Quantile Curves without Crossing

Victor Chernozhukov — MIT
Iván Fernández-Val — Boston University
Alfred Galichon — Ecole Polytechnique

Déjeuner-Séminaire d'Economie

Ecole polytechnique, November 12 2007

Aim of the talk

- Present the methodology and applications of **Quantile Regression**

Aim of the talk

- Present the methodology and applications of **Quantile Regression**
- Indentify and correct a **common problem** for several estimation procedures Quantile Regression

Aim of the talk

- Present the methodology and applications of **Quantile Regression**
- Indentify and correct a **common problem** for several estimation procedures Quantile Regression
- Study the impact of the **correction** on the estimator of the measure

This talk:

1. Quantile Regression and Two applications
   - The Value-at-Risk
   - Engle curves
2. Rearranging Quantile Curves
   - The crossing problem
   - The rearrangement operation
   - Overview of the literature
3. Properties and illustrations
   - Graphical properties
   - Approximation properties
   - Large-sample properties properties

The Value-at-Risk

Aim: measure & manage risk of portfolio's contingent loss Y.

The Value-at-Risk

Aim: measure & manage risk of portfolio's contingent loss Y.

- $VaR_\alpha(Y) =$ smallest capital amount to cover losses in $\alpha\%$ cases...

The Value-at-Risk

Aim: measure & manage risk of portfolio's contingent loss Y.

- $VaR_\alpha(Y)$ = smallest capital amount to cover losses in $\alpha\%$ cases...
- is robust to tail behaviour (eg. more than variance)

The Value-at-Risk

Aim: measure & manage risk of portfolio's contingent loss Y.

- $VaR_\alpha(Y)$ = smallest capital amount to cover losses in $\alpha\%$ cases...
- is robust to tail behaviour (eg. more than variance)
- has become a market standard for market risk measurement (Basle II 1st pillar)

The Value-at-Risk

Aim: measure & manage risk of portfolio's contingent loss Y.

- $VaR_\alpha(Y)$ = smallest capital amount to cover losses in $\alpha\%$ cases...
- is robust to tail behaviour (eg. more than variance)
- has become a market standard for market risk measurement (Basle II 1st pillar)
- is however criticized among both practitioners and academics

VaR and coherent measures

Many authors have pointed out the shortfalls of VaR as a management tool.

VaR and coherent measures

Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** "Merging does not make risk appear larger" (Artzner et al., 1999)

VaR and coherent measures

Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** “Merging does not make risk appear larger” (Artzner et al., 1999)
  $\Rightarrow$ Fails to be satisfied by VaR.

VaR and coherent measures

Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** “Merging does not make risk appear larger” (Artzner et al., 1999)
  $\Rightarrow$ Fails to be satisfied by VaR.
- Alternative “coherent” measures have been proposed.
  - **Expected Shortfall:** average loss beyond given level

VaR and coherent measures

Many authors have pointed out the shortfalls of VaR as a management tool.

- Desirable property: **subadditivity** "Merging does not make risk appear larger" (Artzner et al., 1999)
  $\Rightarrow$ Fails to be satisfied by VaR.
- Alternative "coherent" measures have been proposed.
  - **Expected Shortfall:** average loss beyond given level
  - **distortion measures:** weighted loss average; higher weights toward higher losses.

VaR and Quantile estimation

Recall the definition of a quantile function.

VaR and Quantile estimation

Recall the definition of a quantile function.
Let $F_Y(y) = Pr(Y \leq y)$: distribution function of Y, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.

VaR and Quantile estimation

Recall the definition of a quantile function.
Let $F_Y(y) = Pr(Y \leq y)$: distribution function of Y, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.

- The VaR is is precisely $VaR_\alpha(Y) = Q_Y(\alpha)$.

VaR and Quantile estimation

Recall the definition of a quantile function.
Let $F_Y(y) = Pr(Y \leq y)$: distribution function of Y, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.

