Non-Crossing Quantile Regression with Shape Constraints

Haruki Kono [1]

MIT

Seminar at ***, DATE

[1]Email: hkono@mit.edu

Quantile Regression

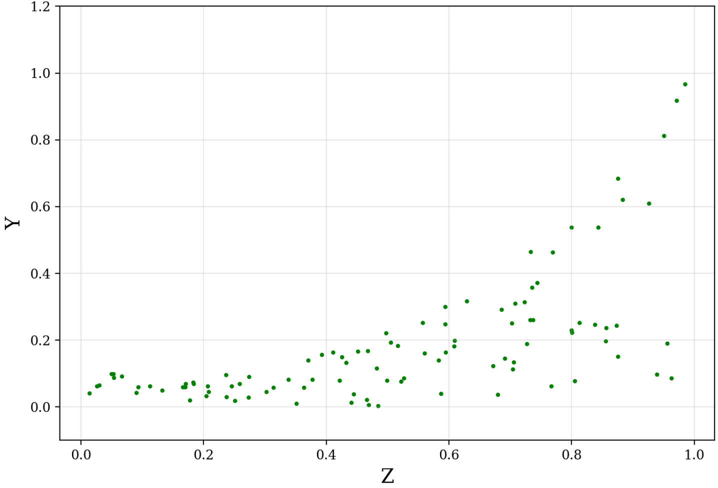
- Economic outcomes are often **heterogeneous**.
  - e.g., household consumption, income inequality
- **Quantile regression** models the conditional quantiles of an outcome Y given covariates $Z = z$:

$$Q_{Y|Z}(u|z) = \begin{array}{l}\text{the outcome level such that} \\ 100u\% \text{ of outcomes among units with } Z = z \\ \text{are} \leq \text{that level.}\end{array}$$

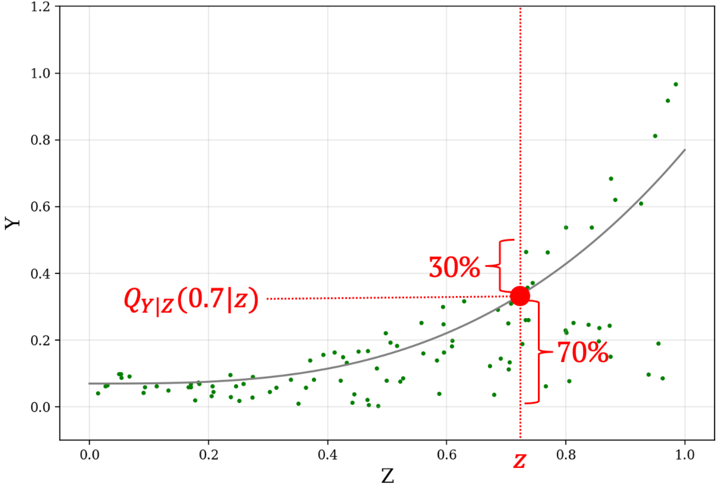
- Running quantile regressions at many quantiles recovers heterogeneous effects of covariates.
  - e.g., how an income change (Z) affects the lower/upper quantiles of consumption (Y).

Illustrative Example

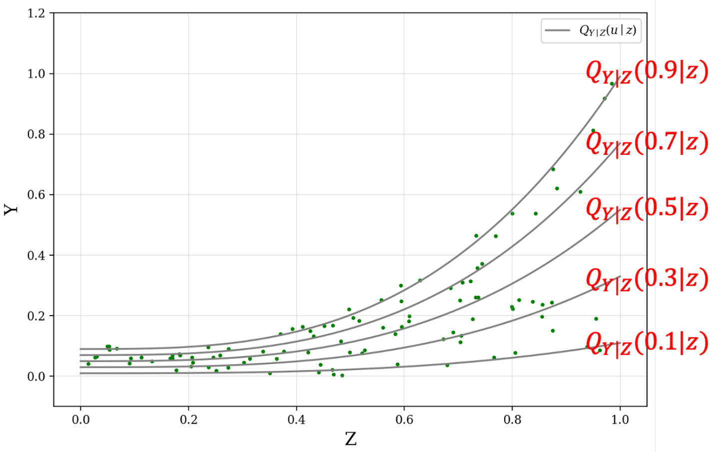
- Generate pairs of outcome Y_i and covariate Z_i from some DGP.



True Conditional Quantile $Q_{Y|Z}(0.7|z)$

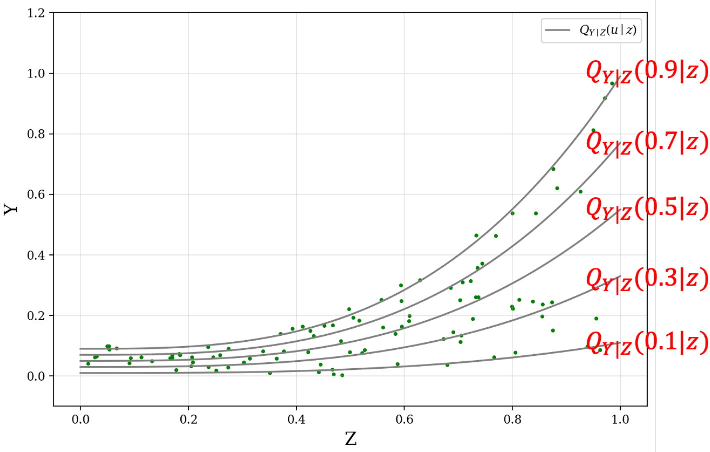


True Conditional Quantile Curves $Q_{Y|Z}(u|z)$



- $Q_{Y|Z}(0.9|z)$ is grows more rapidly with z than $Q_{Y|Z}(0.1|z)$.

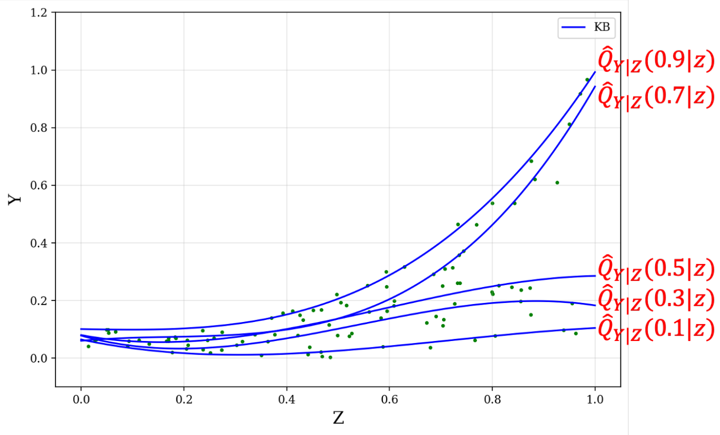
Properties of $Q_{Y|Z}(u|z)$



- Different curves do not cross — e.g., $Q_{Y|Z}(0.7|z) \geq Q_{Y|Z}(0.5|z)$.
- Each curve is increasing — Z affects Y positively.

