{k+1}(y, x) - u_k(y, x))$; while the last term simplifies to $1(Q'(u_{K(y,x)}, x) \leq 0)(1 - u_{K(y,x)})$. A further simplification yields the expression given in the proposition.

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

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

Combining this fact with claim (2), we get the value of the derivative as in claim (3).

To show absolute continuity of $F(y|x)$ with $f(y|x)$ being the Radon-Nykodym derivative, it suffices to show that for each y', $\int_{\infty}^{y'} f(y|x)dx = \int_{-\infty}^{y'} dF(y|x)$, cf. Theorem 31.8 in (**?**). 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 $\mathcal{Y}_x^t = \mathcal{Y}_x \backslash V_t^x$. Then, $\int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\} f(y|x)dx = \int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\} dF(y|x)$. As $t \to 0$, $\int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\} dF(y|x) \uparrow \int_{-\infty}^{y'} dF(y|x)$ by the set of critical critical points $\mathcal{Y}_x \backslash \mathcal{Y}_x^*$ being finite and having mass zero under $F(y|x)$. Therefore, $\int_{-\infty}^{y'} 1\{y \in \mathcal{Y}_x^t\} f(y|x)dx \uparrow \int_{-\infty}^{y'} f(y|x)dx = \int_{-\infty}^{y'} dF(y|x)$.

Claim (4) follows by noting that at the regions where $s \to Q(s,x)$ is increasing and one-to-one, and one has $F(y|x) = \int_{Q(s,x) \le y} ds = \int_{s \le Q^{-1}(y,x)} ds = Q^{-1}(y,x)$. Inverting the equation $u = F(F^{-1}(u|x)|x) = Q^{-1}(F^{-1}(u|x),x)$ 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 $\inf\{y : P(\alpha + \beta Y_x \le y) \ge u\} = \alpha + \beta F^{-1}(u|x)$.

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

Claim (7). The proof of continuity of $F(y|x)$ is subsumed in the proof of Proposition 3. So 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. $y_t = F^{-1}(y_t|x_t)$, converges to y, by the standard argument. □

A.2. **Proof of Propositions 2-5.** In the proofs that follow we will repeatedly use the equivalence of the continuous convergence and the uniform convergence:

Lemma 1. *Let D and D' be complete separable metric spaces, and D is 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 $f_n(x_n) \to f_0(x)$.*

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 $u \in B_\epsilon(u_k(y,x))$ for all small enough $t \ge 0$

$$1\{Q(x,u) + th_t(u,x) \le y\} \le 1\{Q(x,u) + t(h(u_k(y,x),x) + \delta) \le y\},$$

for all $k \in 1,...,K(y,x)$, and for all $u \notin \cup_k B_\epsilon(u_k(y,x))$

$$1\{Q(x,u) + th_t(u,x) \le y\} = 1\{Q(x,u) \le y\}.$$

Therefore,

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

but by the change of variables this is equal to

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

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.

For t small enough, we have that

$$J_k\cap[y-t[h(u_k(y,x),x)+\delta],y]=[y-t[h(u_k(y,x),x)+\delta],y],$$

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 less 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.

It follows similarly that the result holds uniformly in $(x,y)\in K$, a compact subset of $\mathcal{YX}^*$, as on such a set the functions $(x,y)\to|Q'(u_k(y,x),x)|$ is uniformly continuous and so is $(x,y)\mapsto K(y,x)$. To see the latter, note since K excludes a neighborhood of critical points $(\mathcal{Y}\setminus\mathcal{Y}_x^*,x\in\mathcal{X})$, K is the union of a finite number of compact sets $(K_1,...,K_M)$ over any of which the function $K(y,x)$ is constant: that is, $K(y,x)=k_j$ for some integer $k_j>0$, for all $(x,y)\in K_j$, for all $j\le M$. $\square$

Proof of Proposition 3. For a fixed x, Proposition 2, Step 1 of the proof below and an application of the Hadamard differentiability of the quantile operator shown by [References] give the result. The proof below, particularly Step 2, accommodates uniformity over $x\in\mathcal{X}$.

