Quantile on Quantiles

Martina Pons

University of Zurich

martina.pons@econ.uzh.ch

November 7th 2025

Motivation

- Since Koenker and Bassett (1978), quantile regression has been widely used for policy evaluation.
- Yet many real-world policy objectives are inherently **multidimensional**.
  - The **UN Sustainable Development Goals** call for reducing inequality "within and among countries."
  - The **EU Cohesion Policy** aims to foster convergence across regions; yet the within-region component cannot be ignored.
  - **Equality-of-opportunity** principles emphasize compensating for differences due to circumstances while respecting differences due to effort.

Motivation

- The relevance of both dimensions is also reflected in the applied literature, which generally examines heterogeneity along a single dimension.
- **Place-based policies** have been shown to:
  - stimulate local growth and employment in lagging regions (Becker et al., 2010; Busso et al., 2013; Ehrlich and Seidel, 2018),
  - but also increase within-region inequality (Lang et al., 2023; Albanese et al., 2023).
- $\rightarrow$ The two dimensions are **interdependent**: policies may improve outcomes along one dimension while worsening them along the other.
- To capture these trade-offs, we have to model both dimensions together.

This paper suggests a method to **simultaneously study distributional effects and inequalities within and between groups.**

Why Modeling Two-Dimensional Inequality is Challenging

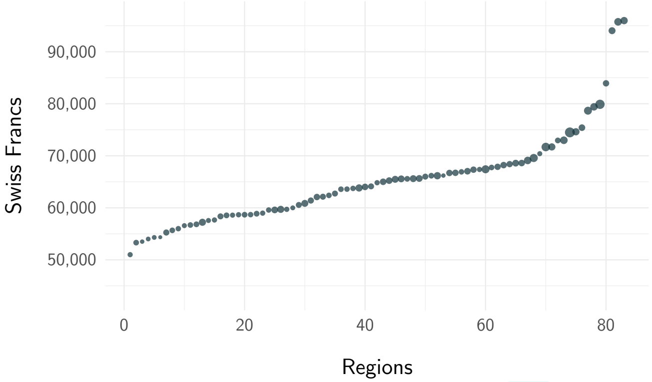
1. **Plausible assumptions only yield partial orderings of groups.**
   - A region can display high mobility for some parts of the parental income distribution but low mobility for others (Chetty and Hendren, 2018a,b).
   - Swiss regions.

Example - Yearly Income across Regions

- Researchers tackle this difficulty by focusing on the group mean or median outcome
  1. Compare average income across regions.
  2. Compare median income across regions.

Example - Yearly Income across Regions

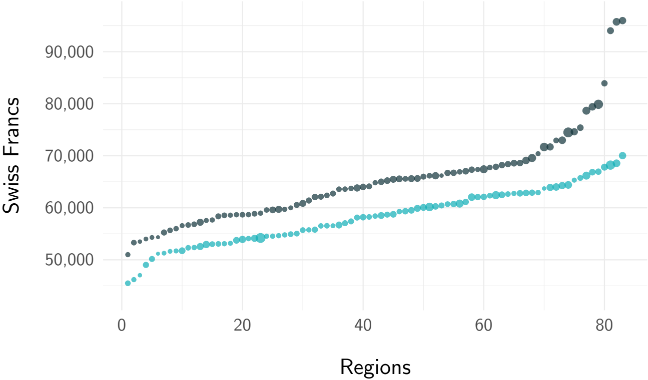
- Researchers tackle this difficulty by focusing on the group mean or median outcome
  1. Compare average income across regions.
  2. Compare median income across regions.



Data: Swiss Federal Statistical Office. Regions are defined by 2-digit ZIP codes. obs = 4.2 million. Data

Example - Yearly Income across Regions

- Researchers tackle this difficulty by focusing on the group mean or median outcome
  1. Compare average income across regions.
  2. Compare median income across regions.



Data: Swiss Federal Statistical Office. Regions are defined by 2-digit ZIP codes. obs = 4.2 million. Data

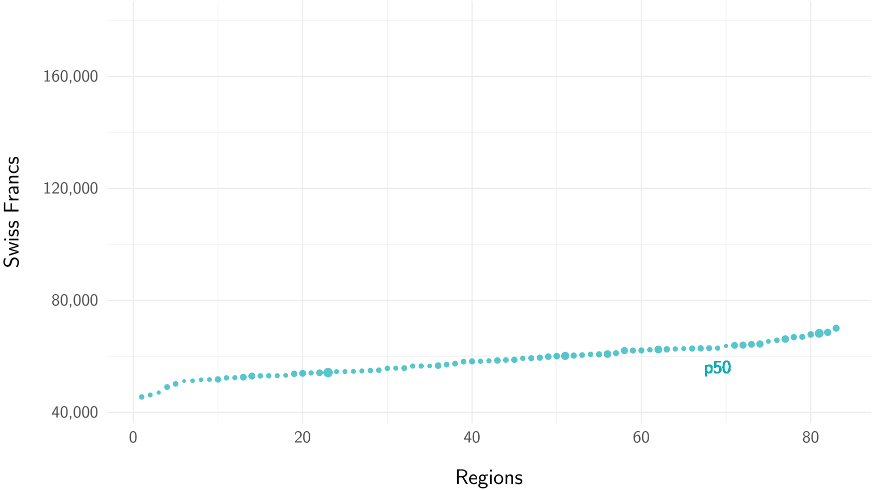
Example - Yearly Income across Regions

- Researchers tackle this difficulty by focusing on the group mean or median outcome
  1. Compare average income across regions.
  2. Compare median income across regions.
- **Limitations**: Both measures impose strong assumptions on the welfare function
  1. Averages ignore the distributional shape.
  2. Median solely reflects the heterogeneity at one point of the distribution, potentially overlooking the labor market situation of a considerable portion of workers.

Example - Yearly Income across Regions

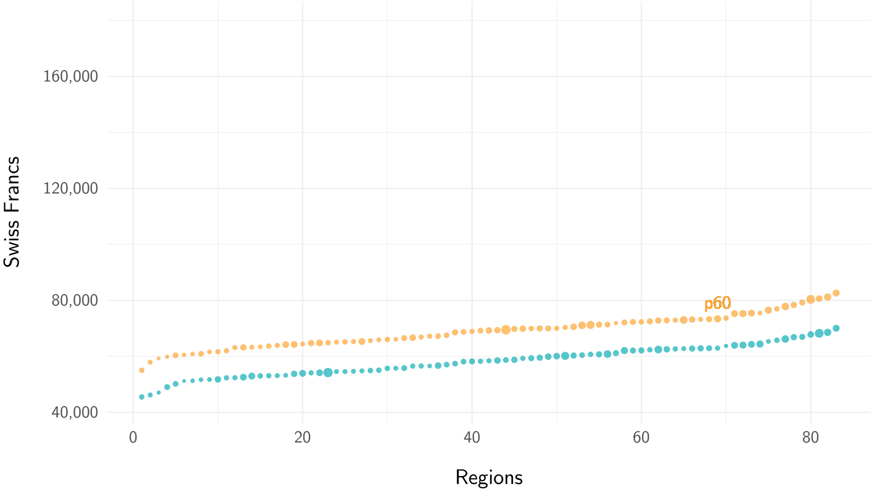
- Researchers tackle this difficulty by focusing on the group mean or median outcome
  1. Compare average income across regions.
  2. Compare median income across regions.
- **Limitations**: Both measures impose strong assumptions on the welfare function
  1. Averages ignore the distributional shape.
  2. Median solely reflects the heterogeneity at one point of the distribution, potentially overlooking the labor market situation of a considerable portion of workers.
- **Solution**: analyze between heterogeneity at different points of the within distribution using a **two-level quantile function.**

Example - Yearly Income across Regions



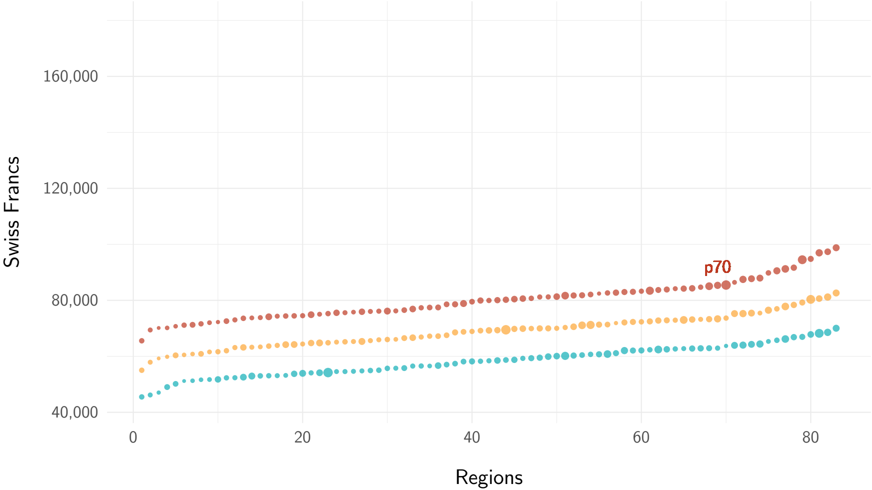
Data: Swiss Federal Statistical Office. Regions are defined by 2-digit ZIP codes. obs = 4.2 million. Data

Example - Yearly Income across Regions



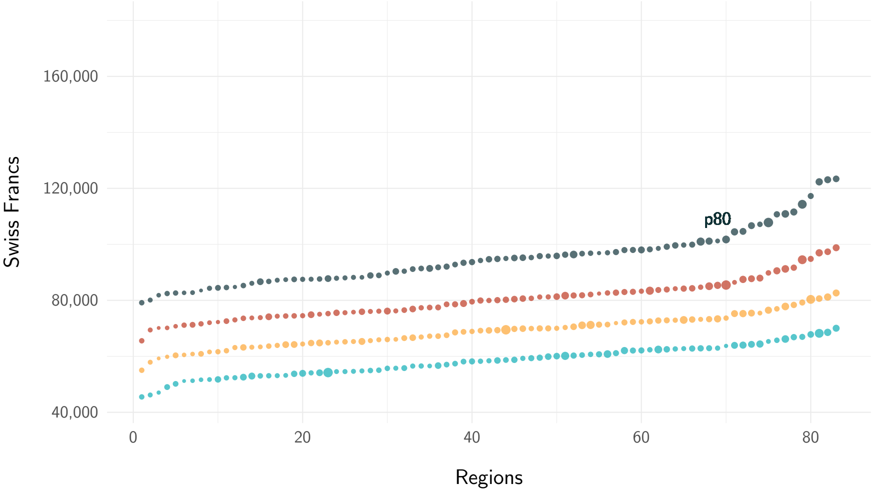
Data: Swiss Federal Statistical Office. Regions are defined by 2-digit ZIP codes. obs = 4.2 million. Data

