Quantile on Quantiles

Martina Pons

University of Bern

martina.pons@unibe.ch

August 26th, 2024

Motivation

- This paper suggests a method to study **distributional effects along multiple dimensions** simultaneously (within and between groups).
- Let groups be geographical regions (counties or commuting zones)
- Large body of literature focusing on inequalities / distributional effects:
  - **Within groups:** Trade shocks (Autor et al., 2021), minimum wages (Autor et al., 2016; Engbom and Moser, 2022), place-based policies (Lang et al., 2023; Albanese et al., 2023).
  - **Between groups:** Place-based policies (Becker et al., 2010; Busso et al., 2013)

Motivation

- This paper suggests a method to study **distributional effects along multiple dimensions** simultaneously (within and between groups).
- Let groups be geographical regions (counties or commuting zones)
- Large body of literature focusing on inequalities / distributional effects:
  - **Within groups:** Trade shocks (Autor et al., 2021), minimum wages (Autor et al., 2016; Engbom and Moser, 2022), place-based policies (Lang et al., 2023; Albanese et al., 2023).
  - **Between groups:** Place-based policies (Becker et al., 2010; Busso et al., 2013)
- Inequalities (along both dimensions) have welfare effects:
  - **Within groups:** Individuals compare themselves with their peers, neighbors, and co-workers (Galí, 1994; Luttmer, 2005; Card et al., 2012). Within-group inequality is associated with a higher murder rate (Glaeser et al., 2009) and lower future outcomes for children (Chetty and Hendren, 2018b).
  - **Between groups:** Neighborhood quality matter for future outcomes (Chetty et al., 2016; Chetty and Hendren, 2018a,b).

Motivation

- Modelling welfare as a function of the unconditional distribution of the outcome ignores the role of inequalities within smaller regions and between these regions.
  - If we keep the unconditional income distribution constant, it is possible to reduce within-region inequality by moving people across space into a more segregated spatial allocation.
- These two dimensions are **interdependent** and there are **tradeoffs**.
  $\rightarrow$ **Hence, we should model them together.**

Motivation

This paper...

- suggests a method to **simultaneously** study **distributional effects within and between groups**.
- Introduces a quantile model with **two quantile indices**: one for the heterogeneity within groups and one for the heterogeneity between groups.
- proposes a **two-step quantile regression estimator** with within-group regression in the first stage and between-group regression in the second stage.

Related Literature

- Within distribution (Galvao and Wang, 2015; Chetverikov, Larsen, and Palmer, 2016; Melly and Pons, 2023).
  - Model also the between distribution.
- Multidimensional heterogeneity (Arellano and Bonhomme, 2016; Frumento, Bottai, and Fernández-Val, 2021; Liu and Yang, 2021; 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).

Example: Business Training

- Consider an experiment designed to improve small business outcomes (e.g., sales, profit, income).
- Quantile regression of sale on the treatment dummy identifies the treatment effect at different points of the sales distribution (high-performing vs. low-performing firms).
- It might be different to be a median business in a highly-performing market compared to a lower-performing one.

Example: Business Training

- Consider an experiment designed to improve small business outcomes (e.g., sales, profit, income).
- Quantile regression of sale on the treatment dummy identifies the treatment effect at different points of the sales distribution (high-performing vs. low-performing firms).
- It might be different to be a median business in a highly-performing market compared to a lower-performing one.
  - A poor-performing market might have poor locations and low consumer traffic.

Example: Business Training

- Consider an experiment designed to improve small business outcomes (e.g., sales, profit, income).
- Quantile regression of sale on the treatment dummy identifies the treatment effect at different points of the sales distribution (high-performing vs. low-performing firms).
- It might be different to be a median business in a highly-performing market compared to a lower-performing one.
  - A poor-performing market might have poor locations and low consumer traffic.
- Identify the quantile treatment effect over the distribution of income within the market and over the distribution of markets.
  - With rank invariance over treatment states, the method identifies the effect for a median (or other percentile) business (in his market) over the distribution of markets.

Model (Simplified)

Let $j = 1, \ldots, m$ be the groups and $i = 1, \ldots, n$ be the individuals. Consider the following structural function for the outcome variable

$$y_{ij} = q(x_{1ij}, x_{2j}, v_j, u_{ij}), \tag{1}$$

where $q(\cdot)$ is **strictly increasing** in the third and fourth arguments.