- The VaR is is precisely $VaR_\alpha(Y) = Q_Y(\alpha)$.
- Distorsion measures can be written $\rho(Y) = \int_0^1 \varphi(u) Q_Y(u) du$, where φ is increasing.

VaR and Quantile estimation

Recall the definition of a quantile function.
Let $F_Y(y) = Pr(Y \leq y)$: distribution function of Y, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.

- The VaR is is precisely $VaR_\alpha(Y) = Q_Y(\alpha)$.
- Distorsion measures can be written $\rho(Y) = \int_0^1 \varphi(u) Q_Y(u) du$, where φ is increasing.
  Example. Expected Shortfall: $\varphi(u) = 1\{u \geq \alpha\}/(1-\alpha)$.

VaR and Quantile estimation

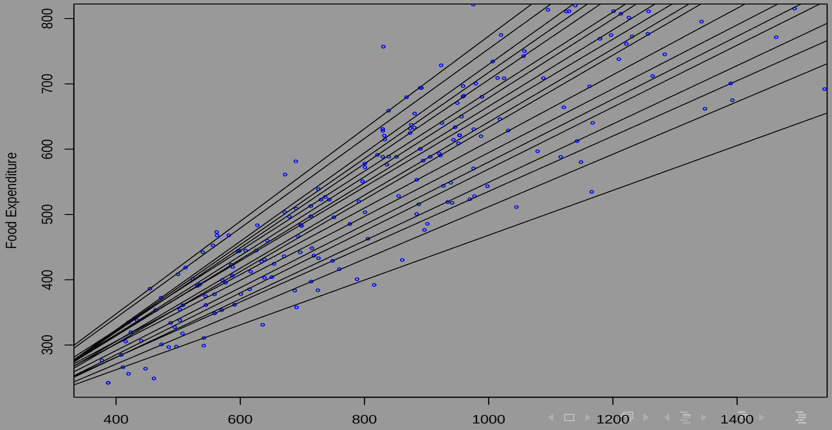
Recall the definition of a quantile function.
Let $F_Y(y) = Pr(Y \leq y)$: distribution function of Y, and $Q_Y(u) = F_Y^{-1}(u)$ quantile function.

- The VaR is is precisely $VaR_\alpha(Y) = Q_Y(\alpha)$.
- Distorsion measures can be written $\rho(Y) = \int_0^1 \varphi(u) Q_Y(u) du$, where φ is increasing.
  Example. Expected Shortfall: $\varphi(u) = 1\{u \geq \alpha\}/(1-\alpha)$.
- Thus estimation of VaR/distorsion measure requires estimation of **quantile function**.

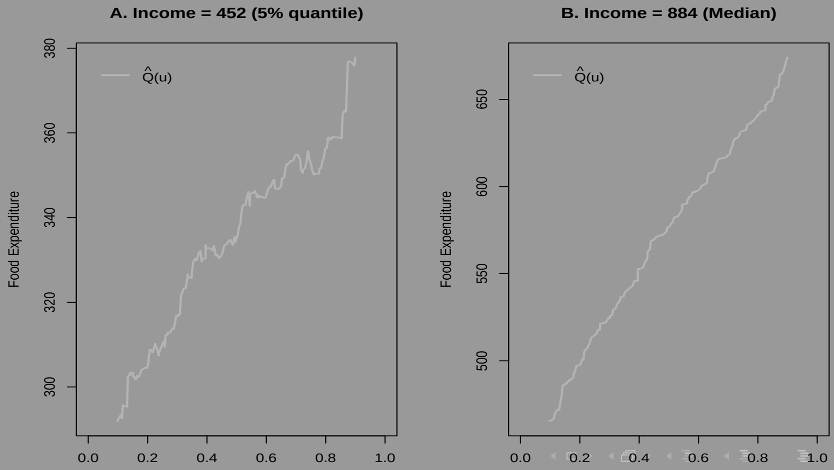
Engel Curves

- Y response variable, X regressor, the u-th quantile of Y given $X = x$
$$Q_0(u|x) = \inf\{y : F(y|x) \geq u\}.$$
- QR estimates a linear approximation to the conditional quantile
$$Q(u|x) = x'\beta(u)$$
- QR fits for different quantiles provide a description of the **entire conditional distribution**
Example: Buchinsky (1994) uses QR to describe the evolution of the wage distribution in the U.S.
- Here, y=food expenditure ; x = household income.