Example - Yearly Income across Regions



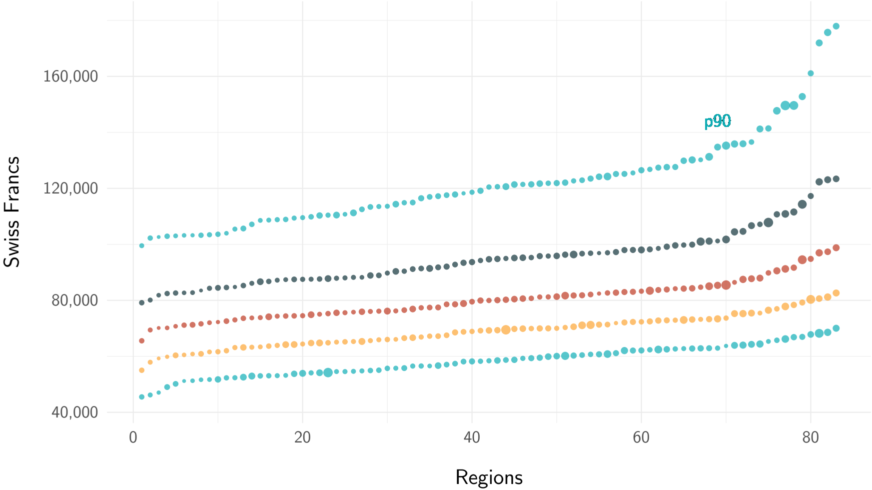
Data: Swiss Federal Statistical Office. Regions are defined by 2-digit ZIP codes. obs = 4.2 million. Data

Example - Yearly Income across Regions



Data: Swiss Federal Statistical Office. Regions are defined by 2-digit ZIP codes. obs = 4.2 million. Data

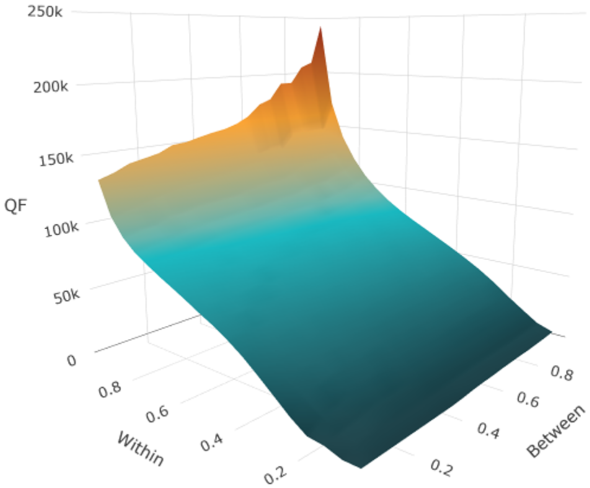
Example - Yearly Income across Regions



Data: Swiss Federal Statistical Office. Regions are defined by 2-digit ZIP codes. obs = 4.2 million. Data

Two-Dimensional Quantile Function

Data



Why Modeling Two-Dimensional Inequality is Challenging

1. **Plausible assumptions only yield partial orderings of groups.**
   - A region can display high mobility for some parts of the parental income distribution but low mobility for others (Chetty and Hendren, 2018a,b).
   - Swiss Regions example.
2. **Comparisons are incomplete without additional normative structure.**
   - Evaluating inequality across multiple dimensions requires assumptions about how society trades off improvements in one dimension against deteriorations in another (Atkinson and Bourguignon, 1987). More

This paper...

Suggests a method to **simultaneously** study **distributional effects within and between groups**.

Contribution

1. Construct an **outcome model** that captures the complete distributional structure and allows for unrestricted heterogeneity across groups.
   - Outcomes are summarized by a **two-dimensional quantile function** reflecting within- and between-group heterogeneity.
2. Introduce a flexible and tractable **welfare criterion**.
   - *Generalized social marginal welfare weights* (Saez and Stantcheva, 2016) explicitly model how society trades off between within- and between-group inequality.
   - The two-dimensional quantile function is the unique minimal sufficient statistic for welfare comparison within a broad class of social welfare criteria.
3. Propose a **two-step quantile regression estimator** with within-group regressions in the first stage and between-group regressions in the second stage, and derive **uniform asymptotic results**.

Today's Presentation

- Literature Review
- Outcome and Welfare Model
- Quantile Model and Estimator
- Asymptotic Results
- Empirical Application

Related Econometrics Literature

- Within Distribution and Quantile Panel Data Models (Galvao and Wang, 2015; Chetverikov, Larsen, and Palmer, 2016; Melly and Pons, 2025).
  - Model also the between distribution. More on Melly and Pons (2024)
- Multidimensional heterogeneity (Arellano and Bonhomme, 2016; Frumento, Bottai, and Fernández-Val, 2021; Liu, 2024; Fernández-Val, Gao, Liao, and Vella, 2022).
  - Allow the effect of individual-level and group-level variables to vary across *both* dimensions.
- Quantile regression with generated dependent variables/regressors (Chen et al., 2003; Ma and Koenker, 2006; Bhattacharya, 2020; Chen et al., 2021).
  - Provide uniform asymptotic results for the entire quantile regression process.

An Outcome Model

Let $j = 1, \ldots, m$ be the groups and $i = 1, \ldots, n$ be the individuals.

Let each individual's outcome be

$$y_{ij} = q(u_{ij}, v_j),$$

- u_{ij}: within-group rank
- v_j: is a vector containing group characteristics or circumstances.

An Outcome Model

Let $j = 1, \ldots, m$ be the groups and $i = 1, \ldots, n$ be the individuals.

Let each individual's outcome be

$$y_{ij} = q(u_{ij}, v_j),$$

- u_{ij}: within-group rank
- v_j: is a vector containing group characteristics or circumstances.

Goal: construct a bivariate function.

An Outcome Model

Let $j = 1, \ldots, m$ be the groups and $i = 1, \ldots, n$ be the individuals.

Let each individual's outcome be

$$y_{ij} = q(u_{ij}, v_j),$$

- u_{ij}: within-group rank
- v_j: is a vector containing group characteristics or circumstances.

Goal: construct a bivariate function.

- **Within dimension**: $u_{ij} | v_j \sim U(0, 1)$, and impose strict monotonicity of $q(\cdot, v_j)$. Yields a group-level quantile function $q(u, v_j)$.

An Outcome Model

Let $j = 1, \ldots, m$ be the groups and $i = 1, \ldots, n$ be the individuals.

Let each individual's outcome be

$$y_{ij} = q(u_{ij}, v_j),$$

- u_{ij}: within-group rank
- v_j: is a vector containing group characteristics or circumstances.

Goal: construct a bivariate function.

- **Within dimension**: $u_{ij}|v_j \sim U(0,1)$, and impose strict monotonicity of $q(\cdot, v_j)$. Yields a group-level quantile function $q(u, v_j)$.
- **Between dimensions**: A scalar v_j would not work!

A Naive Model

Consider a naive version of the model

$$y_{ij} = q(u_{ij}, v_j),$$

where $q(\cdot)$ is also strictly increasing in **scalar** v_j.

A Naive Model

Consider a naive version of the model

$$y_{ij} = q(u_{ij}, v_j),$$

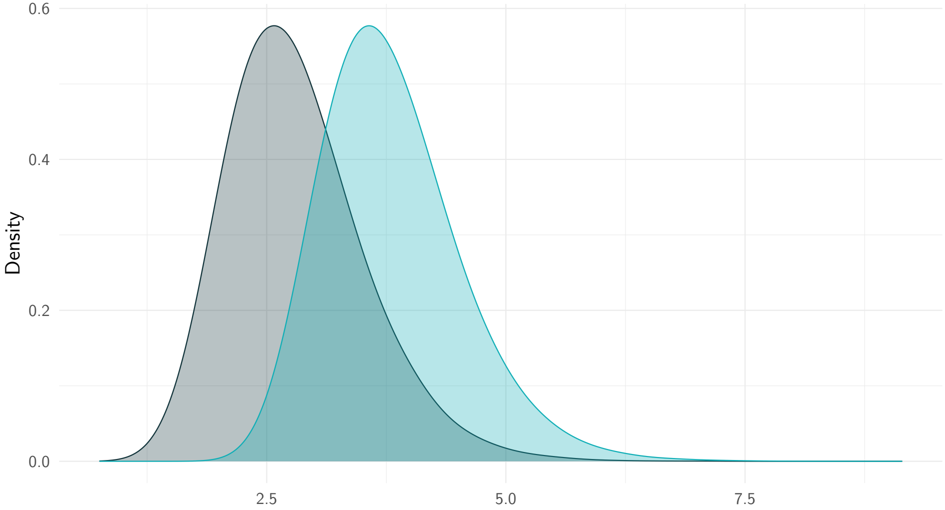
where $q(\cdot)$ is also strictly increasing in **scalar** v_j.

Take two groups $j = \{h, l\}$ with $v_h > v_l$, then strict monotonicity w.r.t. v_j implies

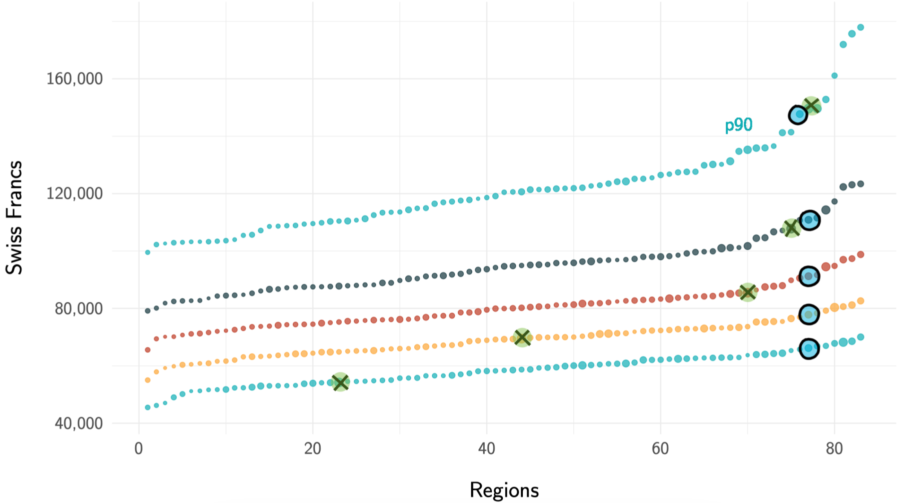
$$q(v_h, u) > q(v_l, u), \quad \text{for all } u \in (0, 1)$$

$\rightarrow$ Groups can be ordered unambiguously.