Step 1. Let K be the subset of the set $\mathcal{YX}^*$. Let (y_t, x_t) be a convergent sequence in this set, to a point, say (y, x). Then for every $(y_t, x_t) \to (y, x)$ we have for $\epsilon_t := \|h_t\|_\infty + \|Q(x_t, \cdot) - Q(x, \cdot)\|_\infty + |y_t - y| \to 0$ that

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

where the last step follows from the absolute continuity of $F(y|x)$, the distribution function of $Q(U, x)$. This also verfies that $F(y|x)$ is continuous in (y, x). This implies uniform convergence of $F(y|x, h_t)$ to $F(y, x)$ by Lemma 3. This in turn implies 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,

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

using Step 1 and Proposition 2, and using the continuity properties of $D_h(y, x)$. Further, uniformly over K, by Taylor expansion

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

and, as shown below,

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

Observe that the left hand side of (A.5) equals that of (A.4) plus that of (A.3). The result follows.

It only remains to show equation (A.5) holds uniformly in K. Note that for any right-continuous cdf F, we have that $u \leq F(F^{-1}(u)) \leq u + F(F^{-1}(u)) - F(F^{-1}(u)-)$, and for any continuous, strictly increasing cdf F, we have that $F(F^{-1}(u)) = u$. Therefore, write

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

where in (1) we used that $F(F^{-1}(u|x,h_t)|x) = F(F^{-1}(u|x,h_t)-|x)$ due to $F(y|x)$ continuous and strictly increasing in y, and in (2) we used Proposition 2. □

The following lemma due to Pratt (1960) will be useful.

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 $\int f_n \to \int f$.*

A.3. **Proof of Lemma 2.** This lemma was proven by Pratt (1960). □

Lemma 3 (Boundedness and Integrability Properties)**.** *We have that for all x and y:*

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

and

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

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

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

and for some constant C, for any $y_t \to y \in \mathcal{Y}_0$,

$$\Delta(y_t,x) \to 2\|h\|_\infty f(y|x) \text{ for a.e } x \text{ in } \mathcal{X}_0 \text{ and } \int_{\mathcal{X}_0} \Delta(y_t,x,t)dx \to \|h_\infty\|C.$$

A simple consequence of the uniform integrability and of Propositions 2 and 3 is the following result:

A.4. **Proof of Lemma 3.** To show the first result note that

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

immediately follows from the equivariance property noted in Claim 5 of Proposition 1.

By Fubini,

$$\int_{\mathcal{Y}} \Delta(y,x_t,t)dy = \int_0^1 \underbrace{\left(\int_{\mathcal{Y}} \frac{1\{|Q(x_t,u)-y| \leq t\|h_t\|_\infty\}}{t}dy\right)}_{=:\ f_t(u)} du. \tag{A.9}$$

But $f_t(u)$ is majorized by $\|h_t\|_\infty$, that is $f_t(u) \leq \|h_t\|_\infty$, and for almost every u, for large enough t, $f_t(u) = \|h_t\|_\infty$ which converges to $\|h\|_\infty$ and, trivially, $\int_0^1 \|h_t\|_\infty du \to \|h\|_\infty$. By Lemma 2 the result follows the right hand side of (A.9) converges to $\|h\|_\infty$.

Let $m(q|u)$ denote the density of $Q(x,u)$, which by hypothesis is bounded by some constant m. By Fubini,

$$\int_{\mathcal{X}_0} \Delta(y_t,x,t)dx = \int_0^1 \int_{\mathcal{X}_0} \frac{1\{|Q(x_t,u)-y_t| \leq t\|h_t\|_\infty\}}{t}dxdu \tag{A.10}$$

$$= \int_0^1 \underbrace{\left(\int_{\mathcal{X}_0} \frac{1\{|q-y_t| \leq t\|h_t\|_\infty\}}{t}m(q|u)dq\right)}_{=:f_t(u)} duLeb(\mathcal{X}_0) \tag{A.11}$$

But since $m(x|u)$ is the density of $Q(x,u)$, for a.e. u, $f_t(u) \to m(y|u)\|h_t\|_\infty$ and $f_t(u) \leq m\|h_t\|_\infty$ where, trivially, $m\|h_t\|_\infty$ converges to $m\|h\|_\infty$ as well as $\int m\|h_t\|du \to m\|h\|_\infty$.