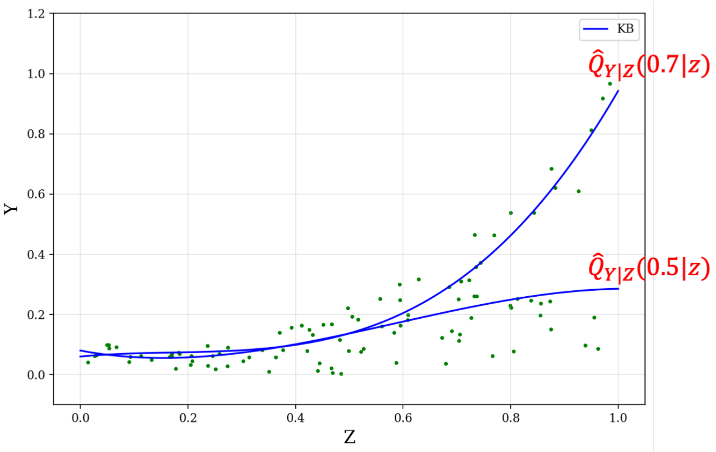
Koenker and Bassett's Estimator

- Koenker and Bassett (1978) study identification/estimation of $Q_{Y|Z}(u|z)$ for each $u \in (0,1)$ from observations (Z_i, Y_i).



- While widely used in theoretical and applied work, **KB's estimator often violates fundamental constraints**.

Quantile Crossing



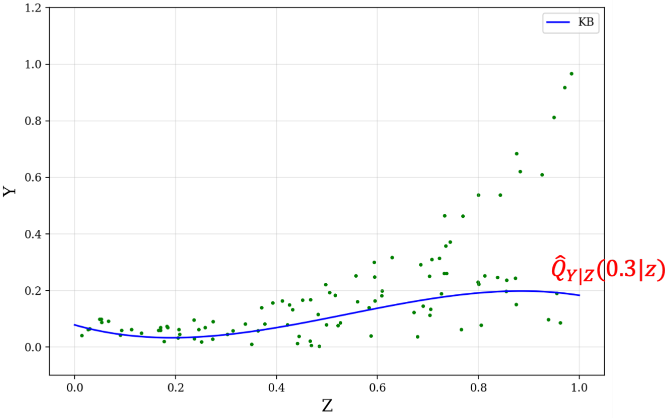
- Estimated quantile curves may **intersect**.

$$\hat{Q}_{Y|Z}(0.5|z) \overset{?}{>} \hat{Q}_{Y|Z}(0.7|z)$$

- Hard to interpret estimates. Negative probability?

Violations of Economic Structure

- Even though $Q_{Y|Z}(u|z)$ is increasing in z in population, estimates are not necessarily so.



- Often **inconsistent** with underlying economic theory

Examples

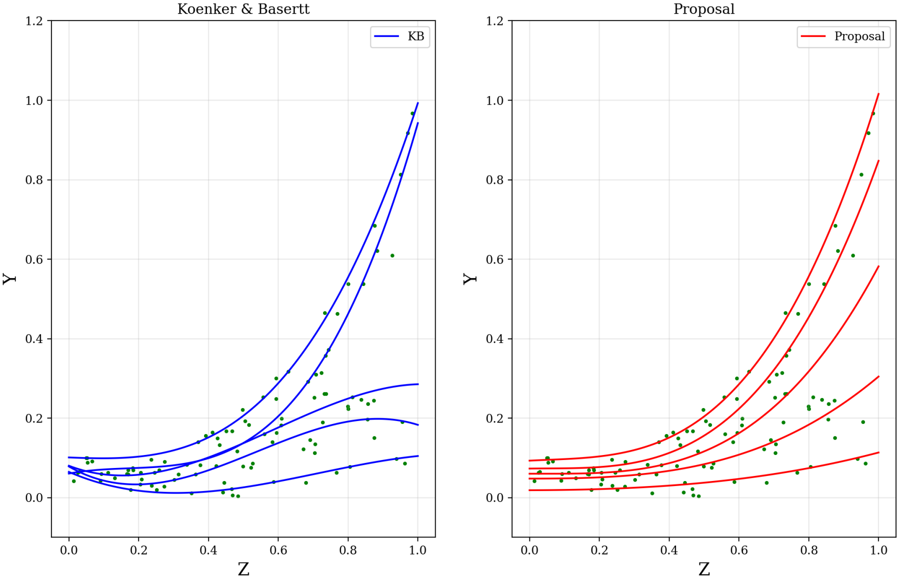
- Production analysis
  - ◇ Y : production
  - ◇ Z : inputs (e.g., labor, capital)
  - ◇ Constraint : monotonicity/concavity in inputs
- First-price auction
  - ◇ Y : bid
  - ◇ Z : auction-specific covariates (e.g., quality of good)
  - ◇ Constraint : monotonicity of bidding strategy
  - ◇ Will be discussed in more detail later.

Goal

Estimate $Q_{Y|Z}(u|z)$ enforcing constraints on

- the map $u \mapsto Q_{Y|Z}(u|z)$
  - ◇ e.g., non-crossing, bidding monotonicity
- the map $z \mapsto Q_{Y|Z}(u|z)$
  - ◇ e.g., shape of production function

KB vs. Proposed Estimator



Model and Identification

- ◇ Introduce a **novel identification** result of quantile regression.

Shape-Constrained Estimation

- ◇ Propose a way to estimate conditional quantile functions **respecting shape constraints**.

Asymptotic Properties

- ◇ Establish **uniform consistency** and **asymptotic normality**.

Empirical Application

- ◇ Apply the estimator to a real-world **auction** dataset.

1 Model and Identification

2 Shape-Constrained Estimation

3 Asymptotic Properties

4 Empirical Application

5 Summary

Setup

- Observe iid (Y_i, Z_i) for $i = 1, \ldots, n$.
  - ◇ Y_i takes values in $\mathcal{Y} \subset \mathbb{R}$.
  - ◇ Z_i takes values in $\mathcal{Z} \subset \mathbb{R}^q$.
- Conditional quantile function $Q_{Y|Z}(u|z)$ is characterized by

$$u = \mathbb{P}\left(Y \leq Q_{Y|Z}(u|z) \mid Z = z\right)$$

for each $(u, z) \in (0, 1) \times \mathcal{Z}$.

Linear Conditional Quantile Functions

- $X_i \coloneqq f(Z_i) = (1, X_{i,-1}) \in \mathcal{X} \subset \mathbb{R}^p$.
  - e.g., $f(z) = (1, z), (1, z, z^2, z^3)$, any transformations.
  - Assume $\mathbb{E}[X_i X_i']$ is full-rank.
- Assume the conditional quantile function of Y given $X = x$ is **linear**:

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

  for some continuous function $\beta(\cdot) : (0, 1) \to \mathbb{R}^p$.
  - **Linear** in x; **nonlinear** in u.
  - Flexible if f is rich.
  - Will discuss misspacification shortly.
- Interested in the **quantile regression coefficient** $\beta(\cdot)$.

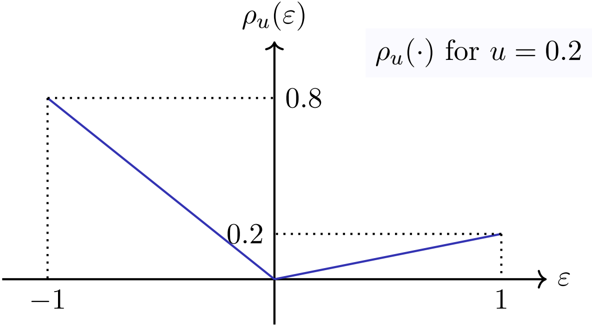
Classical Identification (Koenker and Bassett, 1978)

Lemma (Koenker and Bassett, 1978)

Let $u \in (0,1)$. If $Q_{Y|X}(u|x) = x'\beta(u)$ holds, then $\beta(u)$ is the unique solution to

$$\min_{b \in \mathbb{R}^p} \mathbb{E}\left[\rho_u(Y - X'b)\right]$$

where $\rho_u(\varepsilon) := \varepsilon(u - \mathbf{1}\{\varepsilon < 0\})$ is the check loss function.