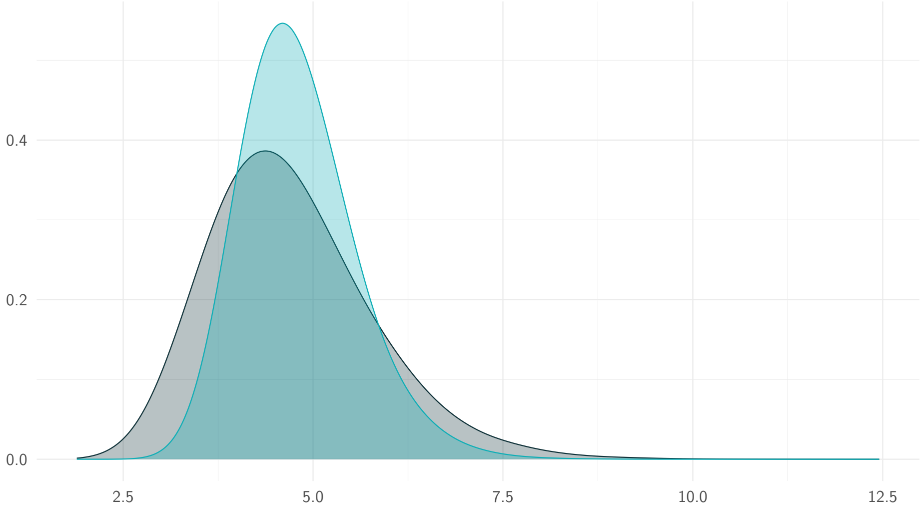
One-dimensional v_j



Why a Scalar v_j Would Not Work



Two-dimensional v_j



Outcome Model

Let v_j be a vector.

Even if v_j is multidimensional, after fixing u we can find a scalar valued function $v_j(u)$ such that

$$q(u, v_j) = q(u, v_j(u)).$$

- This reparameterization imposes no restriction on the model.
- Simply maps multidimensional v_j into a single index.

Outcome Model

Let v_j be a vector.

Even if v_j is multidimensional, after fixing u we can find a scalar valued function $v_j(u)$ such that

$$q(u, v_j) = q(u, v_j(u)).$$

- This reparameterization imposes no restriction on the model.
- Simply maps multidimensional v_j into a single index.

Normalize $v_j(u) \sim U(0,1)$ and assume $q(\cdot,\cdot)$ is increasing in both arguments.
Result: $q(u, v)$ summarizes the entire joint distribution: for each u, it records how the uth group-specific quantiles vary across groups through the dependency on v.

Social Welfare

Welfare is written in terms of *marginal social welfare weights* (Saez and Stantcheva, 2016):

$$W = \int_0^1 \int_0^1 w(u, v)\, q(u, v)\, du\, dv,$$

where $w(u, v) \geq 0$ denotes the social marginal welfare weight assigned to the individual at within-group rank u and group rank v.

- Welfare is a weighted average of the outcomes with weights depending on both rank variables.
- Weights are typically decreasing in both u and v
  - Reflects concern for inequality within and between groups.
- Weights are not necessarily decreasing in the outcome level itself
  - Society may not view all inequalities as equally problematic.

Social Welfare

Many functional forms for $w(u, v)$ are possible, each reflecting different trade-offs and areas of focus.

Examples:

- Two-dimensional Gini Social Welfare Function more
- Equality of Opportunity (Roemer, 1998). more
- Utilitarian more
- (unconditional) rank-dependent welfare function. more

$q(u, v)$ is the unique minimal sufficient statistic for welfare comparison within a broad class of social welfare criteria. Formal Result

$q(u, v)$ as the empirical primitive: once it is known, any welfare evaluation can be computed.

Distributional Policy Evaluation

Consider a policy indexed by $D \in \{0, 1\}$. Assuming that the potential outcome surfaces $q_d(u, v)$ are identified, the welfare impact of the policy is

$$\Delta W = \int_0^1 \int_0^1 w(u, v) \left[q_1(u, v) - q_0(u, v)\right] du\, dv.$$

Hence, this provides a complete statistic for assessing how policies affect welfare across multiple dimensions of heterogeneity.

Quantile Model

Generalize the model to include covariates:

$$
\begin{aligned}
y_{ij} &= q(x_{ij}, v_j, u_{ij}) \\
&= x_{ij}'\beta(u_{ij}, v_j) + \alpha(u_{ij}, v_j)
\end{aligned}
$$

- x_{ij}: vector of covariates
- $\alpha(u_{ij}, v_j)$: intercept.

Normalize

$$
\begin{aligned}
&u_{ij} | x_{ij}, v_j \sim U(0,1) \\
&v_j(u) | x_{ij} \sim U(0,1), \text{ for each } u \in (0,1)
\end{aligned}
$$

Two-Dimensional Quantile Function

Conditional on x_{ij} and v_j, $q(x_{ij}, v_j, u_{ij})$ is strictly monotonic with respect to u_{ij} so that

$$\begin{aligned} Q(u, y_{ij}|x_{ij}, v_j) =& q(x_{ij}, v_j, u) \\ =& x_{1ij}'\beta(u, v_j) + \alpha(u, v_j) \end{aligned}$$

defines the u-conditional quantile function of y_{ij} conditional on x_{ij}, and v_j.

Two-Dimensional Quantile Function

Conditional on x_{ij} and v_j, $q(x_{ij}, v_j, u_{ij})$ is strictly monotonic with respect to u_{ij} so that

$$\begin{aligned} Q(u, y_{ij}|x_{ij}, v_j) =& q(x_{ij}, v_j, u) \\ =& x_{1ij}'\beta(u, v_j) + \alpha(u, v_j) \end{aligned}$$

defines the u-conditional quantile function of y_{ij} conditional on x_{ij}, and v_j.

By the same argument, the v-conditional quantile function $Q(u, y_{ij} \mid x_{ij}, v_j)$ is defined by:

$$\begin{aligned} Q(v, Q(u, y_{ij}|x_{1ij}, v_j)|x_{ij}) &= q(x_{ij}, v, u) \\ &= x_{ij}'\beta(u, v) + \alpha(u, v). \end{aligned}$$

Interpretation of the coefficients

- $\beta(u, v)$ tells how the (u, v)-conditional quantile function responds to a change in x_{ij} by one unit.
- $\beta(0.5, v)$ gives the effect of x_{ij} on the **conditional quantile function of group medians**, with groups with the highest medians positioned at the top and those with the lowest medians at the bottom of the distribution.

Estimator

1. **First stage**: group-by-group quantile regression of the outcome on x_{ij} for quantiles u. For each group j and quantile u:

$$\hat{\beta}_j(u) \equiv \left(\hat{\beta}_{1,j}(u), \hat{\beta}_{2,j}(u)'\right)' = \underset{(b_1,b_2)\in\mathbb{R}^{dim(x)+1}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \rho_u(y_{ij} - b_1 - x_{ij}' b_2),$$

where $\rho_u(x) = (u - 1\{x < 0\})x$ for $x \in \mathbb{R}$ is the check function.

Estimator

1. **First stage**: group-by-group quantile regression of the outcome on x_{ij} for quantiles u. For each group j and quantile u:

$$\hat{\beta}_j(u) \equiv \left(\hat{\beta}_{1,j}(u), \hat{\beta}_{2,j}(u)'\right)' = \underset{(b_1,b_2)\in\mathbb{R}^{dim(x)+1}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \rho_u(y_{ij} - b_1 - x_{ij}'b_2),$$

where $\rho_u(x) = (u - 1\{x < 0\})x$ for $x \in \mathbb{R}$ is the check function. Save the fitted values for each quantile and each group j.

Estimator

1. **First stage**: group-by-group quantile regression of the outcome on x_{ij} for quantiles u. For each group j and quantile u:

$$\hat{\beta}_j(u) \equiv \left(\hat{\beta}_{1,j}(u), \hat{\beta}_{2,j}(u)'\right)' = \underset{(b_1,b_2)\in\mathbb{R}^{dim(x)+1}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \rho_u(y_{ij} - b_1 - x_{ij}' b_2),$$

where $\rho_u(x) = (u - 1\{x < 0\})x$ for $x \in \mathbb{R}$ is the check function. Save the fitted values for each quantile and each group j.

2. **Second stage**: for each quantile u regress the first-stage fitted values on x_{ij} using quantile regression for each quantile v:

$$\hat{\delta}(\hat{\beta}(u), v) = \underset{(a,b)\in\mathbb{R}^{dim(x)+1}}{\arg\min} \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} \rho_v(\hat{y}_{ij}(u) - x_{ij}' b - a),$$

where $\delta = (\alpha, \beta')'$ and $\hat{y}_{ij}(u) = \hat{\beta}_{1,j}(u) + x_{ij}'\hat{\beta}_{2,j}(u)$.

Estimator - Example

- $m = 100 \implies$ 100 groups
- quantile of interest: $\{0.1, 0.2, \ldots, 0.9\} \implies$ 9 quantiles of interest.

1. **First stage**: **9** group-by-group quantile regression of the y_{ij} on x_{ij}. ($\mathbf{9} \times \mathbf{100} = \mathbf{900}$ first step regressions).
   Obtain **9** vectors of fitted values.
2. **Second stage**: for each quantile u regress the first-stage fitted values (**9** vectors) on x_{ij} using quantile regression for each decile $\{0.1, 0.2, \ldots, 0.9\}$. ($\mathbf{9} \times \mathbf{9} = \mathbf{81}$ second step regressions)

Computing time

Asymptotics

- Show uniform consistency and weak convergence of the entire quantile regression process.
- Asymptotic framework where n and $m \to \infty$.
- Suggest testing procedure to test for uniform hypotheses.

Challenges:

- Non-smooth quantile regression objective function.
- Generated dependent variable.
- Dimension of the first stage increases with the number of groups.
- Different rate of convergence of first step estimator.

Use results in Chen, Linton, and Van Keilegom (2003); Angrist, Chernozhukov, and Fernández-Val (2006); Volgushev, Chao, and Cheng (2019); Galvao, Gu, and Volgushev (2020).

Asymptotic Distribution

Let $\mathcal{T}$ be a compact subset of $(0,1)$. Show that uniformly in $\tau = (u, v) \in \mathcal{T} \times \mathcal{T}$,

$$\sqrt{m}\left(\hat{\delta}(\hat{\beta},\tau) - \delta_0(\beta_0,\tau)\right)$$

$$= -\Gamma_1(\delta_0,\beta_0,\tau)^{-1}\sqrt{m}\left(\frac{1}{m}\sum_{j=1}^{m}\bar{\Gamma}_{2j}(\delta_0,\beta_0,\tau)[\hat{\beta}_j(u) - \beta_{j,0}(u)] + M_{mn}(\delta_0,\beta_0,\tau)\right)$$

$$+ \underbrace{o_p(1)}_{\text{negligible}}$$

1. In blue: first-stage error
2. In yellow: second-stage noise

The first-stage quantile regression bias is of order $1/\sqrt{n} \implies$ the number of observations per group must diverge to infinity.

more

Asymptotic Distribution

If $\frac{\sqrt{m}\log n}{n} \to 0$ and other assumptions are satisfied ▸ more

First stage error:

$$\sup_{\tau \in \mathcal{T}\times\mathcal{T}} \left\| \frac{1}{m}\sum_{j=1}^{m} \bar{\Gamma}_{2j}(\delta_0, \beta_0, \tau)\left(\hat{\beta}_j(u) - \beta_{j,0}(u)\right) \right\| = o_p\left(\frac{1}{\sqrt{m}}\right), \tag{1}$$

Second stage noise:

$$\sqrt{m}\,(M_{mn}(\delta_0, \beta_0, \cdot)) \rightsquigarrow \mathbb{G}(\cdot), \text{ in } \ell^\infty(\mathcal{T}\times\mathcal{T}),$$

where $\mathbb{G}$ is a mean-zero Gaussian process with a uniformly continuous sample path and covariance function $\Omega_2(\tau, \tau') = (\min(v, v') - vv')\mathbb{E}[x_{ij}x_{ij}']$.