Engel Curves by Quantile Regression



Engel Curves by Quantile Regression



Quantile regression

Given covariate X (information at period t), estimate QR model

$$Q_{Y|X}(u|x) = x'\beta(u).$$

Quantile regression

Given covariate X (information at period t), estimate QR model

$$Q_{Y|X}(u|x) = x'\beta(u).$$

- estimate using

$$\hat{\beta}(u) = \arg\min_{\beta \in \mathbb{R}^d} \sum_{k=1}^{n} u\left(Y_k - X_k'\beta\right)^+ + (1-u)\left(Y_k - X_k'\beta\right)^-$$

Quantile regression

Given covariate X (information at period t), estimate QR model

$$Q_{Y|X}\left(u|x\right) = x'\beta\left(u\right).$$

- estimate using

$$\hat{\beta}\left(u\right) = \arg\min_{\beta\in\mathbb{R}^d} \sum_{k=1}^{n} u\left(Y_k - X_k'\beta\right)^+ + \left(1-u\right)\left(Y_k - X_k'\beta\right)^-$$

- Autoregressive case- the covariate X captures past information. Several models: Quantile Autoregression, CaViaR, Dynamic Quantile...

The crossing problem

In the QR procedure, nothing ensures that $\hat{Q}_{Y|X}(u|x) = x'\hat{\beta}(u)$ be increasing in u.

The crossing problem

In the QR procedure, nothing ensures that $\hat{Q}_{Y|X}(u|x) = x'\hat{\beta}(u)$ be increasing in u.

- In fact it may be non-monotonic if
  - the QR model is misspecified, or
  - the sample size is small

The crossing problem

In the QR procedure, nothing ensures that $\hat{Q}_{Y|X}(u|x) = x'\hat{\beta}(u)$ be increasing in u.

- In fact it may be non-monotonic if
  - the QR model is misspecified, or
  - the sample size is small
- VaR context: a higher confidence level would require less capital!

The crossing problem

In the QR procedure, nothing ensures that $\hat{Q}_{Y|X}(u|x) = x'\hat{\beta}(u)$ be increasing in u.

- In fact it may be non-monotonic if
  - the QR model is misspecified, or
  - the sample size is small
- VaR context: a higher confidence level would require less capital!
- can have adverse managemental effects / lack of trust for the tool...

A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

- draw $U \sim \mathcal{U}[0,1]$ and take $Y_x := \widehat{Q}(U|x)$ (bootstrap).

A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

- draw $U \sim \mathcal{U}[0,1]$ and take $Y_x := \widehat{Q}(U|x)$ (bootstrap).
- take distribution function $\widehat{F}(y|x) = Pr(Y_x \leq y)$, ie.

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

A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

- draw $U \sim \mathcal{U}[0,1]$ and take $Y_x := \widehat{Q}(U|x)$ (bootstrap).
- take distribution function $\widehat{F}(y|x) = Pr(Y_x \leq y)$, ie.

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

- invert to recover *rearranged quantile* $\widehat{F}^{-1}(u|x)$

A proposed solution

Suppose we use the (flawed) estimator $\hat{Q}_{Y|X}(u|x)$ to simulate $Y|X = x$.