Conclude by Lemma 2 that the right hand side of (A.11) converges to $C\|h\|_\infty$. □

A.5. **Proof of Proposition 4.** Define $m_t(y|x,y') := 1(y \leq y')g(y,x)D_{h_t}(y,x,t)$ and $m(y|x,y') := 1(y \leq y')g(y,x)D_h(y,x)$. To show claim (1), we need to show that for any $y'_t \to y'$ and $x_t \to x$

$$\int_{\mathcal{Y}} m_t(y|x_t,y'_t)dy \to \int_{\mathcal{Y}} m(y|x,y')dy, \tag{A.12}$$

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 which integral converges to a finite number by Lemma 3. Moreover, by Proposition 2, for almost every y we have, $m_t(y|x_t,y_t) \to m(y|x,y')$. Conclude that (A.12) holds by Lemma 2.

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

$$\int_{\mathcal{Y}} m(y|x, y')dy \to \int_{\mathcal{Y}} m(y|x, y')dy, \tag{A.13}$$

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.13) holds by Lemma 2.

The proof of claim (2) follows similarly to the proof of claim (1).

[The part that follows can be omitted because it is also similar.]

To show claim (3), define $m_t(u|x, u') = 1\{u \leq u'\}g(u, x)\tilde{D}_{h_t}(u, x)$ and $m(u|x, u') = 1\{u \leq u'\}g(u, x)\tilde{D}_h(u, x)$. To show claim (1), we need to show that for any $u'_t \to u'$ and $x_t \to x$

$$\int_{\mathcal{U}} m_t(u|x_t, u'_t)dy \to \int_{\mathcal{U}} m(u|x, u')du, \tag{A.14}$$

and that the limit is continuous in (x, u'). 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, and which integral converges to that of $g(u, x)\|h\|_\infty$ by the dominated convergence theorem. Moreover by Proposition 2, for almost every u we have $m_t(u|x_t, u'_t) \to m(u|x, u')$. Conclude that (A.14) holds by Lemma 2.

In order to check continuity, we need to show that for any $u'_t \to u'$ and $x_t \to x$

$$\int_{\mathcal{U}} m(u|x, u'_t)du \to \int_{\mathcal{U}} m(u|x, u')du, \tag{A.15}$$

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 and which integral converges by dominated convergence theorem. Conclude that (A.15) holds by Lemma 2.

The proof of claim (4) follows similarly to the proof of claim (3). □

A.6. **Proof of Proposition 5.** This simply follows by the delta method (Van der Vaart, 1998). Here we remind ourselves what it is. To show the first part, consider the map $g_n(h) = n^{-1/2}(F(y|x, n^{-1/2}h) - F(y|x)$. The sequence of maps satisfies $g_{n'}(h_{n'}) \to D_h(x, y)$

for every subsequence $h_{n'} \to h$ in $\ell^\infty(\mathcal{UX}^*)$, where h is continuous. It follows by the extended continuous mapping theorem that $g_n(\sqrt{n}(\hat{q}(u,x) - Q(u,x))) \Rightarrow D_{G(\cdot,x)}(y,x)$. Conclude similarly for the second part. □

A.7. **Proof of Proposition 6.** This follows by the delta method. □

References

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

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

[3] Doss, Hani; Gill, Richard D. An elementary approach to weak convergence for quantile processes, with applications to censored survival data. J. Amer. Statist. Assoc. 87 (1992), no. 419, 869–877.

[4] Dudley, Richard M.; Norvaisa, Rimas Differentiability of six operators on nonsmooth functions and p-variation. With the collaboration of Jinghua Qian. Lecture Notes in Mathematics, 1703. Springer-Verlag, Berlin, 1999. viii+277 pp.

[5] Fan, J., Yao, Q., and Tong, H., "Estimation of conditional densities and sensitivity measures in. nonlinear dynamical systems," *Biometrika*, 83, pp. 189–206, 1996.