Asymptotic Distribution

If $\frac{\sqrt{m}\log n}{n} \to 0$ and other assumptions are satisfied more

First stage error:

$$\sup_{\tau \in \mathcal{T} \times \mathcal{T}} \left\| \frac{1}{m} \sum_{j=1}^{m} \bar{\Gamma}_{2j}(\delta_0, \beta_0, \tau) \left(\hat{\beta}_j(u) - \beta_{j,0}(u) \right) \right\| = o_p\left(\frac{1}{\sqrt{m}} \right), \tag{1}$$

Second stage noise:

$$\sqrt{m}\left(M_{mn}(\delta_0, \beta_0, \cdot)\right) \rightsquigarrow \mathbb{G}(\cdot), \text{ in } \ell^{\infty}(\mathcal{T} \times \mathcal{T}),$$

where $\mathbb{G}$ is a mean-zero Gaussian process with a uniformly continuous sample path and covariance function $\Omega_2(\tau, \tau') = (\min(v, v') - vv')\mathbb{E}[x_{ij} x_{ij}']$.
Hence,

$$\sqrt{m}\left(\hat{\delta}(\hat{\beta}, \cdot) - \delta_0(\beta_0, \cdot)\right) \rightsquigarrow \Gamma_1^{-1}(\cdot)\mathbb{G}(\cdot) \quad \text{in } \ell^{\infty}(\mathcal{T} \times \mathcal{T}),$$

with $\Gamma_1 = \Gamma_1(\delta_0, \beta_0, \tau)$. Degenerate Distribution Inference

Inference

- I suggest a clustered bootstrap procedure, where entire groups are resampled with replacement.
- First stage is unaffected; hence, fitted values can be resampled.
- I prove the validity of the bootstrap.
- Functional inference:
  - Kolmogorov-Smirnov and Cramér-von-Mises Tests for homogeneity over (u, v). Critical values are estimated using bootstrap. More on KS and CvM Tests
  - Functional confidence band can be constructed by inverting the acceptance region of the Kolmogorov-Smirnov test statistic (Chernozhukov et al., 2013). More functional Confidence Intervals

Empirical Application

- Build on McKenzie and Puerto (2021).
- Estimate the impact of business training on the outcomes of female-owned businesses.
- Sample: 2,922 female-owned businesses operating in 116 different rural markets in Kenya.
- Two-stage randomization:
  1. market-level randomization (markets are assigned to treatment or control markets).
  2. individual-level randomization (firms in the treatment markets are randomly assigned to training).
- Estimate distributional effects both within and between markets.
- Outcome variable: Income from Work.

Empirical Application

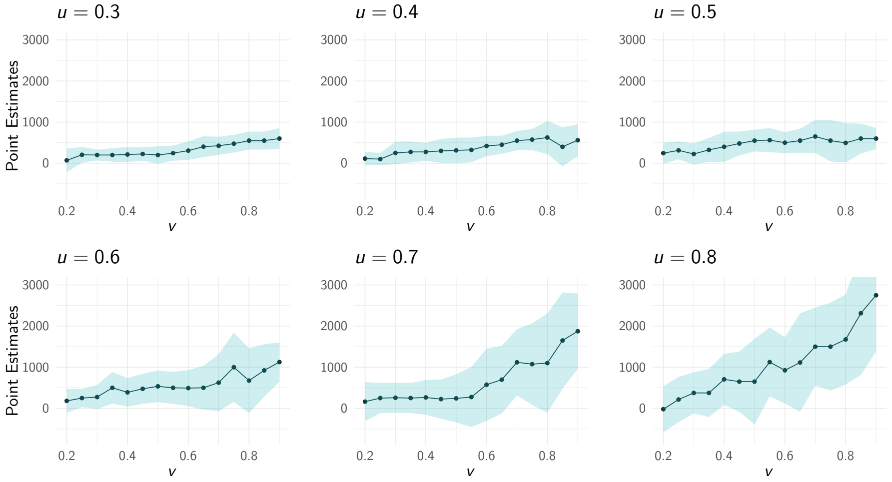
Specification:

$$y_{ij} = \beta_1(u_{ij}, v_j) \cdot D_{ij} + \beta_2(u_{ij}, v_j) \cdot S_{ij} + \alpha(u_{ij}, v_j),$$

- y_{ij}: outcome of firm i operating in market j.
- D_{ij}: treatment indicator.
- S_{ij} binary variable that accounts for potential spillover effects ($= 1$ for individuals in the treatment markets that are assigned to the control group).

Results - Income from Work

▸ More ▸ Rank Corr. ▸ H_0 : Effect Homogeneity ▸ Computing Time ▸ Welfare



1,000 Kenyan Shilling = 7.74 USD

Results - Welfare Gain Under Different Weighting Schemes

Realized outcome vs. counterfactual scenario without treatment intervention.

$$W = \int_0^1 \int_0^1 q(u,v) \cdot w(u,v) dv du,$$

where $w(u,v) = 2(1 - \omega u - (1-\omega)v)$, with $\omega \in \{0.2, 0.5, 0.8\}$.

Weighting Scheme	Welfare Gain (%)
$\omega = 0.2$	11.53
$\omega = 0.5$	13.14
$\omega = 0.8$	15.16
Utilitarian ($w = 1$)	15.33

Back

Conclusion

- Distributional treatment effects are particularly interesting when analyzing treatment effect heterogeneity.
- Heterogeneity manifests itself across various dimensions.
- This paper suggests a method to simultaneously study distributional effects within and between groups while remaining agnostic about social welfare function.
  - Allows us to consider trade-offs between different components of inequality.
  - Ranking groups is a nontrivial task without assuming a welfare function.
- Monte Carlo simulations show good finite sample performance. ▸ Simulations

▸ FAQ

References I

Albanese, G., G. Barone, and G. de Blasio (2023): "The impact of place-based policies on interpersonal income inequality," *Economica*, 90, 508–530.

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

Arellano, M. and S. Bonhomme (2016): "Nonlinear panel data estimation via quantile regressions," *Econometrics Journal*, 19, C61–C94.

Atkinson, A. B. and F. Bourguignon (1987): "Income Distribution and Differences in Needs," in *Arrow and the Foundations of the Theory of Economic Policy*, ed. by G. R. Feiwel, London: Palgrave Macmillan UK, 350–370.

Becker, S. O., P. H. Egger, and M. von Ehrlich (2010): "Going NUTS: The effect of EU Structural Funds on regional performance," *Journal of Public Economics*, 94, 578–590.

Bhattacharya, J. (2020): "Quantile regression with generated dependent variable and covariates," *Working Paper*.

Busso, M., J. Gregory, and P. Kline (2013): "Assessing the incidence and efficiency of a prominent place based policy," *American Economic Review*, 103, 897–947.

Chen, L., A. F. Galvao, and S. Song (2021): "Quantile regression with generated regressors," *Econometrics*, 9.

Chen, X., O. Linton, and I. Van Keilegom (2003): "Estimation of semiparametric models when the criterion function is not smooth," *Econometrica*, 71, 1591–1608.

References II

CHERNOZHUKOV, V. AND I. FERNÁNDEZ-VAL (2005): "Subsampling inference on quantile regression processes," *Sankhya: The Indian Journal of Statistics*, 67, 253–276.

CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND A. GALICHON (2009): "Improving point and interval estimators of monotone functions by rearrangement," *Biometrika*, 96, 559–575.

——— (2010): "Quantile and Probability Curves Without Crossing," *Econometrica*, 78, 1093–1125.

CHERNOZHUKOV, V., I. FERNÁNDEZ-VAL, AND B. MELLY (2013): "Inference on Counterfactual Distributions," *Econometrica*, 81, 2205–2268.

CHERNOZHUKOV, V. AND C. HANSEN (2005): "An IV Model of Quantile Treatment Effects," *Econometrica*, 73, 245–261.

CHETTY, R. AND N. HENDREN (2018a): "The Impact of Neighborhoods on Intergenerational Mobility I: County-Level Estimates," *Quarterly Journal of Economics*, 133, 1107–1162.

——— (2018b): "The impacts of neighborhoods on intergenerational mobility II: County-level How are children," *The Quarterly Journal of Economics*, 133, 1163–1228.

CHETVERIKOV, D., B. LARSEN, AND C. PALMER (2016): "IV Quantile Regression for Group-Level Treatments, With an Application to the Distributional Effects of Trade," *Econometrica*, 84, 809–833.

DONG, Y. AND S. SHEN (2018): "Testing for rank invariance or similarity in program evaluation," *Review of Economics and Statistics*, 100, 78–85.

References III

Ehrlich, M. V. and T. Seidel (2018): "The persistent effects of place-based policy: Evidence from the West-German Zonenrandgebiet," *American Economic Journal: Economic Policy*, 10, 344–374.

Fernández-Val, I., W. Y. Gao, Y. Liao, and F. Vella (2022): "Dynamic Heterogeneous Distribution Regression Panel Models, with an Application to Labor Income Processes," *Working Paper*, 1–45.

Frandsen, B. R. and L. J. Lefgren (2018): "Testing Rank Similarity," *The Review of Economics and Statistics*, 100, 86–91.

Franguridi, G., B. Gafarov, and K. Wüthrich (2024): "Bias correction for quantile regression estimators," *Working Paper*.

Frumento, P., M. Bottai, and I. Fernández-Val (2021): "Parametric Modeling of Quantile Regression Coefficient Functions With Longitudinal Data," *Journal of the American Statistical Association*, 116, 783–797.

Galvao, A. F., J. Gu, and S. Volgushev (2020): "On the unbiased asymptotic normality of quantile regression with fixed effects," *Journal of Econometrics*, 218, 178–215.

Galvao, A. F. and K. Kato (2016): "Smoothed quantile regression for panel data," *Journal of Econometrics*, 193, 92–112.

Galvao, A. F. and L. Wang (2015): "Efficient Minimum Distance Estimator for Quantile Regression Fixed Effects Panel Data," *Journal of Multivariate Analysis*, 133, 1–26.

Kaji, T. and J. Cao (2023): "Assessing Heterogeneity of Treatment Effects," *Working Paper*, 1–22.

References IV

Kim, J. H. and B. G. Park (2022): "Testing rank similarity in the local average treatment effects model," *Econometric Reviews*, 41, 1265–1286.

Kitagawa, T. and A. Tetenov (2021): "Equality-Minded Treatment Choice," *Journal of Business and Economic Statistics*, 39, 561–574.

Koenker, R. and G. Bassett (1978): "Regression Quantiles," *Econometrica*, 46, 33.

Lang, V., N. Redeker, and D. Bischof (2023): "Place-based Policies and Inequality Within Regions," *OSF Reprints*.