Misspecified Conditional Quantile Functions

- Check function minimization is well-defined without the linearity.
- Even when the linearity fails, the rest of the talk is **valid** for the pseudo true parameter

$$\beta(u) := \operatorname*{argmin}_{b \in \mathbb{R}^p} \mathbb{E}\left[\rho_u(Y - X'b)\right]$$

  under a regularity condition that is called quasi-linearity in the paper.

Classical Estimator (Koenker and Bassett, 1978)

- KB's estimator

$$\hat{\beta}_n^{\mathrm{KB}}(u) = \operatorname*{argmin}_{b \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^{n} \rho_u(Y_i - X_i' b).$$

- Each quantile is estimated **separately**.
  - ◇ KB's characterization extracts info. about $\beta(u)$ separately.

$$\begin{tikzcd}
 & \min_b \mathbb{E}[\rho_{0.1}(Y - X'b)] \arrow[r] & \beta(0.1) \\
 & \min_b \mathbb{E}[\rho_{0.2}(Y - X'b)] \arrow[r] & \beta(0.2) \\
 & \vdots & \vdots \\
(X_i, Y_i) \arrow[uuur] \arrow[uur] \arrow[r] \arrow[ddr] \arrow[dddr] & \min_b \mathbb{E}[\rho_{0.5}(Y - X'b)] \arrow[r] & \beta(0.5) \\
 & \vdots & \vdots \\
 & \min_b \mathbb{E}[\rho_{0.8}(Y - X'b)] \arrow[r] & \beta(0.8) \\
 & \min_b \mathbb{E}[\rho_{0.9}(Y - X'b)] \arrow[r] & \beta(0.9)
\end{tikzcd}$$

- Hard to impose constraints **across** quantiles (e.g., non-crossing).

Key Variational Problem for New Characterization

$$(X_i, Y_i) \longrightarrow \boxed{\text{variational problem}} \longrightarrow \beta(\cdot)$$

- Choose
  - $\psi(\cdot) \in \{f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \mid f : \text{cont.}\} =: C(\mathcal{X} \times \mathcal{Y})$
  - $\sigma(\cdot) = (\sigma_1(\cdot), \ldots, \sigma_p(\cdot))$ where for each $k = 1, \ldots, p$,

$$\sigma_k(\cdot) \in \left\{ f : [0,1] \to \mathbb{R} \mid f : \text{cont.}, \int_0^1 f(u)du = 0 \right\} =: \underbrace{C[0,1]}_{\text{cont.}} \cap \underbrace{L_0^1[0,1]}_{\text{zero-mean}}$$

that minimizes

$$\int_{\mathcal{X} \times \mathcal{Y}} \psi(x,y) \underbrace{dF_{XY}(x,y)}_{\text{dist. of } (X,Y)}$$

subject to

$$uy \leq \psi(x,y) + x'\sigma(u) \text{ for } (u,x,y) \in [0,1] \times \mathcal{X} \times \mathcal{Y}$$

Existence of Solution

Lemma (Carlier et al., 2016)

The infinite-dimensional linear program (LP)

$$\inf_{\substack{\boldsymbol{\psi}(\cdot)\in C(\mathcal{X}\times\mathcal{Y})\\ \boldsymbol{\sigma}(\cdot)\in(C[0,1]\cap L_0^1[0,1])^p}} \int \boldsymbol{\psi}\, dF_{XY} .$$

$$\text{s.t.} \quad uy \leq \boldsymbol{\psi}(x,y) + x'\boldsymbol{\sigma}(u) \text{ for } (u,x,y) \in [0,1]\times\mathcal{X}\times\mathcal{Y}$$

admits a solution with

$$\boldsymbol{\sigma}_\beta(u) := \int_0^u \beta(v)dv - \int_0^1 \left(\int_0^{\tilde{u}} \beta(v)dv\right) d\tilde{u}.$$

- Why useful? — The solution $\sigma_\beta(\cdot)$ tells us something about $\beta(\cdot)$.
- Focus on σ; optimal ψ is recovered from optimal σ.

Properties of The Solution

- The solution

$$\sigma_\beta(u) = \int_0^u \beta(v)dv - \int_0^1 \left(\int_0^{\tilde{u}} \beta(v)dv \right) d\tilde{u}$$

  is an **anti-derivative** of $\beta(\cdot)$, i.e., $\dfrac{\partial \sigma_\beta}{\partial u}(\cdot) = \beta(\cdot)$.

- The map $u \mapsto x'\sigma_\beta(u)$ is **convex** for any $x \in \mathcal{X}$.
  - ◇ $\dfrac{\partial(x'\sigma_\beta(\cdot))}{\partial u} = x'\beta(\cdot) = Q_{Y|X}(\cdot|x)$ is non-decreasing.

Uniqueness of Solution

Theorem

The map $\sigma(\cdot) = \sigma_\beta(\cdot)$ is the **unique** solution to the infinite-dimensional LP such that the map $u \mapsto x'\sigma(u)$ is convex for all $x \in \mathcal{X}$.

- While linking quantile regression to the infinite-dimensional LP is not new (Carlier et al. (2016), Carlier et al. (2017)), the uniqueness of the solution is **novel**.
- Uniqueness is essential to establish the identification of $\beta(\cdot)$.
    - ◇ Leads to a new estimation framework.

New Identification of $\beta(\cdot)$

- The distribution F_{XY} of observables (X, Y) defines

$$\begin{aligned} & \inf_{\psi,\sigma} \int \psi dF_{XY} \\ \text{s.t.} \quad & uy \leq \psi(x,y) + x'\sigma(u) \text{ for } (u,x,y) \in [0,1] \times \mathcal{X} \times \mathcal{Y} \end{aligned}.$$

- The unique solution with the convexity qualification is $\sigma_\beta(\cdot)$.
- The quantile regression coefficient $\beta(\cdot)$ is recovered via

$$\beta(\cdot) = \frac{\partial \sigma_\beta}{\partial u}(\cdot).$$

$$F_{XY} \xrightarrow{\text{def.}} \inf_{\psi,\sigma} \int \psi dF_{XY} \xrightarrow{\exists! \text{ sol.}} \sigma_\beta(\cdot) \xrightarrow{\partial/\partial u} \beta(\cdot)$$

1 Model and Identification

2 Shape-Constrained Estimation

3 Asymptotic Properties

4 Empirical Application

5 Summary

Basic Idea for Estimation

- Start with a simple plug-in estimator.
- Replacing F_{XY} with $\hat{F}_{XY} = \frac{1}{n}\sum_i \delta_{(X_i,Y_i)}$, consider

$$\inf_{\substack{(\psi_i)_{i=1}^n \in \mathbb{R}^n \\ \sigma(\cdot)\in(C[0,1]\cap L_0^1[0,1])^p}} \frac{1}{n}\sum_{i=1}^n \psi_i \quad \text{s.t.} \quad \forall (u,i),\ uY_i \leq \psi_i + X_i'\sigma(u)$$

- For a solution $\hat{\sigma}(\cdot)$, estimate $\beta(\cdot)$ with

$$\hat{\beta}(\cdot) \stackrel{?}{=} \frac{\partial \hat{\sigma}}{\partial u}(\cdot).$$

Two Challenges

$$\inf_{\substack{(\psi_i)_{i=1}^n \in \mathbb{R}^n \\ \sigma(\cdot)\in (C[0,1]\cap L_0^1[0,1])^p}} \frac{1}{n}\sum_{i=1}^n \psi_i \quad \text{s.t.} \quad \forall (u,i),\ uY_i \le \psi_i + X_i'\sigma(u)$$

1. No constraint has been imposed on $\hat{\beta}(\cdot)$ yet.

2. Optimization in a function space is hard.

Two Challenges

$$\inf_{\substack{(\psi_i)_{i=1}^n \in \mathbb{R}^n \\ \boldsymbol{\sigma(\cdot)\in\Sigma_{J_n}}}} \frac{1}{n}\sum_{i=1}^{n}\psi_i \quad \text{s.t.} \quad \begin{cases} \forall (u,i),\ uY_i \le \psi_i + X_i'\sigma(u) \\ \textbf{Linear ineq. cnsts. on } \sigma(\cdot) \end{cases}$$

1. No constraint has been imposed on $\hat{\beta}(\cdot)$ yet.
   - Add corresponding constraints to the LP.
2. Optimization in a function space is hard.
   - Introduce a finite-dim. approx. Σ_{J_n} of $(C[0,1] \cap L_0^1[0,1])^p$.