- x_{1ij}: individual-level variables
- x_{2j}: group-level variables
- u_{ij}: ranks individuals within a group
- v_j: ranks the groups

Model (Simplified)

Let $j = 1, \dots, m$ be the groups and $i = 1, \dots, n$ be the individuals. Consider the following structural function for the outcome variable

$$y_{ij} = q(x_{1ij}, x_{2j}, v_j, u_{ij}), \tag{2}$$

where $q(\cdot)$ is strictly increasing in the third and fourth arguments.

Normalize

$$u_{ij} | x_{1ij}, x_{2j}, v_j \sim U(0,1)$$
$$v_j | x_{1ij}, x_{2j} \sim U(0,1).$$

Conditional on x_{1ij}, x_{2j}, v_j, $q(x_{1ij}, x_{2j}, v_j, u_{ij})$ is strictly monotonic with respect to u_{ij} so that

$$Q(\tau_1, y_{ij} | x_{1ij}, x_{2j}, v_j) = q(x_{1ij}, x_{2j}, v_j, \tau_1) \tag{3}$$

is the τ_1-conditional quantile function of the outcome y_{ij}

Model (Simplified)

Assume there are no covariates.

Take two groups $j = \{h, l\}$ with $v_h > v_l$. Strict monotonicity of $q(v_j, \tau_1)$ with respect to v_j implies:

$$q(v_h, \tau_1) > q(v_l, \tau_1), \quad \text{for all } \tau_1 \in (0, 1).$$

Hence, if a group has a higher first decile, it must also have a higher ninth decile.

A model with a **univariate** v_j restricts the evolution of the group ranks at different values of τ_1 (**constant** ranks over τ_1).

Satisfied if all groups share the same outcome distribution up to a **location parameter**.

Model (Simplified)

Assume there are no covariates.

Take two groups $j = \{h, l\}$ with $v_h > v_l$. Strict monotonicity of $q(v_j, \tau_1)$ with respect to v_j implies:

$$q(v_h, \tau_1) > q(v_l, \tau_1), \quad \text{for all } \tau_1 \in (0, 1).$$

Hence, if a group has a higher first decile, it must also have a higher ninth decile.

A model with a **univariate** v_j restricts the evolution of the group ranks at different values of τ_1 (**constant** ranks over τ_1).

Satisfied if all groups share the same outcome distribution up to a **location parameter**.

- Allow for different **mean** and **variance** across groups → **bivariate** v_j

Model (Simplified)

Assume there are no covariates.

Take two groups $j = \{h, l\}$ with $v_h > v_l$. Strict monotonicity of $q(v_j, \tau_1)$ with respect to v_j implies:

$$q(v_h, \tau_1) > q(v_l, \tau_1), \quad \text{for all } \tau_1 \in (0, 1).$$

Hence, if a group has a higher first decile, it must also have a higher ninth decile.

A model with a **univariate** v_j restricts the evolution of the group ranks at different values of τ_1 (**constant** ranks over τ_1).

Satisfied if all groups share the same outcome distribution up to a **location parameter**.

- Allow for different **mean** and **variance** across groups → **bivariate** v_j
- Allow for different outcome **distribution** over groups → **infinitely dimensional** v_j.

A more general model

Consider the linear specification

$$y_{ij} = x_{1ij}'\beta(u_{ij}, v_j) + x_{2j}'\gamma(u_{ij}, v_j) + \alpha(u_{ij}, v_j),$$

where $\alpha(u_{ij}, v_j)$ is the intercept.

Maintain assumptions on u_{ij}, but allow v_j to be **infinitely dimensional**.

- τ_1-CQF of the outcome y_{ij} conditional on x_{1ij}, x_{2j}, and v_j:

$$Q(\tau_1, y_{ij}|x_{1ij}, x_{2j}, v_j) = x_{1ij}'\beta(\tau_1, v_j) + x_{2j}'\gamma(\tau_1, v_j) + \alpha(\tau_1, v_j).$$

A more general model

Consider the linear specification

$$y_{ij} = x_{1ij}'\beta(u_{ij}, v_j) + x_{2j}'\gamma(u_{ij}, v_j) + \alpha(u_{ij}, v_j),$$

where $\alpha(u_{ij}, v_j)$ is the intercept.

Maintain assumptions on u_{ij}, but allow v_j to be **infinitely dimensional**.