Liao, Y. and X. Yang (2018): "Uniform Inference for Characteristic Effects of Large Continuous-Time Linear Models," *Working Paper*, 1–52.

Liu, X. (2024): "A quantile-based nonadditive fixed effects model," *Working Paper*, 1–34.

Lu, X. and L. Su (2023): "Uniform inference in linear panel data models with two-dimensional heterogeneity," *Journal of Econometrics*, 235, 694–719.

Ma, L. and R. Koenker (2006): "Quantile regression methods for recursive structural equation models," *Journal of Econometrics*, 134, 471–506.

McKenzie, D. and S. Puerto (2021): "Growing Markets through Business Training for Female Entrepreneurs: A Market-Level Randomized Experiment in Kenya," *American Economic Journal: Applied Economics*, 13, 297–332.

Melly, B. and M. Pons (2025): "Minimum Distance Estimation of Quantile Panel Data Models," *Working Paper*.

Roemer, J. E. (1998): *Equality of Opportunity*.

References V

SAEZ, E. AND S. STANTCHEVA (2016): "Generalized social marginal welfare weights for optimal tax theory," *American Economic Review*, 106, 24–45.

VOLGUSHEV, S., S.-K. CHAO, AND G. CHENG (2019): "Distributed inference for quantile regression processes," *The Annals of Statistics*, 47, 1634–1662.

Example I - Income heterogeneity within and between regions

- Groups: 83 Swiss regions (2-digit zip code)
- Data: Administrative data on the universe of Swiss residents
- Restrict to individuals aged 29 to 64 (4.2 million observations)

back

Rank Correlation - Income from Work

Table: Correlation of Ranks over u

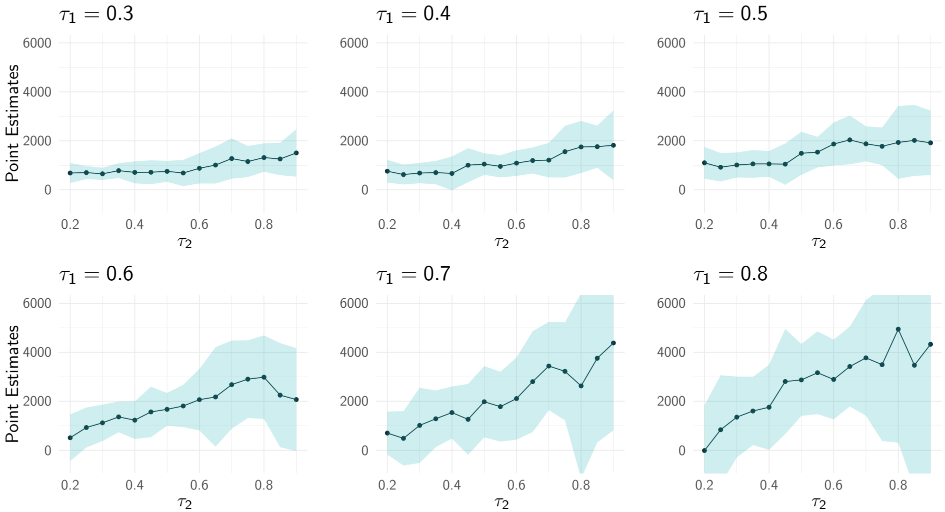
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.2	1							
0.3	0.74	1						
0.4	0.65	0.87	1					
0.5	0.53	0.76	0.85	1				
0.6	0.49	0.66	0.72	0.82	1			
0.7	0.42	0.6	0.66	0.69	0.83	1		
0.8	0.36	0.51	0.58	0.62	0.77	0.88	1	
0.9	0.32	0.44	0.42	0.47	0.59	0.6	0.69	1

Note:
The table shows the correlation matrix of the ranks at different values of u.

Back

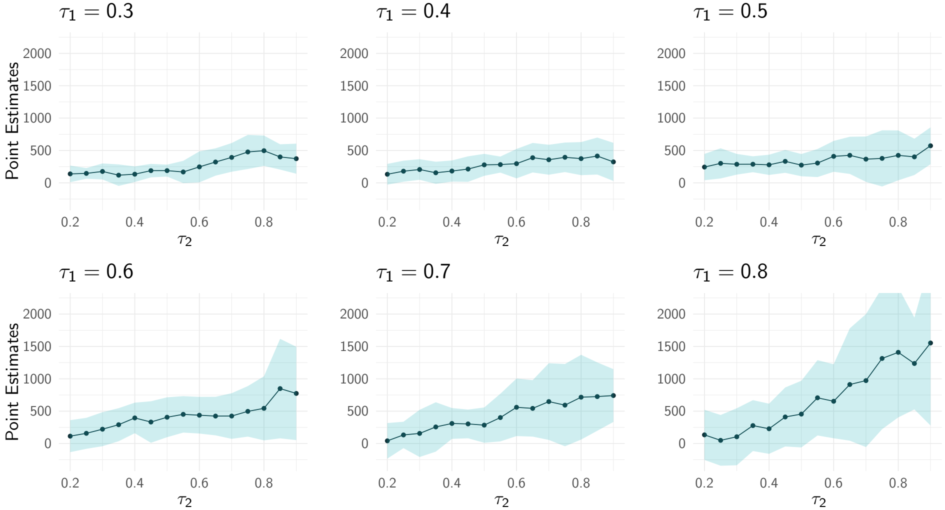
Additional Results - Sales

Back



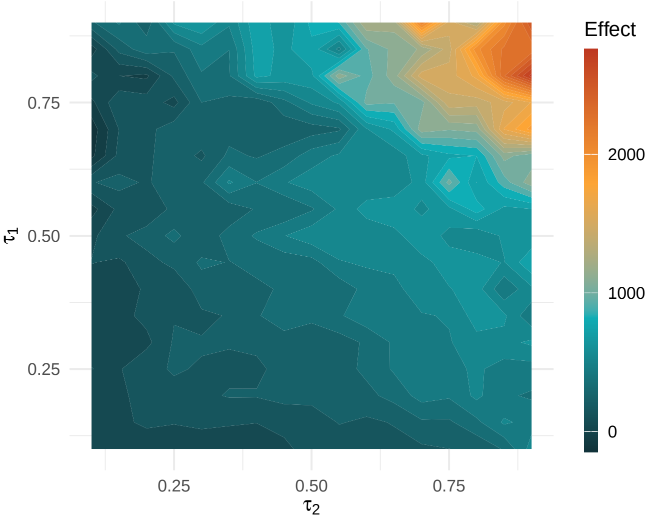
Additional Results - Profits

Back



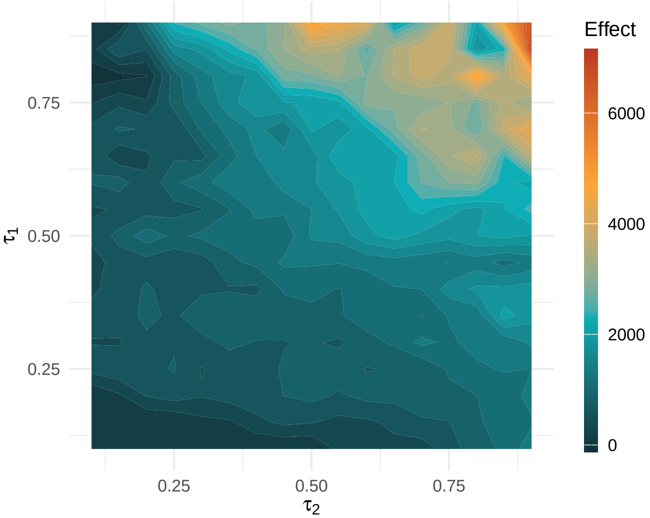
Additional Results - Income from Work

Back



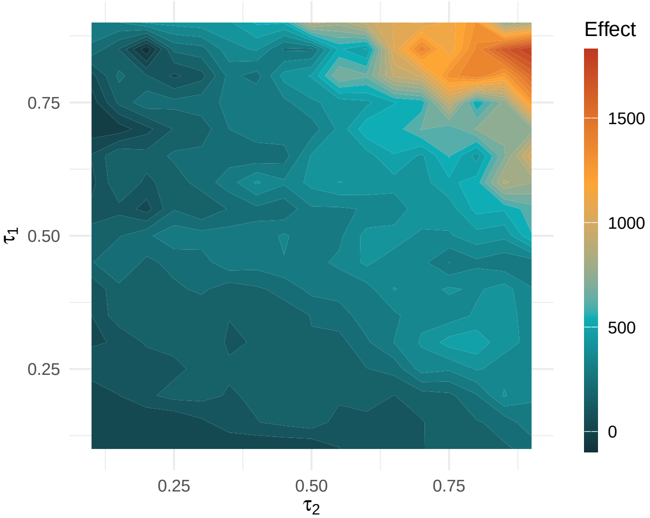
Additional Results - Sales

▸ Back



Additional Results - Profits

Back



Test of the H_0 of Homogeneous Effects Homogeneity

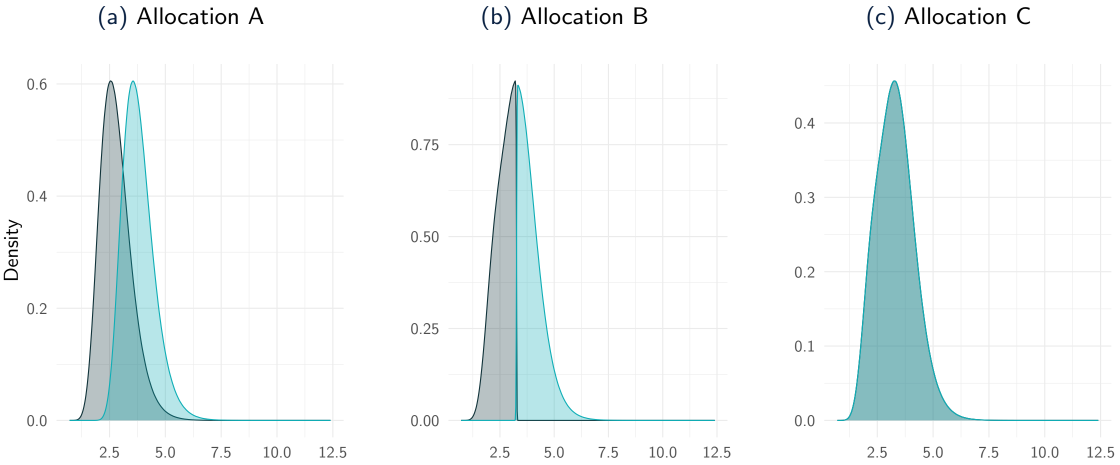
Table: *P*-Values of Cramér-von Mises and Kolmogorov-Smirnov Tests

	Income	Profits	Sales
Cramér-von Mises	0.024	0.027	0.024
Kolmogorov-Smirnov	0.006	0.009	0.012

Note:
The table shows the p-values of the Cramér-von Mises and Kolmogorov-Smirnov tests for the null hypothesis that the coefficients are homogeneous over both dimensions. The test is performed with the parametric bootstrap with 1000 replications.