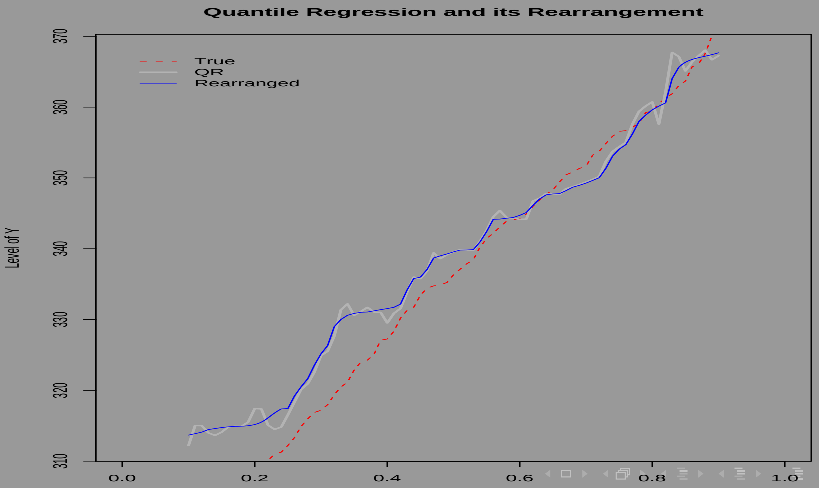
- draw $U \sim \mathcal{U}[0,1]$ and take $Y_x := \widehat{Q}(U|x)$ (bootstrap).
- take distribution function $\widehat{F}(y|x) = Pr(Y_x \leq y)$, ie.

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

- invert to recover *rearranged quantile* $\widehat{F}^{-1}(u|x)$

If the original estimator $\hat{Q}_X(u|x)$ is monotonic, then $\widehat{F}^{-1}(u|x)$ coincides with it.

The rearrangement: illustration



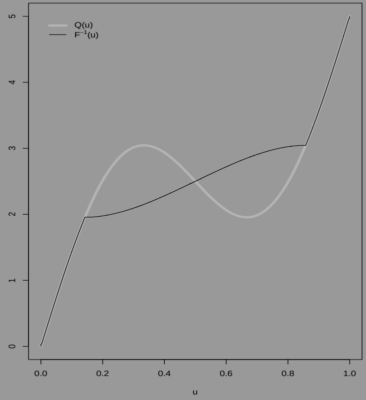
Literature review

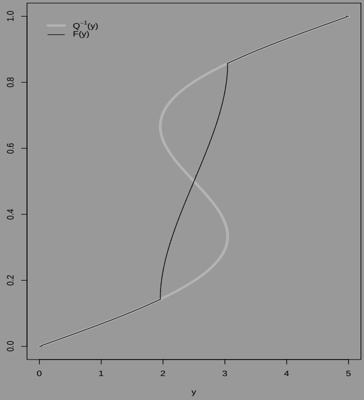
- **Quantile Regression:** Koenker & Bassett (1978). Dynamic, autoregressive context, VaR: Chernozhukov & Umantsev (2001), Koenker & Xiao (2006), Engle & Manganelli (2007), Gourieroux & Jasiak (2007).
- **Increasing rearrangement:** Hardy, Littlewood & Polya (1930's), Mossino & Temam (1979). In Statistics: Fougeres (1997), C F-V & G (2006), Dette, Neumeyer, and Pilz (2006).
- **Other monotonization procedure:** location-scale model of He (1997). Dynamic Quantile Model, Gourieroux & Jasiak (2007). Constraint optimization Koenker & Ng (2005).

Analytical Properties

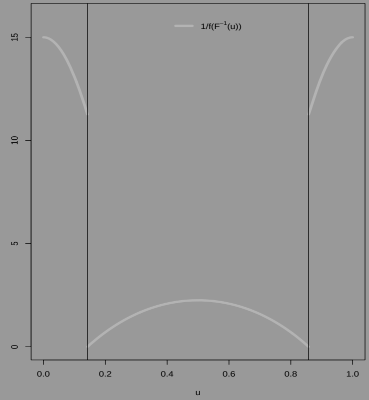
- As an example, take $Q(u)$ to be a non-monotone function of u - slope changes sign twice in $[0, 1]$.
- Rearranged curve is monotonically increasing and coincides with $Q(u)$ for points where $Q^{-1}(y)$ is uniquely defined.
- The derivative of the rearranged curve is a proper density function, continuous at the regular values of $Q(u)$.