[6] Gill, Richard D.; Johansen, Soren. A survey of product-integration with a view toward application in survival analysis. Ann. Statist. 18 (1990), no. 4, 1501–1555.

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

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

[9] Koenker, Roger Confidence intervals for regression quantiles. Asymptotic statistics (Prague, 1993), 349–359, Contrib. Statist., Physica, Heidelberg, 1994

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

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

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

[13] Pratt, J.W.: On interchanging limits and integrals. *Ann. Math. Stat.* 31, 74 (1960)

[14] Rudin, W., *Functional Analysis*, 2nd edition, New York: McGraw-Hill, Inc., 1991.

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

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

[17] van der Vaart, A. W. Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics, 3. Cambridge University Press, Cambridge, 1998. xvi+443 pp

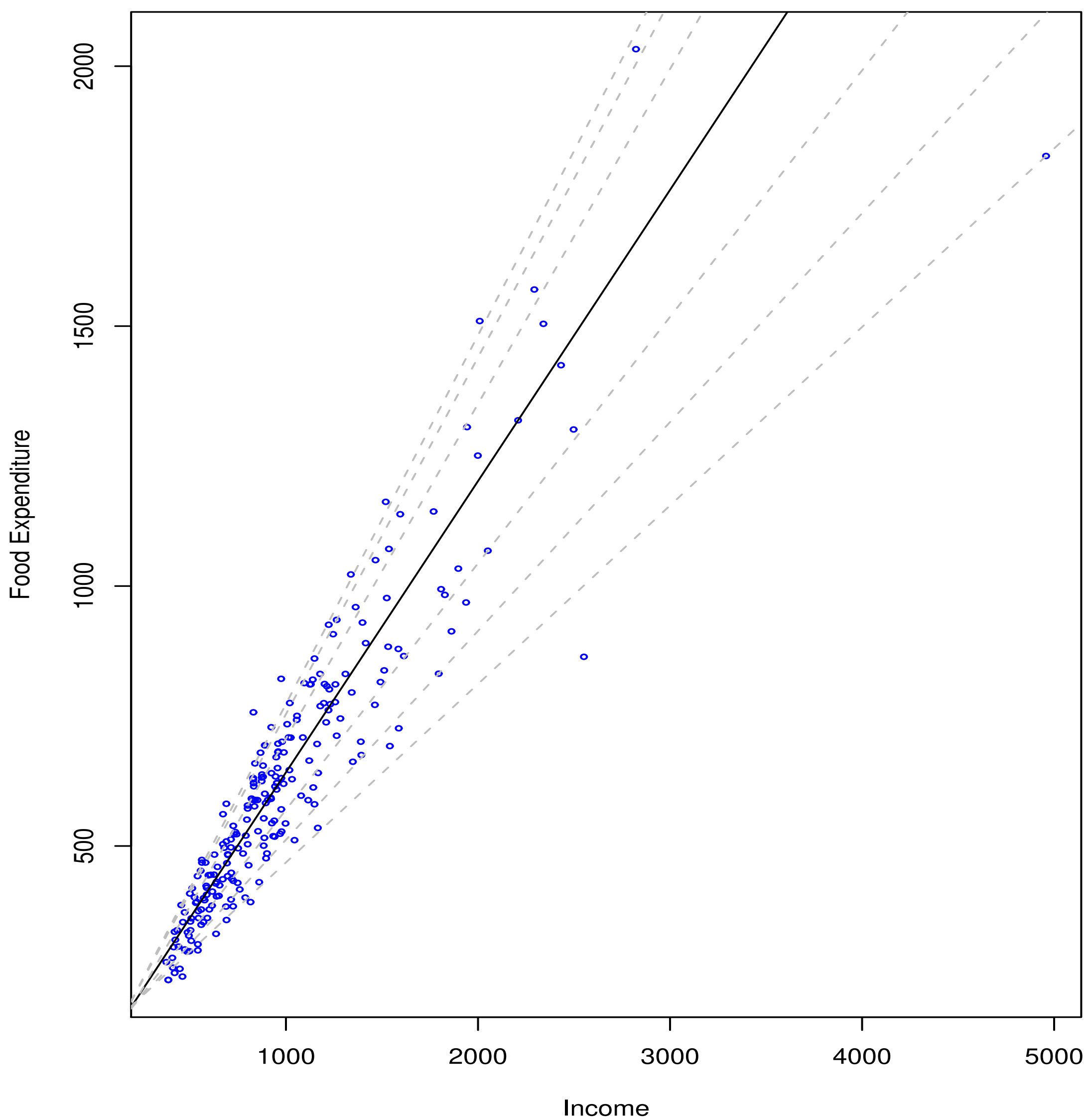

FIGURE 3. Scatterplot and quantile regression fit of the Engel food expenditure data. The plot shows a scatterplot of the Engel data on food expenditure vs. household income for a sample of 235 19th century working-class Belgium households. Superimposed on the plot are the $\{0.05, 0.10, 0.25, 0.75, 0.90, 0.95\}$ quantile regression curves as dashed lines, and the median fit as a solid line.