▸ Back

Why Modeling Two-Dimensional Inequality is Challenging



back

Minimal Sufficiency of $q(u, v)$ I

Definition (Class of welfare functionals)

Let $\mathcal{W} = L^1_+((0,1)^2)$ denote the set of nonnegative integrable weight functions on $(0,1)^2$. Each $w \in \mathcal{W}$ defines a welfare functional as in Equation (14), for any measurable $q : (0,1)^2 \to \mathbb{R}$ such that $W_w(q) < \infty$. Two outcome surfaces q_1, q_2 are *$\mathcal{W}$-equivalent* if $W_w(q_1) = W_w(q_2)$ for all $w \in \mathcal{W}$.

◂ Back

Minimal Sufficiency of $q(u, v)$ II

Theorem

Let $\mathcal{W}$ be as in the Definition. Then:

1. **Sufficiency.** *If $q_1 = q_2$ almost everywhere, then $W_w(q_1) = W_w(q_2)$ for all $w \in \mathcal{W}$. Hence, all welfare comparisons in $\mathcal{W}$ depend only on $q(u, v)$.*
2. **Identification completeness.** *If $q_1 \neq q_2$ on a subset of $(0, 1)^2$ with positive measure, there exists $w \in \mathcal{W}$ such that $W_w(q_1) \neq W_w(q_2)$; equivalently,*

$$\forall\, w \in \mathcal{W}, \quad \int w(u, v)\left[q_1(u, v) - q_2(u, v)\right] du\, dv = 0 \quad \Longleftrightarrow \quad q_1 = q_2 \text{ a.e.}$$

3. **Uniqueness.** *Any other statistic $T(\cdot)$ that is sufficient for all welfare criteria in $\mathcal{W}$ must coincide almost everywhere with a measurable transformation of $q(u, v)$; that is, there exists a measurable function ϕ such that $T = \phi(q(u, v))$ a.e.*

Utilitarian Welfare Function

Equal weights across all individuals:

$$w(u, v) = 1.$$

- Welfare reduces to the mean outcome: $W = E[Y]$.
- Society is indifferent to inequality.

◂ Back to overview

Rank-Dependent Welfare Function

Weights depend only on unconditional ranks:

$$w(u, v) = \tilde{w}(F_Y(q(u, v))).$$

- Standard welfarist form: $W = \int_0^1 \tilde{w}(\theta)q(\theta)d\theta$.
- Ignores within/between dimensions. Only the overall rank matters.

◂ Back to overview

Two-Dimensional Gini Welfare Function

Weights decay linearly in both within- and between-group ranks:

$$w(u, v) = 2[1 - \omega u - (1 - \omega)v], \quad \omega \in [0, 1].$$

- ω controls the trade-off between within- and between-group inequality.
- $\omega = 1$: welfare reduces to a function of the Gini index in the average group: $W = E[y](1 - I_{Gini})$.

◂ Back to overview

Equality of Opportunity

Weights focus on compensating for differences in circumstances:

$$w(u, v) = w(v), \qquad w'(v) \leq 0.$$

- Society compensates across v (circumstances) but not across u (effort).
- Roemer (1998): all weight on the worst circumstance $w(v) = \lim_{\varepsilon \downarrow 0} \frac{\mathbf{1}\{0 \leq v \leq \varepsilon\}}{\varepsilon}$.

Back to overview

Optimal Treatment Assignment

- Two-dimensional quantile treatment effects can be used to optimally assign groups or individuals to treatment.
- Policymaker decides whom to treat in a given **target** population after observing data from a **sample** population by maximizing a rank-dependent social welfare function (see, e.g., Kitagawa and Tetenov, 2021).

Optimal Treatment Assignment

- Two-dimensional quantile treatment effects can be used to optimally assign groups or individuals to treatment.
- Policymaker decides whom to treat in a given **target** population after observing data from a **sample** population by maximizing a rank-dependent social welfare function (see, e.g., Kitagawa and Tetenov, 2021).
- Point of departure:
  - Kitagawa and Tetenov (2021) assigns treatment based on observable covariates. Baseline outcomes are not always available.
  - Kaji and Cao (2023) considers one-dimensional heterogeneity.

Optimal Treatment Assignment

- Two-dimensional quantile treatment effects can be used to optimally assign groups or individuals to treatment.
- Policymaker decides whom to treat in a given **target** population after observing data from a **sample** population by maximizing a rank-dependent social welfare function (see, e.g., Kitagawa and Tetenov, 2021).
- Point of departure:
  - Kitagawa and Tetenov (2021) assigns treatment based on observable covariates. Baseline outcomes are not always available.
  - Kaji and Cao (2023) considers one-dimensional heterogeneity.
- Goal: select a treatment rule that assigns individuals depending on their ranks (u_{ij}, v_j).
- With the structural model, individual treatment effects are identified.
- Exploit treatment effect heterogeneity within and between groups to allocate the treatment more efficiently.

Optimal Treatment Assignment

- Welfare under treatment rule G depends on the distribution of the outcome y_{ij} under the treatment rule:

$$y_{ij} = 1\{(u_{ij}, v_j) \in G\} y_{ij}(1) + 1\{(u_{ij}, v_j) \notin G\} y_{ij}(0),$$

and the optimal treatment rule solves

$$G^* \in \arg\max_{G \in \mathcal{G}} W(G). \tag{2}$$

- Summing up the welfare weights of each individual in a group provides a unified and welfare-based measure of group rank or priority.

back to policy evaluation

Asymptotics - Intuition

If the first stage parameter vector $\beta_0(u)$ was known, the true parameter vector $\delta_0(\beta_0, \tau)$ of the second stage quantile regression uniquely satisfies:

$$\mathbb{E}[m_{ij}(\delta_0, \beta_0, \tau)] = 0 \tag{3}$$

with $m_{ij}(\delta, \beta, \tau) = x'_{ij}[v - 1(\tilde{x}'_{ij}\beta_j(u) \leq x'_{ij}\delta(\beta(u), v))]$.

Let $M_{mn}(\hat{\delta}, \hat{\beta}, \tau) = \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} m_{ij}(\hat{\delta}, \hat{\beta}, \tau)$.

1. Show that $||M_{mn}(\hat{\delta}, \hat{\beta}, \tau)] - \mathcal{L}(\hat{\delta})|| \leq o_p(m^{-1/2})$, for some linear function $\mathcal{L}(\delta)$.
2. Let $\bar{\delta}$ be the minimizer of $\mathcal{L}(\delta)$ where

$$\sqrt{m}\left(\bar{\delta} - \delta_0\right) = -\Gamma_1(\delta_0, \beta_0)^{-1}\sqrt{m}\left(\frac{1}{m}\sum_{j=1}^{m}\bar{\Gamma}_{2j}(\delta_0, \beta_0)[\hat{\beta}_j - \beta_{j,0}] + M_{mn}(\delta_0, \beta_0)\right)$$

3. Show that $\sqrt{m}\left(\hat{\delta}(\hat{\beta}) - \bar{\delta}\right) = o_p(1)$.

back

Assumptions I

1. **Sampling –** (i) The processes $\{(y_{ij}, x_{ij}) : i = 1, \ldots, n\}$ are i.i.d. across j. (ii) For each j, the observations (y_{ij}, x_{ij}) are i.i.d. across i.
2. **Covariates –** (i) For all $j = 1, \ldots, m$ and all $i = 1, \ldots, n$, $\|x_{ij}\| \leq C$ almost surely. (ii) The eigenvalues of $\mathbb{E}_{i|j}[\tilde{x}_{ij}\tilde{x}'_{ij}]$ and $\mathbb{E}[x_{ij}x'_{ij}]$ are bounded away from zero and infinity uniformly across j.
3. **Conditional distribution I–** The conditional distribution $F_{y_{ij}|x_{1ij},v_j}(y|x,v)$ is twice differentiable w.r.t. y, with the corresponding derivatives $f_{y_{ij}|x_{1ij},v_j}(y|x,v)$ and $f'_{y_{ij}|x_{1ij},v_j}(y|x,v)$. Further, assume that

$$f_y^{max} := \sup_j \sup_{y\in\mathbb{R},\ x\in\mathcal{X}_1} |f_{y_{ij}|x_{1ij},v_j}(y|x,v)| < \infty,$$

and

$$\bar{f}'_y := \sup_j \sup_{y\in\mathbb{R},\ x\in\mathcal{X}_1} |f'_{y_{ij}|x_{1ij},v_j}(y|x,v)| < \infty.$$

where $\mathcal{X}_1$ is the support of x_{1ij}

Assumptions II

4. **Bounded density I** – There exists a constant $f_y^{min} < f_y^{max}$ such that

$$0 < f_{min} \leq \inf_j \inf_{u \in \mathcal{T}} \inf_{x \in \mathcal{X}_1} f_{y_{ij}|x_{1ij}, v_j}(Q(u, y_{ij}|x_{ij}, v_j)|x, v).$$

5. **Group level heterogeneity**– The conditional distribution $F_{Q(u,y_{ij}|x_{ij},v_j)|x_{ij}}(q|x)$ is twice continuously differentiable w.r.t. q, with the corresponding derivatives $f_{Q(u,y_{ij}|x_{ij},v_j)|x_{ij}}(q|x)$ and $f'_{Q(u,y_{ij}|x_{ij},v_j)|x_{ij}}(q|x)$. Further, assume that

$$f_Q^{max} := \sup_{u \in \mathcal{T},\ q \in \mathbb{R},\ x \in \mathcal{X}} |f_{Q(u,y_{ij}|x_{ij},v_j)|x_{ij}}(q|x)| < \infty$$

and

$$\bar{f}'_Q := \sup_{u \in \mathcal{T},\ q \in \mathbb{R},\ x \in \mathcal{X}} |f'_{Q(u,y_{ij}|x_{ij},v_j)|x_{ij}}(q|x)| < \infty.$$

where $\mathcal{X}$ is the support of x_{ij}.

Assumptions III

6. **Bounded density II –** There exists a constant $f_Q^{min} < f_Q^{max}$ such that

$$0 < f_{min} \leq \inf_{u,v \in \mathcal{T} \times \mathcal{T}} \inf_{x \in \mathcal{X}} f_{Q(u, y_{ij}|x_{ij}, v_j)|x_{ij}}(x_{ij}' \delta_0(\tau)|x).$$

7. **Compact parameter space –** For all τ, $\beta_{j,0}(u) \in \text{int}(\mathcal{B}_j)$ and $\delta_0(\beta_0, \tau) \in \text{int}(\mathcal{D})$, where $\mathcal{B}_j$ and $\mathcal{D}$ are compact subsets of $\mathbb{R}^{K_1+1}$ and $\mathbb{R}^K$, respectively.
8. **Coefficients –** For all $u, u' \in \mathcal{T}$ and $j = 1, \ldots, m$, $\|\beta_j(u) - \beta_j(u')\| \leq C|u - u'|$. Further, for all $\tau, \tau' \in \mathcal{T} \times \mathcal{T}$ and $\|\delta(\tau) - \delta(\tau')\| \leq C|u - u'| + \leq C|v - v'|$.
9. **Growth rates –** As $m \to \infty$, we have
    1. $\frac{\log m}{n} \to 0$,
    2. $\frac{\sqrt{m} \log n}{n} \to 0$,

back

Inference

- The asymptotic distribution is degenerate, if there is no group-level heterogeneity. More
- In similar settings (Liao and Yang, 2018; Lu and Su, 2023; Fernández-Val et al., 2022) show that the procedure is uniformly valid in the rate of convergence. While Melly and Pons (2025) shows similar results for clustered covariance matrix estimator.
- It is likely that the inference procedure suggested here is valid adaptively.
- However, it is not possible to use the same proof strategy (linearization used to prove the results holds only under heterogeneity).