① Adding Shape Constraints

- Many constraints can be written as **linear** inequalities of $\beta(\cdot)$.
  - ◇ Will see examples in the next slide.
- They imply those of $\sigma(\cdot)$ since $\beta(\cdot) = \dfrac{\partial \sigma}{\partial u}(\cdot)$.
- Add the constraints on $\sigma(\cdot)$ to the LP.

$$\inf_{\substack{(\psi_i)_{i=1}^n \in \mathbb{R}^n \\ \sigma(\cdot) \in (C[0,1] \cap L_0^1[0,1])^p}} \frac{1}{n} \sum_{i=1}^n \psi_i \quad \text{s.t.} \quad \begin{cases} \forall (u,i),\ uY_i \leq \psi_i + X_i' \sigma(u) \\ \textbf{Linear ineq. cnsts. on } \sigma(\cdot) \end{cases} .$$

Example (Non-Crossing)

$$
\begin{aligned}
& \text{Non-crossing} \\
\iff & \text{The map } u \mapsto Q_{Y|X}(u|x) \text{ is non-decreasing for each } x \\
\iff & \forall (u,x),\ \frac{\partial Q_{Y|X}(u|x)}{\partial u} \geq 0 \\
\iff & \forall (u,x),\ \frac{\partial}{\partial u}\big(x'\beta(u)\big) \geq 0 \quad \big[\because Q_{Y|X}(u|x) = x'\beta(u)\big] \\
\iff & \underbrace{\forall (u,x),\ x'\left(\frac{\partial^2 \sigma(u)}{\partial u^2}\right) \geq 0}_{\text{Linear ineq. in } \sigma(\cdot)} \quad \left[\because \beta(\cdot) = \frac{\partial \sigma}{\partial u}(\cdot)\right]
\end{aligned}
$$

Example (Covariate Monotonicity)

- Recall $X = f(Z)$.
- Suppose that Z is a scalar for simplicity.

$$\text{The map } z \mapsto Q_{Y|Z}(u|z) \text{ is non-decreasing for each } u$$

$$\iff \forall(u,z),\ \frac{\partial Q_{Y|Z}(u|z)}{\partial z} \geq 0$$

$$\iff \forall(u,z),\ \frac{\partial}{\partial z}\left(f(z)'\beta(u)\right) \geq 0 \quad \left[\because Q_{Y|Z}(u|z) = f(z)'\beta(u)\right]$$

$$\iff \underbrace{\forall(u,z),\ \left(\frac{\partial f(z)}{\partial z}\right)'\left(\frac{\partial \sigma(u)}{\partial u}\right) \geq 0}_{\text{Linear ineq. in } \sigma(\cdot)} \quad \left[\because \beta(\cdot) = \frac{\partial \sigma}{\partial u}(\cdot)\right]$$

Example (Bidding Monotonicity)

- The monotonicity of the equilibrium bidding strategy can be written as a linear inequality constraint on the quantile function of bids.
- Will be discussed in detail later.

① Adding Shape Constraints

- Augmented LP

$$\inf_{\substack{(\psi_i)_{i=1}^n \in \mathbb{R}^n \\ \sigma \in (C[0,1] \cap L_0^1[0,1])^p}} \frac{1}{n} \sum_{i=1}^{n} \psi_i \quad \text{s.t.} \quad \begin{cases} \forall (u,i),\ uY_i \le \psi_i + X_i' \sigma(u) \\ \textbf{Linear ineq. cnsts. on } \boldsymbol{\sigma} \end{cases}$$

- Any feasible $\sigma(\cdot)$ produces $\beta(\cdot) = \dfrac{\partial \sigma}{\partial u}(\cdot)$ satisfying the imposed constraints.

② Finite-Dim. Approx. of $\sigma(\cdot) \in (C[0,1] \cap L_0^1[0,1])^p$

- Optimizing over $\sigma(\cdot) \in (\underbrace{C[0,1]}_{\text{cont.}} \cap \underbrace{L_0^1[0,1]}_{\text{zero-mean}})^p$ is **infeasible**.
- For an integer J, approximate the space with

$$\begin{aligned}\Sigma_J &:= (\text{space of zero-mean polynomials of order} \le J)^p \\ &\subset (C[0,1] \cap L_0^1[0,1])^p\end{aligned}$$

  - ◇ Σ_J is J-dimensional and increasing in J.
  - ◇ Any function in $(C[0,1] \cap L_0^1[0,1])^p$ can be approximated by some function in Σ_J **arbitrarily well** for large J.
  - ◇ Functional bases other than polynomials can also be used.
- Expected that the LP over Σ_J yields an **approximate solution**.

Feasible Constrained Estimator

- For some choice $J = J_n$, consider

$$\inf_{\substack{(\psi_i)_{i=1}^n \in \mathbb{R}^n \\ \boldsymbol{\sigma(\cdot)\in\Sigma_{J_n}}}} \frac{1}{n}\sum_{i=1}^{n} \psi_i \quad \text{s.t.} \quad \begin{cases} \forall (u,i),\ uY_i \leq \psi_i + X_i'\sigma(u) \\ \textbf{Linear ineq. cnsts. on } \sigma(\cdot) \end{cases}$$

- Find a solution $\hat{\sigma}_n(\cdot) \in \Sigma_{J_n}$ and estimate $\beta(\cdot)$ with

$$\hat{\beta}_n(\cdot) := \frac{\partial \hat{\sigma}_n}{\partial u}(\cdot)$$

- $\hat{\beta}_n(\cdot)$ satisfies the shape constraints by construction.
  - e.g., non-crossing, monotonicity in z, bidding monotonicity.
- By choosing $J_n \to \infty$, Σ_{J_n} becomes as flexible as $(C[0,1] \cap L_0^1[0,1])^p$.

1 Model and Identification

2 Shape-Constrained Estimation

3 Asymptotic Properties

4 Empirical Application

5 Summary

Uniform Consistency

Theorem

Assume

- the imposed shape constraints are correct in population;
- **the non-crossing condition is imposed**;
- $J_n \to \infty$;
- other regularity conditions.

Then, for any compact set $K \subset (0,1)$,

$$\lim_{n\to\infty} \sup_{u\in K} \left\| \hat{\beta}_n(u) - \beta(u) \right\| = 0 \quad \text{a.s.}$$

holds.

Role of Non-Crossing Condition

- When a function is estimated using a ${J_n}^{\text{th}}$-order basis expansion, consistency requires controlling the growth rate of J_n.
- In contrast, consistency of $\hat{\beta}_n(\cdot)$ imposes **no** rate condition on J_n.
- **Non-crossing**, or monotonicity of $u \mapsto x'\hat{\beta}_n(u)$, **automatically regularizes**.