- τ_1-CQF of the outcome y_{ij} conditional on x_{1ij}, x_{2j}, and v_j:

$$Q(\tau_1, y_{ij}|x_{1ij}, x_{2j}, v_j) = x_{1ij}'\beta(\tau_1, v_j) + x_{2j}'\gamma(\tau_1, v_j) + \alpha(\tau_1, v_j).$$

- Multidimensionality of v_j requires restricting the relationship between the τ_1-CQF and $v_j = (v_j^{(1)}, v_j^{(2)}, \dots)$.

A more general model

Consider the linear specification

$$y_{ij} = x_{1ij}'\beta(u_{ij}, v_j) + x_{2j}'\gamma(u_{ij}, v_j) + \alpha(u_{ij}, v_j),$$

where $\alpha(u_{ij}, v_j)$ is the intercept.

Maintain assumptions on u_{ij}, but allow v_j to be **infinitely dimensional**.

- τ_1-CQF of the outcome y_{ij} conditional on x_{1ij}, x_{2j}, and v_j:

$$Q(\tau_1, y_{ij}|x_{1ij}, x_{2j}, v_j) = x_{1ij}'\beta(\tau_1, v_j) + x_{2j}'\gamma(\tau_1, v_j) + \alpha(\tau_1, v_j).$$

- Multidimensionality of v_j requires restricting the relationship between the τ_1-CQF and $v_j = (v_j^{(1)}, v_j^{(2)}, \dots)$.
  $\Longrightarrow$ For each τ_1, there is a **scalar-valued function** $v_j(\tau_1)$ such that $q(x_{1ij}, x_{2j}, v_j, \tau_1) = q(x_{1ij}, x_{2j}, v_j(\tau_1), \tau_1)$.

A more general model

Consider the linear specification

$$y_{ij} = x_{1ij}'\beta(u_{ij}, v_j) + x_{2j}'\gamma(u_{ij}, v_j) + \alpha(u_{ij}, v_j),$$

where $\alpha(u_{ij}, v_j)$ is the intercept.
Maintain assumptions on u_{ij}, but allow v_j to be **infinitely dimensional**.

- τ_1-CQF of the outcome y_{ij} conditional on x_{1ij}, x_{2j}, and v_j:

$$Q(\tau_1, y_{ij}|x_{1ij}, x_{2j}, v_j) = x_{1ij}'\beta(\tau_1, v_j) + x_{2j}'\gamma(\tau_1, v_j) + \alpha(\tau_1, v_j).$$

- Multidimensionality of v_j requires restricting the relationship between the τ_1-CQF and $v_j = (v_j^{(1)}, v_j^{(2)}, \dots)$.
  $\implies$ For each τ_1, there is a **scalar-valued function** $v_j(\tau_1)$ such that $q(x_{1ij}, x_{2j}, v_j, \tau_1) = q(x_{1ij}, x_{2j}, v_j(\tau_1), \tau_1)$.

- Imposing monotonicity w.r.t. $v_j(\tau_1)$, and with proper normalization, we obtain the τ_2–CQF of $Q(\tau_1, y_{ij}|x_{1ij}, x_{2j}, v_j)$

$$Q(\tau_2, Q(\tau_1, y_{ij}|x_{1ij}, x_{2j}, v_j)|x_{1ij}, x_{2j}) = x_{1ij}'\beta(\tau_1, \tau_2) + x_{2j}'\gamma(\tau_1, \tau_2) + \alpha(\tau_1, \tau_2).$$

Interpretation of the coefficients

- $\beta(\tau_1, \tau_2)$ tells how the (τ_1, τ_2)-conditional quantile function responds to a change in x_{1ij} by one unit.
- $\beta(0.5, \tau_2)$ gives the effect of x_{1ij} 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.
- With rank invariance over treatment states, the coefficients can be interpreted as individual effects
  - $\beta(\tau_1, \tau_2)$ gives the effects for individuals at the τ_1 percentile of their groups, belonging to a group at the τ_2 percentile, where this second distribution is viewed from their perspective.
- Interpretation of $\gamma(\tau_1, \tau_2)$ follows the same argument.