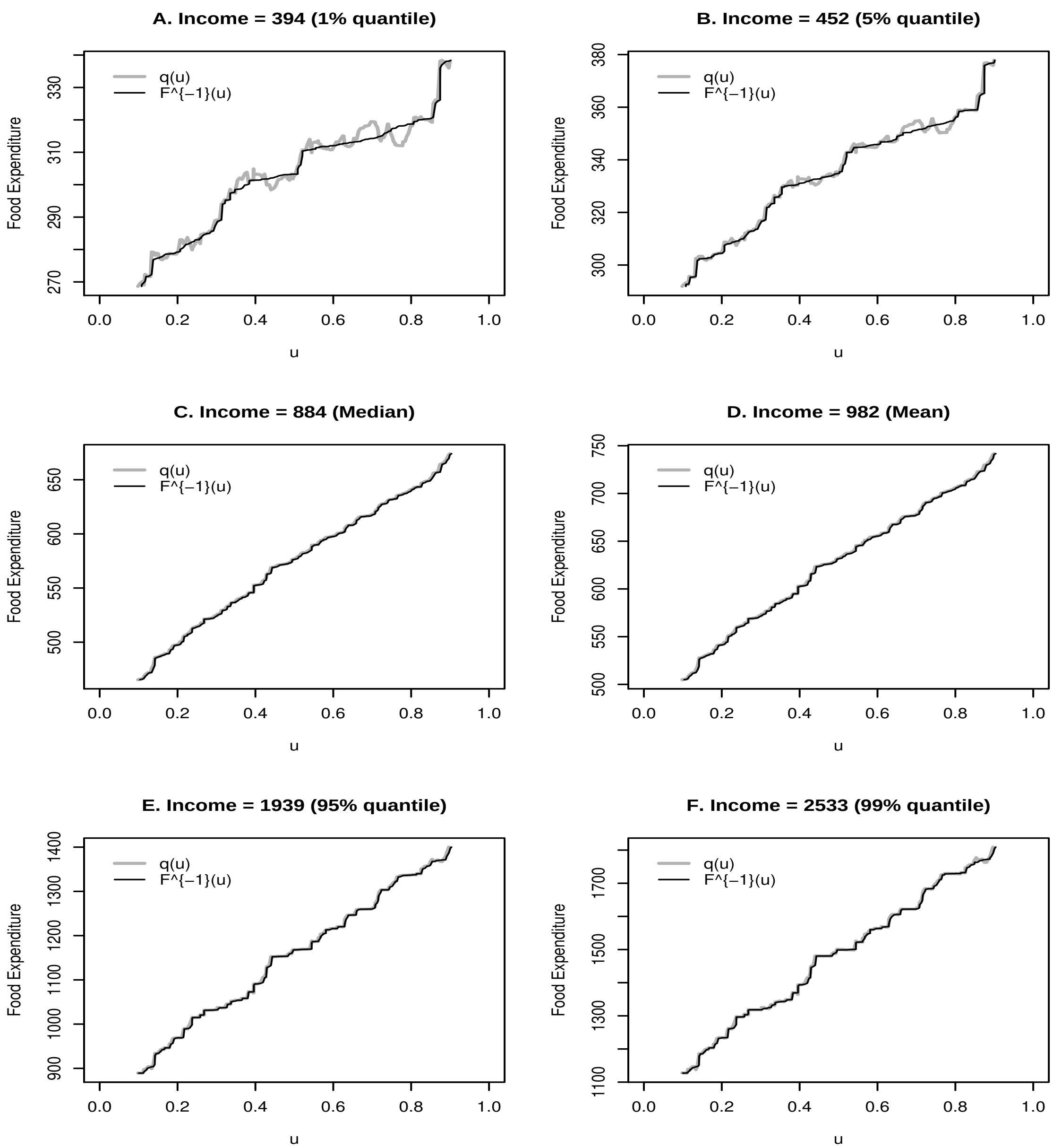

Figure 4. Quantile regression processes and rearranged quantile processes for the Engel food expenditure data. The plot shows uncorrected and rearranged quantile regression estimates of the conditional quantile function of food expenditure given income as a function of the quantile index, for different values of the distribution of income. Quantile regression estimates are plotted with a thick gray line, whereas the corresponding monotonized curves are plotted in black.

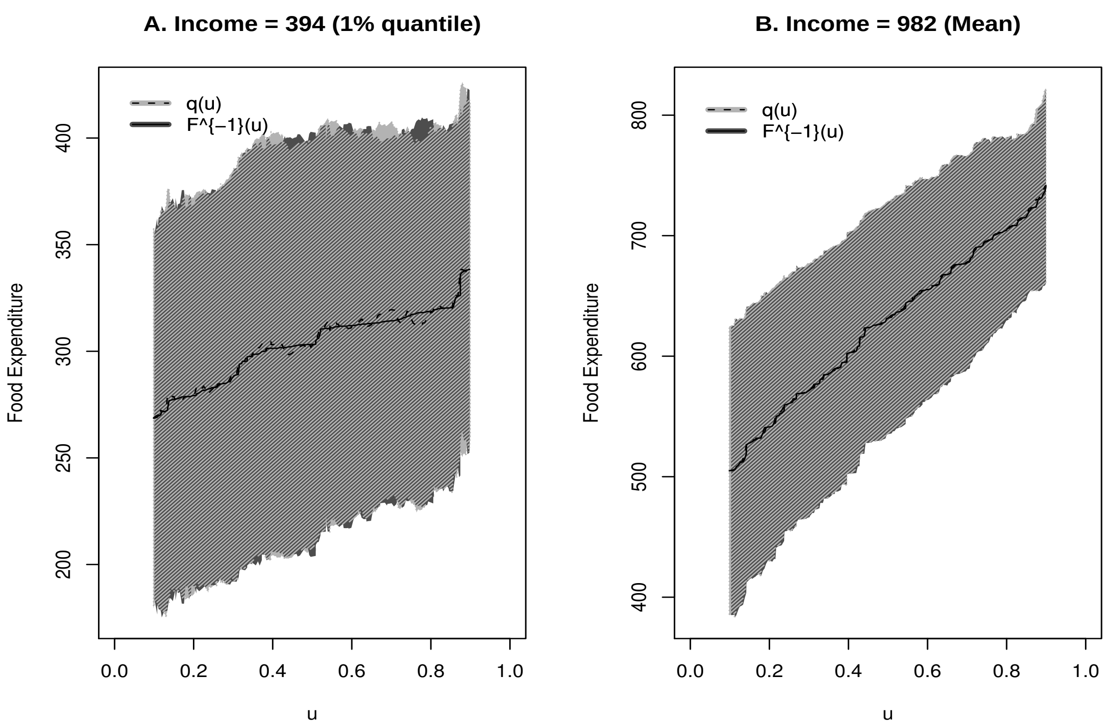

FIGURE 5. Simultaneous 90% confidence bands for quantile regression processes and rearranged quantile processes for the Engel food expenditure data. The figure shows quantile-uniform bands for uncorrected and rearranged quantile regression estimates of the conditional quantile function of food expenditure given income as a function of the quantile index, for two different values of the distribution of income. Quantile regression bands are plotted in light gray, whereas the corresponding monotonized bands are in dark gray.