Asymptotic Normality

- Distributional theory for $\hat{\beta}_n(\cdot)$ is hard in nonparametric setting.
- Assume the polynomial approximation is **exact**.

Assumption

There exists $\sigma(\cdot) \in \Sigma_{J_*}$ for some $J_* > 0$ such that

$$\beta(\cdot) = \frac{\partial \sigma}{\partial u}(\cdot),$$

i.e., $\beta(\cdot)$ is a $(J_* - 1)^{\text{th}}$-order polynomial.

Asymptotic Normality

Theorem

Assume

- the polynomial approximation is **exact**;
- shape constraints are correct in population;
- the non-crossing condition is imposed;
- $J_n = \bar{J} \geq J_*$;
- other regularity conditions.

Then,

$$\sqrt{n}\left(\hat{\beta}_n(u) - \beta(u)\right) \Rightarrow N\left(0, v(u)\right)$$

holds for each $u \in (0, 1)$.

- Uniform version holds.
- Asymptotic variance is consistently estimable.

1 Model and Identification

2 Shape-Constrained Estimation

3 Asymptotic Properties

4 Empirical Application

5 Summary

U.S. Timber Auction

- First-price auctions organized by US Forest Service in 1979.
- Focus on auctions with two bidders.
- Contains all bids, timber volumes, and appraisal values.

First-Price Sealed-Bid Auctions

- **Independent-private-value paradigm** with two bidders.
  - ◇ Auction $\ell = 1, \ldots, L$ has characteristic X_ℓ.
  - ◇ Bidders $i = 1, 2$ come to auction ℓ.
  - ◇ Bidder i draws private value $V_{\ell i}$ from $F_{V|X}(\cdot|X_\ell)$.
  - ◇ Bidder i submits bid $B_{\ell i} = s(V_{\ell i}|X_\ell)$.
  - ◇ Bidder with higher bid wins and pays his own bid.
- Observe $(X_\ell, B_{\ell i})$ for each (ℓ, i).
  - ◇ $X_\ell = \begin{pmatrix} 1 & \log(\text{volume}_\ell) & \text{appraisal}_\ell \end{pmatrix}'$
- Interested in
  - ◇ **bidding strategy** $s(\cdot|x)$
  - ◇ **private value distribution** $F_{V|X}(\cdot|x)$

Equilibrium

- Assume bidders are **symmetric**.
- A Bayes-Nash eq. bidding strategy **uniquely exists**, is **increasing**, and is given by $s(\cdot|x) = \xi^{-1}(\cdot|x)$ where

$$\xi(b|x) := b + \frac{F_{B|X}(b|x)}{f_{B|X}(b|x)},$$

  where $F_{B|X}$ and $f_{B|X}$ are the conditional distribution and density of bids, respectively.

- The private value quantile function is identified via

$$\underbrace{Q_{V|X}(u|x)}_{\text{quantile of valuations}} = \underbrace{Q_{B|X}(u|x)}_{\text{quantile of bids}} + u\frac{\partial Q_{B|X}}{\partial u}(u|x).$$

Bid Quantile Regression

- Estimation of ξ and $Q_{V|X}$ requires the conditional bid dist.
- Gimenes and Guerre (2022) model it with a linear quantile regression model
$$Q_{B|X}(u|x) = x'\beta(u).$$
  - ◇ Holds if bidders have linear-additive random valuations:
  $$V_{\ell i} = X_\ell' \gamma + \varepsilon_{\ell i}$$
  where γ is a constant coefficient and $\varepsilon_{\ell i}$ is a preference shock.
  - ◇ More generally, it allows for random coefficient valuations.

Unconstrained Estimator (Gimenes and Guerre, 2022)

- Gimenes and Guerre (2022) propose estimators of $\beta(\cdot)$ and its derivative $\dfrac{\partial \beta}{\partial u}(\cdot)$.
- Compute estimators $\hat{Q}^{\mathrm{GG}}_{B|X}$, $\hat{Q}^{\mathrm{GG}}_{V|X}$, and $\hat{\xi}^{\mathrm{GG}}$.
- Gimenes and Guerre (2022)'s estimators are **unconstrained**.
  - $\hat{Q}^{\mathrm{GG}}_{B|X}(\cdot|x)$ and $\hat{Q}^{\mathrm{GG}}_{V|X}(\cdot|x)$ can be non-monotone.
  - $\hat{\xi}^{\mathrm{GG}}(\cdot|x)$ can be non-monotone.

Shape Constraints Imposed by Theory

- **Bid non-crossing**: the map $u \mapsto Q_{B|X}(u|x)$ is non-decreasing.
- **Monotonicity of equilibrium bidding strategy**:

$$\forall (u, x),\ \frac{\partial}{\partial u}\left(x'\left(\beta(u) + u\frac{\partial \beta}{\partial u}(u)\right)\right) \geq 0 \ : \ \text{linear in } \beta(\cdot)$$

- These imply the **valuation non-crossing**.

Constrained Estimator

- Consider

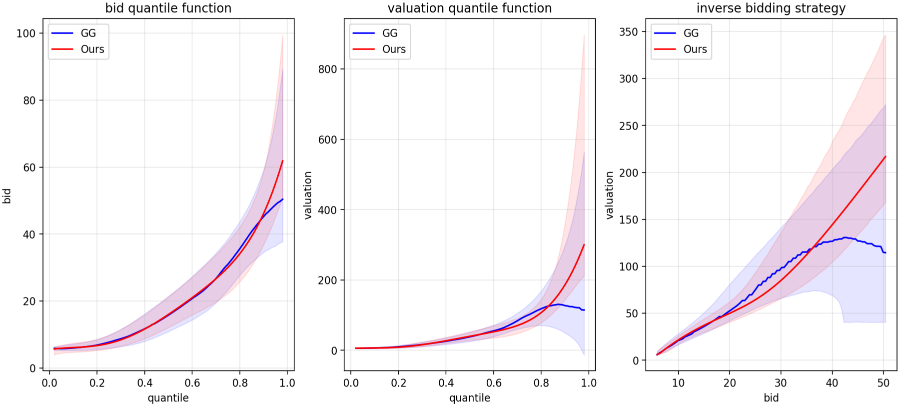
$$\inf_{\substack{(\psi_{\ell i})_{\ell.i}\in\mathbb{R}^{LI}\\ \sigma(\cdot)\in\Sigma_{J_n}}} \frac{1}{LI}\sum_{\ell=1}^{L}\sum_{i=1}^{I}\psi_{\ell i} \quad \text{s.t.} \quad \begin{cases} \forall(u,\ell,i),\ uB_{\ell i}\leq \psi_{\ell i}+X_\ell'\sigma(u) \\ \textbf{Bid non-crossing} \\ \textbf{Bidding monotonicity} \end{cases}$$

- For a solution $\hat{\sigma}_n$, define an estimator

$$\hat{\beta}_n(\cdot) := \frac{\partial\hat{\sigma}_n}{\partial u}(\cdot).$$

- $\hat{\beta}_n(\cdot)$ satisfies bid/valuation non-crossing and bid monotonicity.

Estimates at Data Point $x = (1, 4.32, 5.67)'$



1. Model and Identification
2. Shape-Constrained Estimation
3. Asymptotic Properties
4. Empirical Application
5. Summary

Summary

- Introduced a novel identification result of quantile regression.
- Proposed an estimator based on the identification and imposed various shape restrictions such as non-crossing and other economic constraints.
- Showed the asymptotic properties.
- Obtained estimates respecting desired constraints in a real dataset.