Estimator

1. **First stage**: group-by-group quantile regression of the outcome on the variables that vary within groups for quantiles $\tau_1 \in \mathcal{T}$. For each group j and quantile $\tau_1 \in \mathcal{T}$:

$$\hat{\beta}_j(\tau_1) \equiv \left(\hat{\beta}_{1,j}(\tau_1), \hat{\beta}_{2,j}(\tau_1)'\right)' = \underset{(b_1,b_2)\in\mathbb{R}^{dim(x_1)+1}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \rho_{\tau_1}(y_{ij} - b_1 - x'_{1ij}b_2),$$

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

Estimator

1. **First stage**: group-by-group quantile regression of the outcome on the variables that vary within groups for quantiles $\tau_1 \in \mathcal{T}$. For each group j and quantile $\tau_1 \in \mathcal{T}$:

$$\hat{\beta}_j(\tau_1) \equiv \left(\hat{\beta}_{1,j}(\tau_1), \hat{\beta}_{2,j}(\tau_1)'\right)' = \underset{(b_1,b_2)\in\mathbb{R}^{dim(x_1)+1}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \rho_{\tau_1}(y_{ij} - b_1 - x_{1ij}' b_2),$$

where $\rho_\tau(x) = (\tau - 1\{x < 0\})x$ 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 the variables that vary within groups for quantiles $\tau_1 \in \mathcal{T}$. For each group j and quantile $\tau_1 \in \mathcal{T}$:

$$\hat{\beta}_j(\tau_1) \equiv \left(\hat{\beta}_{1,j}(\tau_1), \hat{\beta}_{2,j}(\tau_1)'\right)' = \underset{(b_1,b_2)\in\mathbb{R}^{dim(x_1)+1}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} \rho_{\tau_1}(y_{ij} - b_1 - x_{1ij}'b_2),$$

where $\rho_\tau(x) = (\tau - 1\{x < 0\})x$ 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 $\tau_1 \in \mathcal{T}$ regress the first-stage fitted values on all variables using quantile regression for each quantile $\tau_2 \in \mathcal{T}$:

$$\hat{\delta}(\hat{\beta}(\tau_1), \tau_2) = \underset{(a,b,g)\in\mathbb{R}^{dim(x)+1}}{\arg\min} \frac{1}{mn}\sum_{j=1}^{m}\sum_{i=1}^{n} \rho_{\tau_2}(\hat{y}_{ij}(\tau_1) - x_{2j}'g - x_{1ij}'b - a),$$

where $\delta = (\alpha, \beta', \gamma')'$ and $\hat{y}_{ij}(\tau_1) = \hat{\beta}_{1,j}(\tau_1) + x_{1ij}'\hat{\beta}_{2,j}(\tau_1)$.

Asymptotics

Asymptotic framework where n and $m \to \infty$.

Challenges:

- non-smooth quantile regression objective function with a 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); Volgushev, Chao, and Cheng (2019); Galvao, Gu, and Volgushev (2020).

Asymptotic Distribution

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

$$= -\Gamma_1(\delta_0,\beta_0,\tau)^{-1}\sqrt{m}\left(\frac{1}{m}\sum_{j=1}^{m}\bar{\Gamma}_{2j}(\tau,\delta_0,\beta_0)[\hat{\beta}_j(\tau_1)-\beta_{j,0}(\tau_1)] + M_{mn}(\delta_0,\beta_0,\tau)\right)$$

$$+ o_p(1)$$

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:

$$\frac{1}{m}\sum_{j=1}^{m}\bar{\Gamma}_{2j}(\tau,\delta_0,\beta_0)[\hat{\beta}_j(\tau_1)-\beta_{j,0}(\tau_1)] = o_p(m^{-1/2})$$

Second stage noise:

$$\sqrt{m}\left(M_{mn}(\delta_0,\beta_0,\tau)\right) \xrightarrow{d} N\left(0,\Omega_2(\tau)\right),$$

where $\Omega_2(\tau) = \mathbb{E}\left[[\tau_2 - 1(\tilde{x}_{ij}'\beta_{j,0}(\tau_1) \leq x_{ij}'\delta_0(\beta_0,\tau)]^2 x_{ij}x_{ij}'\right]$.

Asymptotic Distribution

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

First stage error:

$$\frac{1}{m}\sum_{j=1}^{m}\bar{\Gamma}_{2j}(\tau,\delta_0,\beta_0)[\hat{\beta}_j(\tau_1)-\beta_{j,0}(\tau_1)] = o_p(m^{-1/2})$$