Back

Kolmogorov–Smirnov

Consider the $H_0 : \delta_k(\tau) = \bar{\delta}_k, \ \forall \ u, v \in \mathcal{T} \times \mathcal{T}$.
Test statistic:

$$t^{KS} = \sup_{\tau \in \mathcal{T} \times \mathcal{T}} \sqrt{\left(\hat{\delta}_k(\tau) - \bar{\delta}_k\right)' \hat{V}_k(\tau)^{-1} \left(\hat{\delta}_k(\tau) - \bar{\delta}_k\right)},$$

with $\bar{\delta}_k = \int_v \int_u \hat{\delta}(u, v) dudv$ and where $\hat{V}_k(\tau)$ is a bootstrap estimate of the asymptotic variance of $\hat{\delta}_k(\tau)$.

Kolmogorov–Smirnov

- To obtain the critical values, I follow Chernozhukov and Fernández-Val (2005) and use the bootstrap to mimic the test statistic.
- To impose the null, I use the parametric bootstrap based on the estimated quantile regression process.
- For each bootstrap iteration, construct the test statistic:

$$t_b^{KS} = \sup_{\tau \in \mathcal{T} \times \mathcal{T}} \sqrt{\left(\hat{\delta}_k^{*b}(\tau) - \hat{\delta}_k^{*b}\right)' \hat{V}_k(\tau)^{-1} \left(\hat{\delta}_k^{*b}(\tau) - \hat{\delta}_k^{*b}\right)}, \tag{4}$$

where $\hat{\delta}_k^{*b} = \int_v \int_u \hat{\delta}^{*b}(u, v) du dv$.

- The critical values of a test with size α are the $1 - \alpha$ quantile of $\{t_b^{KS} : 1 \leq b \leq B\}$.

back

Functional Confidence Intervals

Following Chernozhukov et al. (2013), it is possible to construct functional confidence intervals that cover the entire function with a pre-specified rate by inverting the acceptance region of the KS statistics

$$t_b^{KS} = \sup_{\tau \in \mathcal{T} \times \mathcal{T}} \sqrt{\left(\hat{\delta}_k^{*b}(\tau) - \hat{\delta}_k(\tau)\right)' \hat{V}_k(\tau)^{-1} \left(\hat{\delta}_k^{*b}(\tau) - \hat{\delta}_k(\tau)\right)}.$$

The $(1-\alpha)$ functional confidence bands for a coefficient $\hat{\delta}_k(\tau)$ can be constructed by

$$\hat{\delta}_k(\tau) \pm \hat{t}^*_{1-\alpha} \cdot \sqrt{\hat{V}_k(\tau)},$$

where $\hat{t}^*_{1-\alpha}$ is the $1-\alpha$ quantile of $\{t_b^{KS} : 1 \leq b \leq B\}$.

back

Simulations

- Data generating process:

$$y_{ij} = 1 + x_{1ij} + \gamma \cdot x_{2j} + \eta_j(1 - 0.1 \cdot x_{1ij} - 0.1 \cdot x_{2j}) + \nu_{ij}(1 + 0.1 \cdot x_{1ij} + 0.1 \cdot x_{2j})$$

  with $x_{1ij} = 1 + h_j + w_{ij}$, where $h_j \sim U[0,1]$ and $w_{ij}, x_{2j}, \eta_j, \nu_{ij}$ are $N(0,1)$.

  Let F be the standard normal cdf.
  - $\beta(u,v) = 1 + 0.1 \cdot F^{-1}(u) - 0.1 \cdot F^{-1}(v)$
  - $\gamma(u,v) = 1 + 0.1 \cdot F^{-1}(u) - 0.1 \cdot F^{-1}(v)$.

- $(m,n) = \{(25,25),(200,25),(25,200),(200,200),(200,400)\}$
- Set of quantiles $\{0.25, 0.5, 0.75\}$
- $2,000$ Monte Carlo simulations.
- 100 bootstrap repetitions.

back

Simulations - Bias and Standard Deviation I

	β			γ		
$u \setminus v$	0.25	0.5	0.75	0.25	0.5	0.75
			(m, n) = (25,25)			
0.25	-0.023	0.004	0.034	-0.030	-0.006	0.018
	(0.119)	(0.110)	(0.117)	(0.243)	(0.222)	(0.239)
0.5	-0.021	-0.001	0.027	-0.029	-0.010	0.014
	(0.114)	(0.106)	(0.111)	(0.240)	(0.219)	(0.235)
0.75	-0.029	-0.005	0.024	-0.031	-0.012	0.014
	(0.114)	(0.112)	(0.119)	(0.246)	(0.222)	(0.236)
			(m, n) = (25,200)			
0.25	-0.010	0.000	0.007	-0.004	0.006	0.019
	(0.071)	(0.067)	(0.072)	(0.237)	(0.215)	(0.232)
0.5	-0.010	-0.002	0.005	-0.004	0.004	0.018
	(0.067)	(0.066)	(0.070)	(0.237)	(0.215)	(0.235)
0.75	-0.010	-0.004	0.006	-0.007	0.004	0.017
	(0.070)	(0.069)	(0.072)	(0.237)	(0.217)	(0.238)

Simulations - Bias and Standard Deviation II

	β			γ		
$u \setminus v$	0.25	0.5	0.75	0.25	0.5	0.75
			(m, n) = (200,25)			
0.25	-0.023	0.004	0.030	-0.018	0.003	0.022
	(0.043)	(0.040)	(0.042)	(0.082)	(0.072)	(0.078)
0.5	-0.024	-0.001	0.023	-0.018	0.001	0.019
	(0.041)	(0.037)	(0.040)	(0.078)	(0.072)	(0.077)
0.75	-0.032	-0.007	0.020	-0.020	-0.002	0.018
	(0.043)	(0.038)	(0.042)	(0.079)	(0.072)	(0.078)
			(m, n) = (200,200)			
0.25	-0.005	0.001	0.006	-0.004	0.001	0.003
	(0.028)	(0.026)	(0.028)	(0.076)	(0.073)	(0.079)
0.5	-0.005	0.000	0.006	-0.004	0.000	0.003
	(0.028)	(0.025)	(0.028)	(0.076)	(0.073)	(0.079)
0.75	-0.006	0.000	0.006	-0.005	0.001	0.002
	(0.028)	(0.026)	(0.028)	(0.077)	(0.073)	(0.079)

Simulations - Bias and Standard Deviation III

	β			γ		
$u \setminus v$	0.25	0.5	0.75	0.25	0.5	0.75
(m, n) = (200,200)						
0.25	-0.005	0.001	0.006	-0.004	0.001	0.003
	(0.028)	(0.026)	(0.028)	(0.076)	(0.073)	(0.079)
0.5	-0.005	0.000	0.006	-0.004	0.000	0.003
	(0.028)	(0.025)	(0.028)	(0.076)	(0.073)	(0.079)
0.75	-0.006	0.000	0.006	-0.005	0.001	0.002
	(0.028)	(0.026)	(0.028)	(0.077)	(0.073)	(0.079)
(m, n) = (200,400)						
0.25	-0.003	0.000	0.003	-0.004	-0.003	0.002
	(0.026)	(0.023)	(0.026)	(0.077)	(0.073)	(0.079)
0.5	-0.003	0.000	0.003	-0.004	-0.003	0.002
	(0.025)	(0.023)	(0.025)	(0.077)	(0.073)	(0.079)
0.75	-0.004	-0.001	0.003	-0.005	-0.004	0.002
	(0.026)	(0.024)	(0.026)	(0.077)	(0.073)	(0.079)

Simulations - Standard Errors I

Table: Bootstrap Standard Errors relative to Standard Deviation

	β			γ		
$u \setminus v$	0.25	0.5	0.75	0.25	0.5	0.75
(m, n) = (25,25)						
0.25	1.180	1.146	1.200	1.148	1.098	1.261
0.5	1.191	1.133	1.242	1.190	1.115	1.327
0.75	1.213	1.119	1.230	1.167	1.107	1.357
(m, n) = (25,200)						
0.25	1.321	1.231	1.401	1.275	1.138	1.649
0.5	1.381	1.229	1.457	1.332	1.138	1.720
0.75	1.358	1.199	1.443	1.352	1.126	1.724

back

Simulations - Standard Errors II

Table: Bootstrap Standard Errors relative to Standard Deviation

	β			γ		
$u \setminus v$	0.25	0.5	0.75	0.25	0.5	0.75
(m, n) = (200,25)						
0.25	1.028	1.031	1.048	1.002	1.056	1.043
0.5	1.017	1.052	1.069	1.028	1.052	1.063
0.75	1.025	1.063	1.050	1.027	1.053	1.052
(m, n) = (200,200)						
0.25	1.089	1.081	1.080	1.056	0.995	1.021
0.5	1.064	1.081	1.081	1.052	1.000	1.014
0.75	1.075	1.082	1.095	1.036	1.004	1.018
(m, n) = (200,400)						
0.25	1.081	1.111	1.078	1.044	1.003	1.011
0.5	1.089	1.092	1.088	1.039	1.004	1.009
0.75	1.092	1.092	1.082	1.037	1.005	1.008

Simulations - Coverage Probability I

Table: Coverage Probability of Bootstrap 95% Confidence Interval

	β			γ		
$u \setminus v$	0.25	0.5	0.75	0.25	0.5	0.75
			(m, n) = (25,25)			
0.25	0.970	0.973	0.969	0.948	0.954	0.953
0.5	0.972	0.973	0.970	0.949	0.951	0.948
0.75	0.971	0.968	0.972	0.949	0.958	0.946
			(m, n) = (25,200)			
0.25	0.985	0.987	0.985	0.957	0.959	0.965
0.5	0.986	0.985	0.981	0.956	0.956	0.964
0.75	0.988	0.988	0.987	0.955	0.953	0.954

back

Simulations - Coverage Probability II

Table: Coverage Probability of Bootstrap 95% Confidence Interval

	β			γ		
$u \setminus v$	0.25	0.5	0.75	0.25	0.5	0.75
(m, n) = (200,25)						
0.25	0.916	0.948	0.899	0.929	0.943	0.928
0.5	0.905	0.955	0.925	0.936	0.954	0.932
0.75	0.878	0.952	0.931	0.941	0.959	0.943
(m, n) = (200,200)						
0.25	0.964	0.965	0.954	0.948	0.938	0.940
0.5	0.955	0.961	0.956	0.945	0.940	0.944
0.75	0.961	0.963	0.961	0.947	0.942	0.947
(m, n) = (200,400)						
0.25	0.957	0.958	0.961	0.948	0.936	0.939
0.5	0.963	0.961	0.961	0.946	0.938	0.934
0.75	0.959	0.963	0.959	0.944	0.940	0.931

Simulations - DGP for the KS and CvM Tests

Let

$$y_{ij} = 1 + x_{1ij} + x_{2j} + \eta_j(1 - \psi(x_{1ij} + x_{2j})) + \nu_{ij}(1 + \phi(x_{1ij} + x_{2j})),$$

with $x_{1ij} = 1 + h_j + w_{ij}$, where $h_j \sim U[0,1]$ and $w_{ij}, x_{2j}, \eta_j, \nu_{ij}$ are $N(0,1)$.