A statistical framework living inside economic models.

References I

- Angrist, Joshua, Victor Chernozhukov, and Iván Fernández-Val (2006). "Quantile regression under misspecification, with an application to the US wage structure". In: *Econometrica* 74(2), pp. 539–563.
- Bondell, Howard D, Brian J Reich, and Huixia Wang (2010). "Noncrossing quantile regression curve estimation". In: *Biometrika* 97(4), pp. 825–838.
- Carlier, Guillaume, Victor Chernozhukov, and Alfred Galichon (2016). "Vector Quantile Regression: An Optimal Transport Approach". In: *Annals of Statistics* 44(3), pp. 1165–1192.
- Carlier, Guillaume, Victor Chernozhukov, and Alfred Galichon (2017). "Vector quantile regression beyond the specified case". In: *Journal of Multivariate Analysis* 161, pp. 96–102.
- Chernozhukov, Victor, Iván Fernández-Val, and Alfred Galichon (2010). "Quantile and probability curves without crossing". In: *Econometrica* 78(3), pp. 1093–1125.

References II

- Gimenes, Nathalie and Emmanuel Guerre (2022). "Quantile regression methods for first-price auctions". In: *Journal of Econometrics* 226(2), pp. 224–247.
- Koenker, Roger and Gilbert Bassett (1978). "Regression quantiles". In: *Econometrica: journal of the Econometric Society*, pp. 33–50.
- López, Marco and Georg Still (2007). "Semi-infinite programming". In: *European journal of operational research* 180(2), pp. 491–518.

APPENDIX

Simulation DGP

- $n = 100$.
- $(Z_i, U_i) \sim U[0,1] \otimes U[0,1]$ iid for $i = 1, \ldots n$.
- $Y_i = U_i(Z_i^3 + 0.1)$ for $i = 1, \ldots n$.
- Observe (Z_i, Y_i) for $i = 1, \ldots n$.

Back

Linear Approximation

- For now, suppose that Z_i is a **scalar**.
- $Q_{Y|Z}(u|z)$ admits an expansion

$$
\begin{aligned}
Q_{Y|Z}(u|z) &= a_{00} + a_{01}u + a_{10}z + a_{02}u^2 + a_{11}uz + a_{20}z^2 + \cdots \\
&= \underbrace{(a_{00} + a_{01}u + a_{02}u^2 \cdots)}_{=:\beta_0(u)} \\
&\qquad + \underbrace{(a_{10} + a_{11}u + \cdots)}_{=:\beta_1(u)} z + \underbrace{(a_{20} + \cdots)}_{=:\beta_2(u)} z^2 + \cdots \\
&\approx \underbrace{\begin{pmatrix} 1 & z & z^2 \end{pmatrix}}_{=:x} \underbrace{\begin{pmatrix} \beta_0(u) \\ \beta_1(u) \\ \beta_2(u) \end{pmatrix}}_{=:\beta(u)} \\
&= x'\beta(u).
\end{aligned}
$$

Back

- Higher-order polynomials produce more flexible quantile functions.

Quasi-Linearity

- For $u \in (0,1)$, the uth **quantile regression coefficient** of F_{YX} is defined as

$$\beta(u) \in \operatorname*{argmin}_{\beta \in \mathbb{R}^p} \mathbb{E}\left[\rho_u(Y - X'\beta)\right].$$

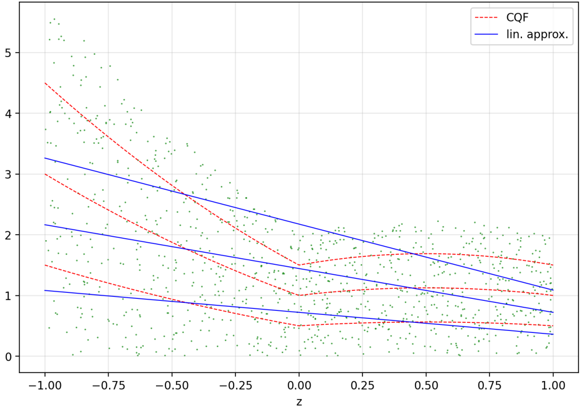
- The map $u \mapsto x'\beta(u)$ is used as a **surrogate** of the CQF without the linearity.
  - ◇ Angrist et al. (2006) showed that it is the best linear approx. of the CQF under a weighted mean-square error
  - ◇ But the map $u \mapsto x'\beta(u)$ is not necessarily increasing.

- The joint distribution F_{YX} satisfies the **quasi-linearity** if $u \mapsto \beta(u)$ is continuous and $u \mapsto x'\beta(u)$ is strictly increasing for each $x \in \mathcal{X}$.
  - ◇ A minimal assumption to consider $u \mapsto x'\beta(u)$ as a surrogate of CQF.
  - ◇ QL is satisfied if $Q_{Y|X}(u|x)$ is linear in x.

Back

Quasi-Linearity

- Allows for deviations from linear CQFs.

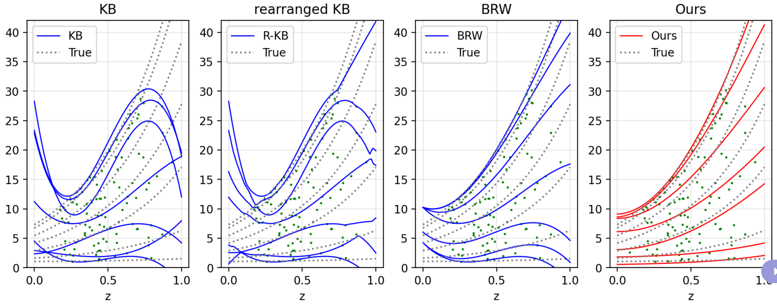


- QL is expected to hold if f is sufficiently rich.

Back

Existing Non-Crossing Estimators

- Rearrangement (Chernozhukov et al. (2010))
  - cannot accommodate other constraints;
  - tend to produce kinked estimates;
  - weaken interpretability due to linearity violation.
- Composite quantile regression (Bondell et al. (2010))
  - imposes monotonicity only on finitely many quantiles;
  - can be computationally demanding when many quantiles are accounted for;
  - cannot accommodate constraints on derivatives.



▸▸ Back

Optimal $\psi(\cdot)$

- Given $\sigma(\cdot)$, the inequality constraint

$$uy \leq \psi(x, y) + x'\sigma(u) \text{ for } (u, x, y) \in [0, 1] \times \mathcal{X} \times \mathcal{Y}$$

gives the optimal $\psi(\cdot)$ as follows

$$\psi(x, y) = \sup_{u \in [0,1]} \left(uy - x'\sigma(u) \right).$$

Back

Naive Empirical Problem

- Population problem:

$$\inf_{\substack{\psi \in C(\mathcal{X}\times\mathcal{Y}) \\ \sigma \in (C[0,1]\cap L_0^1[0,1])^p}} \int \psi \, dF_{XY} \quad \text{s.t.} \quad \forall (u,x,y),\ uy \le \psi(x,y) + x'\sigma(u)$$

- Empirical problem:

$$\inf_{\substack{\psi \in C(\mathcal{X}\times\mathcal{Y}) \\ \sigma \in (C[0,1]\cap L_0^1[0,1])^p}} \frac{1}{n}\sum_{i=1}^n \psi(X_i, Y_i) \quad \text{s.t.} \quad \forall (u,x,y),\ uy \le \psi(x,y) + x'\sigma(u)$$

- Equivalent empirical problem:

$$\inf_{\substack{(\psi_i)_{i=1}^n \in \mathbb{R}^n \\ \sigma \in (C[0,1]\cap L_0^1[0,1])^p}} \frac{1}{n}\sum_{i=1}^n \psi_i \quad \text{s.t.} \quad \forall (u,i),\ uY_i \le \psi_i + X_i'\sigma(u)$$

Back

Semi-Infinite Program

Write the inequality constraints as

$$g(\psi, \sigma; i, u, x) \geq 0 \quad \text{for} \quad (i, u, x) \in \{1, \ldots, n\} \times [0, 1] \times \mathcal{X} =: T.$$

We consider a grid $T_k \coloneqq \{1, \ldots, n\} \times \left\{0, \frac{1}{2^k}, \ldots, \frac{2^k-1}{2^k}, 1\right\} \times \mathcal{X}_k$, where $\mathcal{X}_k \nearrow \mathcal{X}$. Given this grid and a small tolerance level $\varepsilon > 0$, we run the following algorithm (López and Still (2007)).