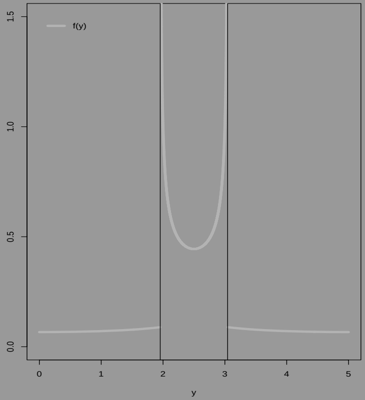
Analytical Properties: Example





Analytical Properties: Example





Estimation Properties

Denote $Q_0(u|x)$: true conditional quantile curve

Estimation Properties

Denote $Q_0(u|x)$: true conditional quantile curve

- For $p \in [1, \infty]$, rearrangement inequality:

$$\int_0^1 |Q_0(u|x) - \widehat{F}^{-1}(u|x)|^p du \leq \int_0^1 |Q_0(u|x) - x'\widehat{\beta}(u)|^p du.$$

Estimation Properties

Denote $Q_0(u|x)$: true conditional quantile curve

- For $p \in [1, \infty]$, rearrangement inequality:

$$\int_0^1 |Q_0(u|x) - \widehat{F}^{-1}(u|x)|^p du \leq \int_0^1 |Q_0(u|x) - x'\widehat{\beta}(u)|^p du.$$

- This property is **independent** of the sample size (holds in population).

Estimation Properties

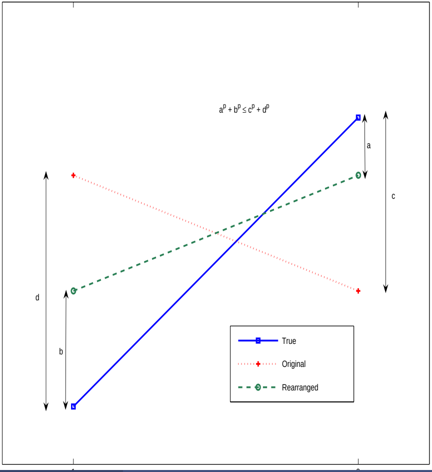
Denote $Q_0(u|x)$: true conditional quantile curve

- For $p \in [1, \infty]$, rearrangement inequality:

$$\int_0^1 |Q_0(u|x) - \widehat{F}^{-1}(u|x)|^p du \leq \int_0^1 |Q_0(u|x) - x'\widehat{\beta}(u)|^p du.$$

- This property is **independent** of the sample size (holds in population).
- Rearranged quantile curves have a smaller estimation error than the original curves whenever the latter are not monotone.

Illustration



Statistical properties (large sample)

- Fix an x, and suppose

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

where G is a Gaussian process. The population curve $u \to x'\beta(u)$ is assumed to be increasing.

Statistical properties (large sample)

- Fix an x, and suppose

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

  where G is a Gaussian process. The population curve $u \rightarrow x'\beta(u)$ is assumed to be increasing.
- Recall $\widehat{F}(y|x) = \int_0^1 1\{x'\hat{\beta}(u) \leq y\}du$. One has

Statistical properties (large sample)

- Fix an x, and suppose

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

where G is a Gaussian process. The population curve $u \to x'\beta(u)$ is assumed to be increasing.

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

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

Statistical properties (large sample)

- Fix an x, and suppose

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

where G is a Gaussian process. The population curve $u \to x'\beta(u)$ is assumed to be increasing.

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

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

- For quantiles, one has

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow x'G(u) \text{ in } \ell^{\infty}((0,1))$$

Statistical properties (large sample)

- Fix an x, and suppose

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

where G is a Gaussian process. The population curve $u \to x'\beta(u)$ is assumed to be increasing.