- ϕ regulate effect heterogeneity over u
- ψ regulate effect heterogeneity over v.

Test the null hypotheses that $\beta(\tau) = \bar{\beta}$ and that $\gamma(\tau) = \bar{\gamma}$.

- Simulations on the set of quantiles $0.1, 0.2, \ldots, 0.9$.
- Impose the null using the parametric bootstrap based on the estimated quantile regression process.
- $(m, n) = \{(25, 25), (200, 25), (25, 200), (200, 200)\}$
- 1,000 Monte Carlo simulations.
- 100 bootstrap repetition.

Simulations - Rejection Probability of the KS and CvM Tests

Table: Rejection Probability of the Kolmogorov-Smirnov Test

(ϕ, ψ)	(0, 0)	(0, 0.1)	(0.1, 0)	(0.1, 0.1)	(0.2, 0.2)
$H_0 : \gamma(\tau) = \bar{\gamma}$					
(m, n) = (25,25)	0.007	0.005	0.007	0.009	0.034
(m, n) = (25,200)	0.015	0.013	0.020	0.032	0.173
(m, n) = (200,25)	0.026	0.209	0.251	0.469	0.996
(m, n) = (200,200)	0.046	0.307	0.397	0.826	1.000
$H_0 : \beta(\tau) = \bar{\beta}$					
(m, n) = (25,25)	0.026	0.108	0.101	0.156	0.537
(m, n) = (25,200)	0.056	0.536	0.548	0.885	1.000
(m, n) = (200,25)	0.026	0.767	0.822	0.970	1.000
(m, n) = (200,200)	0.057	1.000	1.000	1.000	1.000

Simulations - Rejection Probability of the KS and CvM Tests

Table: Rejection Probability of the Cramér-von Mises Test

(ϕ, ψ)	(0, 0)	(0, 0.1)	(0.1, 0)	(0.1, 0.1)	(0.2, 0.2)
$H_0 : \gamma(\tau) = \bar{\gamma}$					
(m, n) = (25,25)	0.014	0.026	0.022	0.027	0.165
(m, n) = (25,200)	0.023	0.030	0.035	0.047	0.381
(m, n) = (200,25)	0.044	0.381	0.414	0.789	1.000
(m, n) = (200,200)	0.061	0.446	0.430	0.895	1.000
$H_0 : \beta(\tau) = \bar{\beta}$					
(m, n) = (25,25)	0.038	0.223	0.231	0.373	0.921
(m, n) = (25,200)	0.068	0.728	0.844	0.988	1.000
(m, n) = (200,25)	0.048	0.937	0.995	1.000	1.000
(m, n) = (200,200)	0.056	1.000	1.000	1.000	1.000

Simulations - Computing Time

2000 simulations.
100 bootstrap repetitions.
Set of quantile $\{0.25, 0.5, 0.75\}$.
AMD Ryzen Threadripper 3960X 24-Core Processor

(m, n)	**Computing Time**
(25, 25)	18.70 sec
(25, 200)	32.70 sec
(200, 25)	1.30 min
(200, 200)	10.01 min
(200, 400)	32.16 min

back to FAQ   Empirical Application Computing Time   Back to Conclusion

Questions

- Convergence Rate ▸ more
- Growth Condition ▸ more
- Degenerate Distribution ▸ more
- Smoothed Quantile Regression and Bias Correction ▸ more
- Link to Melly and Pons (2025) ▸ more
- Computing Time Empirical Application Simulations
- Endogenous treatment and instrumental variables more
- Quantile Crossing more
- Rank Invariance more

Quantile Crossing

- Ensuring the monotonicity of the estimated two-level quantile functions across both dimensions might require a rearrangement operation, as suggested in Chernozhukov et al. (2009, 2010).
- Due to the nested structure of the problem, rearrangement along the u dimension should be performed after the first stage.
- Monotonicity of the first stage in all groups guarantees that the second stage quantile regression remains monotonic along the u dimension.
- Rearrangement along the v dimension can be implemented subsequent to the second stage.

Back to FAQ

Endogenous Treatment and Instrumental Variables

- The model in the paper assumes that the variation of x_{ij} is exogenous.
- If this is not the case, the estimator suggested here can be easily extended to accommodate instrumental variables.
- Depending on which variables are assumed to be endogenous, either the second stage or both stages could be estimated using an instrumental variable quantile regression estimator (e.g., Chernozhukov and Hansen, 2005).

Back to FAQ

Relation to Melly and Pons (2025)

- Propose a minimum distance approach to quantile panel data models where the unit effects may be correlated with the covariates.
- The model and estimator are flexible and apply to:
  - **Classical panel data**, tracking units over time,
  - **Grouped data**, where individual-level data is available, but often the treatment vars are at the group level.
- We suggest a general framework for quantile panel data models.
- New random effects quantile estimator, new Hausman test, new Hausman-Taylor quantile estimator, new grouped (IV) quantile regression estimator.
- The asymptotic distribution of our estimator is non-standard, as the rate of convergence of a coefficient depends on the presence of group-level heterogeneity and the variation used to identify that coefficient. $\Longrightarrow$ We derive adaptive asymptotic results and inference procedure.

Relation to Melly and Pons (2025)

This paper focuses on simultaneously estimating the effect on the distribution of the outcome within and between groups. In Melly and Pons (2025) the heterogeneity arises from the individual rank variable u_{ij} and the focus is on the within distribution.

Starting from the two-dimensional quantile function and assuming that $(x_{ij}) \perp\!\!\!\perp v_j$, we can obtain the model in Melly and Pons (2025) by integrating over v_j:

$$
\begin{aligned}
\mathbb{E}\left[Q\left(u, y_{ij} | x_{ij}, v_i\right) | x_{ij}\right] =& x_{1ij}' \int \beta(u, v) dF_V(v) + x_{2j}' \int \gamma(u, v) dF_V(v) \\
& + \int \alpha(u, v) dF_V(v) \\
=& x_{ij}' \bar{\beta}(u) + \bar{\alpha}(u).
\end{aligned}
$$

They identify the average effects over groups at the u quantile of the within distribution.

Back to FAQ Back to Literature

Computing Time - Empirical Application

- 17 quantiles: $\{0.1, 0.15, 0.2, \ldots, 0.9\}$
- $m = 116$
- $n \times m = 2922$ (average group size $= 25$).
- Bootstrap standard errors ($r = 1,000$).

Computing time: 2021 MacBook Pro with Apple M1 Pro Chip (8 cores): 2.21 minutes.

▸ Back to FAQ ▸ Simulations Running Time ▸ Estimator ▸ Application Results ▸ Conclusion

Degenerate Distribution

- In similar settings, Galvao et al. (2020), Melly and Pons (2025) show that without group-level heterogeneity, the first stage error dominates, and the estimator convergences at the $1/\sqrt{mn}$ rate (requirement: $\frac{m(\log n)^2}{n} \to 0$).
- Under the stronger growth condition, it is possible to show that $\sqrt{mn}\frac{1}{m}\sum_{j=1}^{m}\bar{\Gamma}_{2,j}(\delta_0,\beta_0,\tau)\left(\hat{\beta}_j(u)-\beta_{j,0}(u)\right)\xrightarrow{d} N(0,\Omega_1(\tau))$.
- Intuitively, without heterogeneity between groups, the estimated group-level conditional quantile functions are identical up to the first stage error, and the estimator should converge at the faster $1/\sqrt{mn}$ rate.
- In this case, it is not possible to use that same proof strategy. The linearization used to derive the asymptotic results relies on the presence of group-level heterogeneity.
- Simulations without group-level heterogeneity show that this is also the case with the non-linear second-step estimator.

▸ Back to Asymptotic ▸ FAQ ▸ More on Inference

Convergence Rate

- The entire coefficient vector converges at the $1/\sqrt{m}$ rate despite mn observations being used for the estimation.
- It is a consequence of modeling heterogeneities between groups:
  - Imposing equality of $\beta(u, v)$ over groups would allow to estimate this coefficient at the $1/\sqrt{mn}$ rate.
  - Since $\beta(u, v)$ is allowed to vary over groups through the dependency on v, between groups variation is necessary for identification.
- Similarly, in the least squares case, it is always possible to estimate the coefficient on x_{ij} at the $1/\sqrt{mn}$ rate by implementing a fixed effects estimator.
- However, this estimator only exploits the within-group variation and cannot identify heterogeneities between groups.
- Ultimately, the between variation, which slows down the convergence rate, has to be used to identify between-group heterogeneity.

▸ Back to FAQ

Growth Conditon

- Nonlinear panel data literature has shown that $m/n \to 0$ is a sufficient condition to obtain asymptotic normality of nonlinear panel data FE estimators.
- Galvao et al. (2020) show that unbiased asymptotic normality of panel data FE QR estimator hold under $m(log(n))^2/n \to 0$.
  - Previous condition in the literature: $m^2 log(m)(log(n))^2/n \to 0$.
- These estimator converge at the $\sqrt{mn}$ rate.
- My estimator converges at the $\sqrt{m}$ rate. Hence, I only need $m \ log(n)/n \to 0$.

Back to FAQ

Smoothed Panel Data Quantile Regression and Bias Correction

- Galvao and Kato (2016) show that the smoothed FE estimator $\sqrt{mn}(\hat{\beta} - \beta_0) \xrightarrow{d} N(bias, V)$ if $m/n \to c$.
- Bias corrected estimator is centered at zero under the same growth condition.
- Smoothed QR estimator requires stronger smoothness conditions on the distribution of the outcome variable and the choice of a bandwidth that is arbitrary.
- This approach is not applicable in this setting as it assumes homogeneity of the coefficients over groups.
- Franguridi, Gafarov, and Wüthrich (2024) derive an explicit formula for the bias of the leading term of the expansion. However, implementation remains a major challenge.

Back to FAQ

Rank Invariance

- With rank invariance, treatment effects of individuals at given points of the distribution are identified.
- The model in this paper continues to identify well-defined parameters even if rank invariance is not satisfied.
- Testing procedure for rank similarity (or rank invariance) have been proposed in the literature (Dong and Shen, 2018; Frandsen and Lefgren, 2018; Kim and Park, 2022).
  - Requirements: Binary treatment, multi-valued instrument or multiple IVs (Frandsen and Lefgren, 2018).

Back to FAQ