- Let $k = 1$.
- Solve the LP subject to the constraints over T_k instead of T. Let $(\psi^{(k)}, \sigma^{(k)})$ be a solution to the problem.
- Stop the algorithm if the inequalities for $(\psi^{(k)}, \sigma^{(k)})$ approximately hold on T, i.e.,

  $$g(\psi^{(k)}, \sigma^{(k)}; i, u, x) > -\varepsilon \quad \text{for} \quad (i, u, x) \in \{1, \ldots, n\} \times [0, 1] \times \mathcal{X}.$$

  Otherwise, repeat Step 2 with $k + 1$.

Back

Uniform Limit Distribution

Under the same assumption, $\sqrt{n}\left(\hat{\beta}_n(\cdot) - \beta(\cdot)\right)$ weakly converges to a tight centered Gaussian process $\mathbb{G}(\cdot)$ in $\ell^\infty([0,1],\mathbb{R}^p)$, of which covariance function is given by

$$\mathbb{E}[\mathbb{G}(u_1)\mathbb{G}(u_2)'] = \left(\frac{\partial m^{\bar{J}}}{\partial u}(u_1)' \otimes I_p\right)(V^{-1}WV^{-1})\left(I_p \otimes \frac{\partial m^{\bar{J}}}{\partial u}(u_2)\right)$$

where

$$W := \mathbb{E}\left[(X \otimes m^{\bar{J}}(U))(X \otimes m^{\bar{J}}(U))'\right],$$

$$V := \mathbb{E}\left[\frac{1}{X'\frac{\partial \beta}{\partial u}(U)}\left(X \otimes \frac{\partial m^{\bar{J}}}{\partial u}(U)\right)\left(X \otimes \frac{\partial m^{\bar{J}}}{\partial u}(U)\right)'\right],$$

and

$$m^{\bar{J}}(u) := \begin{pmatrix} u - \frac{1}{2} & \frac{1}{2}\left(u^2 - \frac{1}{3}\right) & \dots & \frac{1}{\bar{J}}\left(u^{\bar{J}} - \frac{1}{\bar{J}+1}\right)\end{pmatrix}'.$$

Back

Variance estimation

Let $\hat{U}_i := (X_i'\hat{\beta}_n(\cdot))^{-1}(Y_i)$ for $i = 1, \ldots, n$, and define

$$\hat{W}_n := \frac{1}{n}\sum_{i=1}^{n}\left(X_i \otimes m^{\bar{J}}(\hat{U}_i)\right)\left(X_i \otimes m^{\bar{J}}(\hat{U}_i)\right)'$$

and

$$\hat{V}_n := \frac{1}{n}\sum_{i=1}^{n}\frac{1}{X_i'\frac{\partial\hat{\beta}_n}{\partial u}(\hat{U}_i)}\left(X_i \otimes \frac{\partial m^{\bar{J}}}{\partial u}(\hat{U}_i)\right)\left(X_i \otimes \frac{\partial m^{\bar{J}}}{\partial u}(\hat{U}_i)\right)'.$$

Under the same setup,

$$\hat{W}_n \to W \quad \text{and} \quad \hat{V}_n \to V$$

hold almost surely.

Back

Linear Quantile Specification of Bids

- The linear quantile specification of bids assumes the representation

$$V_{\ell i} = X_\ell' \gamma(U_{\ell i}), \quad U_{\ell i} \sim U[0,1], \quad X_\ell \perp\!\!\!\perp U_{\ell i}$$

  where $u \mapsto x'\gamma(u)$ is non-decreasing for each x.

- The linear-additive random representation

$$V_{\ell i} = X_\ell' \gamma + \varepsilon_{\ell i}$$

  is in this class

$$\gamma(U_{\ell i}) = \left(\gamma_1 + \underbrace{F_\varepsilon(U_{\ell i})}_{=\varepsilon_{\ell i}}, \gamma_2, \ldots, \gamma_p\right)'$$

Back

DGP 1 (linear CQF, from Bondell et al. (2010))

(DGP1) The data is independently drawn from

$$Y_i = X_i'(\alpha + \gamma \Phi^{-1}(U_i)), \quad X_i = (1, Z_i'), \quad (Z_i, U_i) \sim U[0,1]^4 \otimes U[0,1]$$

for $i = 1, \ldots, n$, where Φ^{-1} is the quantile function of the standard Gaussian distribution,

$$\alpha = (1,1,1,1,1)' \quad \text{and} \quad \gamma = (1, 0.1, 0.1, 0.1, 0.1)'.$$

- The QR coefficient is

$$\beta(u) = \alpha + \gamma \Phi^{-1}(u).$$

- Set $n = 100$ following Bondell et al. (2010).

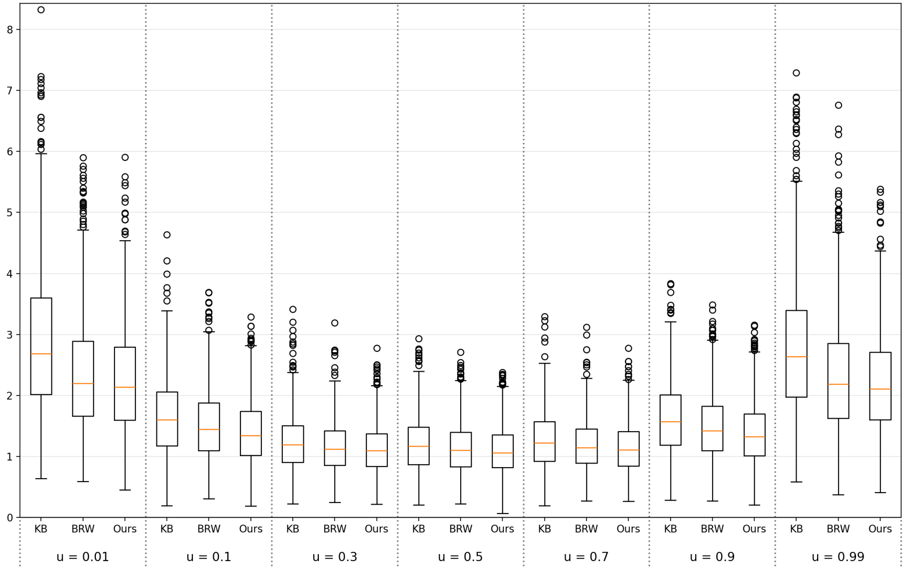
DGP 1

- Let $(u_1, u_2, u_3, u_4, u_5, u_6, u_7) = (0.01, 0.1, 0.3, 0.5, 0.7, 0.9, 0.99)$.
- Compare our estimator estimated with the non-crossing condition with Koenker and Bassett (1978) and Bondell et al. (2010).
- These estimators are evaluated by the loss function

$$\sqrt{\frac{1}{7}\sum_{k=1}^{7}\left(\tilde{\beta}(u_k) - \beta(u_k)\right)^2}.$$

- Generate data and compute the loss for 500 times.