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

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

- For quantiles, one has

$$\sqrt{n}(\widehat{F}^{-1}(u|x) - F^{-1}(u|x)) \Rightarrow x'G(u) \ \text{ in } \ \ell^{\infty}((0,1))$$

(same asymptotic limit as for the original curve $u \mapsto x'\hat{\beta}(u)$).

Statistical properties, comments

- The rearranged curve has the same asymptotic error term as the original curve!...

Statistical properties, comments

- The rearranged curve has the same asymptotic error term as the original curve!...
- Not incompatible with finite sample properties

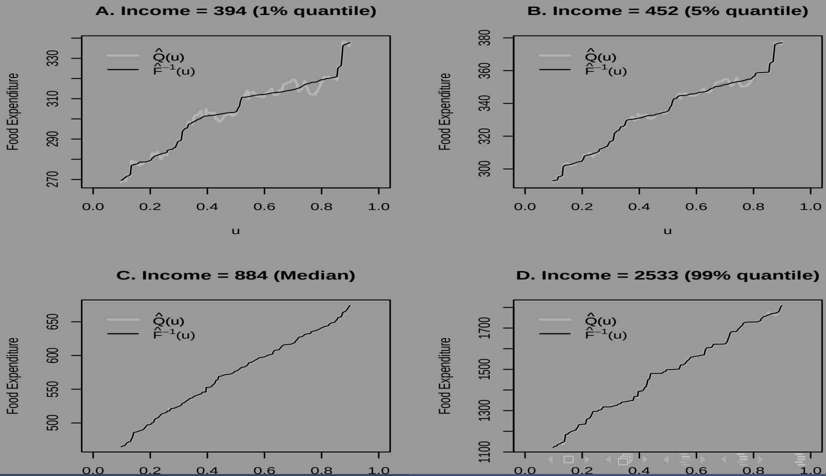
Statistical properties, comments

- The rearranged curve has the same asymptotic error term as the original curve!...
- Not incompatible with finite sample properties
- Convenient for testing purposes, as it does not modify the asymptotic properties of a test, while improving approximation in finite sample.

Empirical Application: Engel Curves

- We use the original Engel (1857) data to estimate the relationship between food expenditure and annual household income.
- Data set is based on 235 budget surveys of 19th century working-class Belgium households .
- Plot of quantile regression process (as a function of u) shows quantile-crossing for 5% percentile of income. No crossing problem for the sample median of income.
- Rearrangement procedure produces monotonically increasing curves - coincides with QR for the median of income.

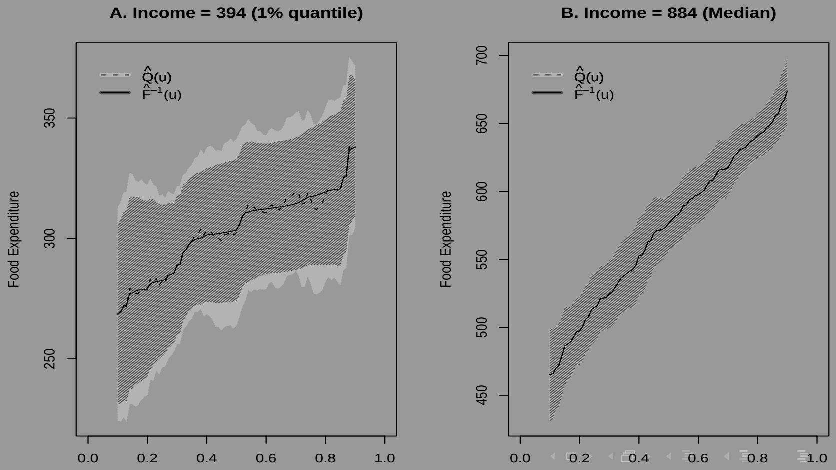
Empirical Application: Engel Curves



Uniform Inference: Engel Curves

- **Quantile-uniform inference** can be performed using the rearranged quantile curves.
- Next figure plots simultaneous 90% confidence intervals for the conditional quantile process of food expenditure for two different values of income, the sample median and the 1 percent sample percentile.
- Bands for QR are obtained by **bootstrap** using 500 repetitions and a grid of quantiles $\{0.10, 0.11, ..., 0.90\}$.
- Bands for rearranged curves are constructed assuming that estimand of QR is monotonically correct.
- Rearranged bands lie within QR bands - points towards lack of monotonicity due to small sample size.