DGP 1



0
1
2
3
4
5
6
7
8
KB
BRW
Ours
KB
BRW
Ours
KB
BRW
Ours
KB
BRW
Ours
KB
BRW
Ours
KB
BRW
Ours
KB
BRW
Ours
u = 0.01
u = 0.1
u = 0.3
u = 0.5
u = 0.7
u = 0.9
u = 0.99

DGP 2 (nonlinear CQF but quasi-linear)

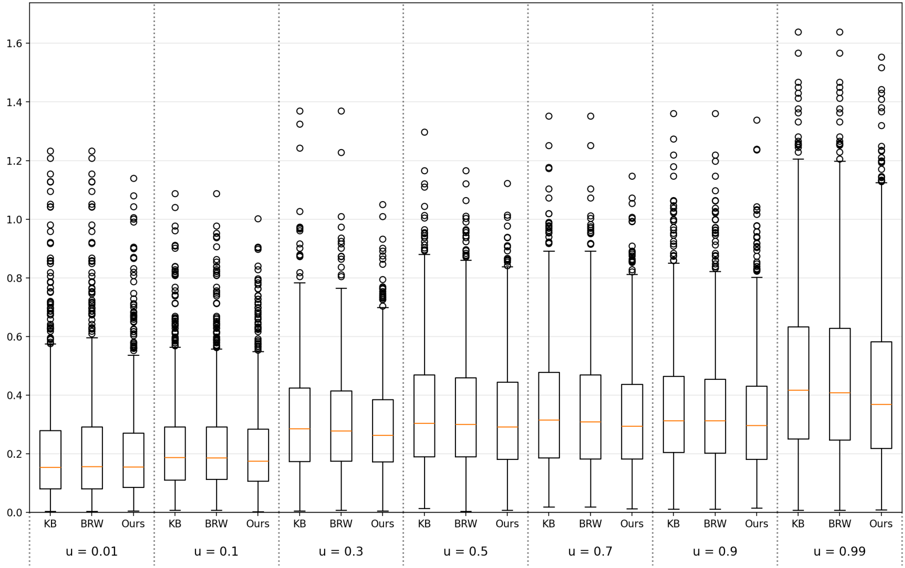
(DGP2) The data is independently drawn from

$$Y_i = (2 - Z_i)(|Z_i| + 1)U_i, \quad X_i = (1, Z_i)', \quad (Z_i, U_i) \sim U[-1,1] \otimes U[0,1]$$

for $i = 1, \ldots, n$. figure

- Recall that the joint distribution F_{YX} is QL, but the CQF $Q_{Y|X}(u|x)$ is nonlinear in x.
- Set $n = 50$.

DGP 2



1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
KB BRW Ours KB BRW Ours KB BRW Ours KB BRW Ours KB BRW Ours KB BRW Ours KB BRW Ours
u = 0.01
u = 0.1
u = 0.3
u = 0.5
u = 0.7
u = 0.9
u = 0.99

DGP 3 (potential misspecification)

(DGP3) The data is independently drawn from

$$Y_i = 1 + U_i(e^{2(Z_i+1)} - 1), \quad (Z_i, U_i) \sim \text{Beta}(3,3) \otimes U[0,1]$$

for $i = 1, \ldots, n$.

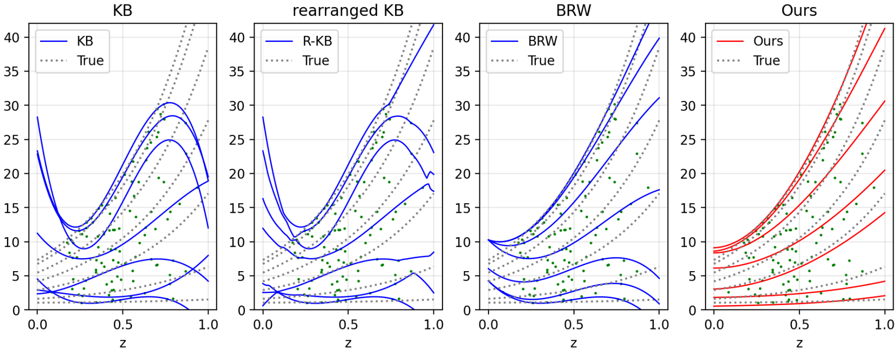
- Apply each estimator to the 3rd-order polynomial regression model

$$\beta(u) = \underset{\beta \in \mathbb{R}^4}{\operatorname{argmin}} \, \mathbb{E}\left[\rho_u\left(Y - X'\beta\right)\right],$$

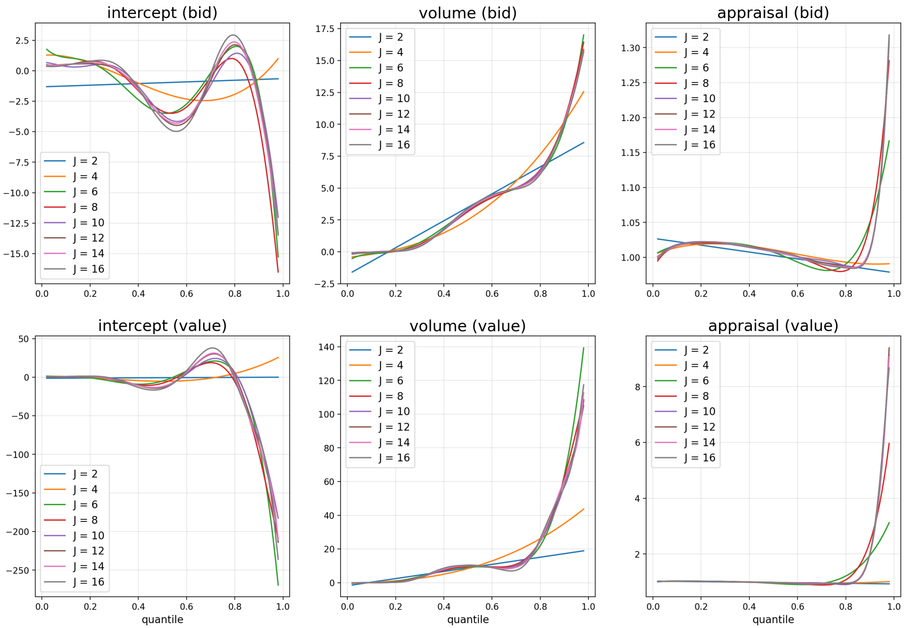
  where $X = (1, Z, Z^2, Z^3)'$.

- The map $z \mapsto 1 + u(e^{2(z+1)} - 1)$ is convex and increasing.
- Impose convexity and monotonicity in z plus non-crossing on $\hat{\beta}_n$.

DGP 3


KB
rearranged KB
BRW
Ours
KB
R-KB
BRW
Ours
True
40
35
30
25
20
15
10
5
0
0.0
0.5
1.0
z

Estimates of $\beta(\cdot)$ and $\gamma(\cdot)$ for $J = \{3, 6, 9, 12, 15\}$



Estimates of Q_B, Q_V, and ξ for $J = \{3, 6, 9, 12, 15\}$