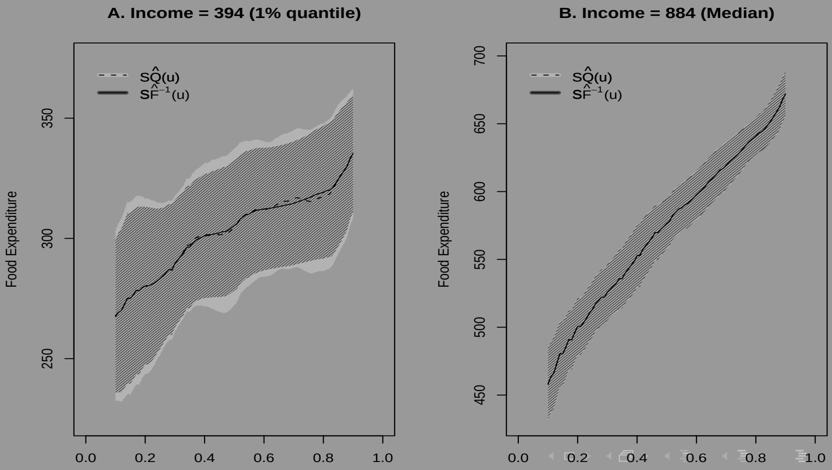
Uniform Inference: Engel Curves



Smoothing: Engel Curves

- Uniform bands can be constructed for smoothed quantile regression and rearranged curves.
- Previous bootstrap procedure is valid for smoothed rearranged curves even **under population non monotonicity.**
- Smoothed estimates with box kernel and bandwidth $= 0.05$.
- Almost perfect overlap of the bands - indication of population monotonicity.
- Smoothing reduces widths of the bands, but it is not enough to monotonize quantile regression curves.

Smoothing: Engel Curves



Estimation Properties: Monte Carlo

- We consider two versions of the location-scale shift model:

$$y_i = x_i'\alpha + (x_i'\gamma)\epsilon_i, \quad Q_0(u|x_i) = x_i'(\alpha + \gamma F_\epsilon^{-1}(u))$$

(1) **Linear:** $x_i = (1, z_i)$
(2) **Piecewise Linear:** $x_i = (1, z_i, 1\{z_i > Med[z]\} \times z_i)$

- Parameters calibrated to Engel application
- 1,000 Monte Carlo samples of $n = 235$ from a normal with same mean and variance as the residuals $\epsilon_i = (y_i - x_i'\alpha)/(x_i'\gamma)$.
- Regressors fixed to the values of income in Engel data set.
- We estimate a linear model: $Q(u|x_i) = x_i'\beta(u)$ for $x_i = (1, z_i)$

Approximation Properties: Monte Carlo

	(1) Correct Specification			(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**

- Each entry of the table gives a Monte Carlo average of

$$L^p(\tilde{Q}) := \left(\int |Q_0(u|x_0) - \tilde{Q}(u|x_0)|^p du \right)^{1/p}$$

for $\widetilde{Q}(u|x_0) = x_0'\hat{\beta}(u)$, $\widetilde{Q}(u|x_0) = \widehat{F}^{-1}(u|x_0)$, and $x_0 = 452$

Conclusion

Further research directions and extensions:

- Probability curves
- Demand curves
- Growth curves
- Yield curves

Thank you!

alfred.galichon@polytechnique.edu