Second stage noise:

$$\sqrt{m}\,(M_{mn}(\delta_0,\beta_0,\tau)) \xrightarrow{d} N\left(0,\Omega_2(\tau)\right),$$

where $\Omega_2(\tau) = \mathbb{E}\left[[\tau_2 - 1(\tilde{x}_{ij}'\beta_{j,0}(\tau_1) \leq x_{ij}'\delta_0(\beta_0,\tau)]^2 x_{ij}x_{ij}'\right]$. Hence,

$$\sqrt{m}\left(\hat{\delta}(\hat{\beta},\tau)-\delta_0(\beta_0,\tau)\right) \xrightarrow{d} N(0,\Gamma_1^{-1}\Omega_2(\tau)\Gamma_1'^{-1})$$

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

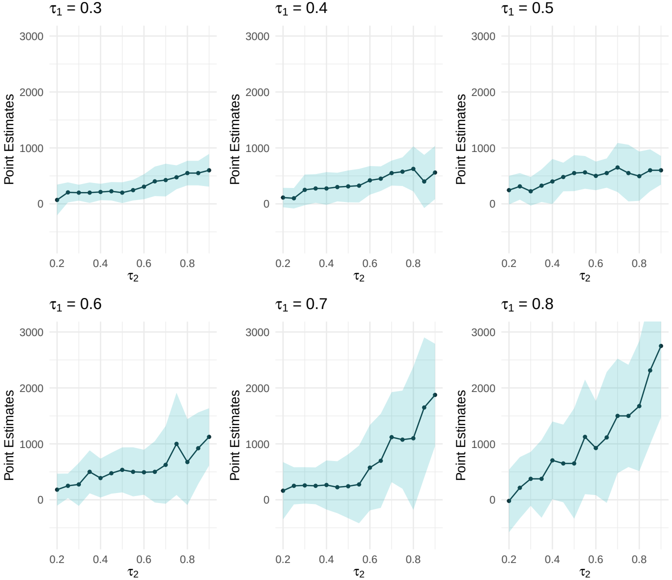
Empirical Application

- Build on McKenzie and Puerto (2021)
- Estimate the impact of business training on the outcomes of female-owned businesses.
- Sample: 3,537 female-owned businesses operating in 157 different rural markets in Kenya.
- Two-stage randomization:
  1. market-level randomization (93 markets are assigned to the treatment markets, and the remaining 64 are control markets)
  2. individual-level randomization. Firms in the treatment markets are randomly assigned to training.
- Estimate distributional effect both within and between markets.
- Outcomes: Sales, Profits, Income from Work.

▸ Specification

Results - Income from Work

Rank Correlation



Conclusion

- This paper suggests a method to simultaneously study distributional effects within and between groups.
  - Allows us to consider tradeoffs 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
- Example of income inequality in Switzerland
- Application to policy evaluation and optimal treatment assignment.
- The road ahead:
  - Uniform results

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.

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

Autor, D. H., D. Dorn, and G. H. Hanson (2021): "When Work Disappears: Manufacturing Decline and the Falling Marriage Market Value of Young Men," *American Economic Review: Insights*, 1, 161–178.

Autor, D. H., A. Manning, and C. L. Smith (2016): "The contribution of the minimum wage to US wage inequality over three decades: A reassessment," *American Economic Journal: Applied Economics*, 8, 58–99.

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," .

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.

Card, D., A. Mas, E. Moretti, and E. Saez (2012): "Inequality at work: The effect of peer salaries on job satisfaction," *American Economic Review*, 102, 2981–3003.

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.

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

References II

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

Chetty, R., N. Hendren, and L. F. Katz (2016): "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment," *American Economic Review*, 106, 855–902.

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.

Dhaene, G. and K. Jochmans (2015): "Split-Panel Jackknife Estimation of Fixed-Effect Models," *Review of Economic Studies*, 82, 991–1030.

Engbom, N. and C. Moser (2022): "Earnings Inequality and the Minimum Wage: Evidence from Brazil," *American Economic Review*, 112, 3803–3847.

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," 1–45.

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.

Galí, J. (1994): "Keeping up with the Joneses : Consumption Externalities , Portfolio Choice , and Asset Prices," *Journal of Money , Credit and Banking*, 26, 1–8.

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.

References III

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.

Glaeser, E. L., M. Resseger, and K. Tobio (2009): "Inequality in cities," *Journal of Regional Science*, 49, 617–646.

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

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

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

Liu, S. and X. Yang (2021): "Human capital externalities or consumption spillovers? The effect of high-skill human capital across low-skill labor markets," *Regional Science and Urban Economics*, 87, 103620.

Luttmer, E. F. (2005): "Neighbors as negatives: Relative earnings and well-being," *Quarterly Journal of Economics*, 120, 963–1002.

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.

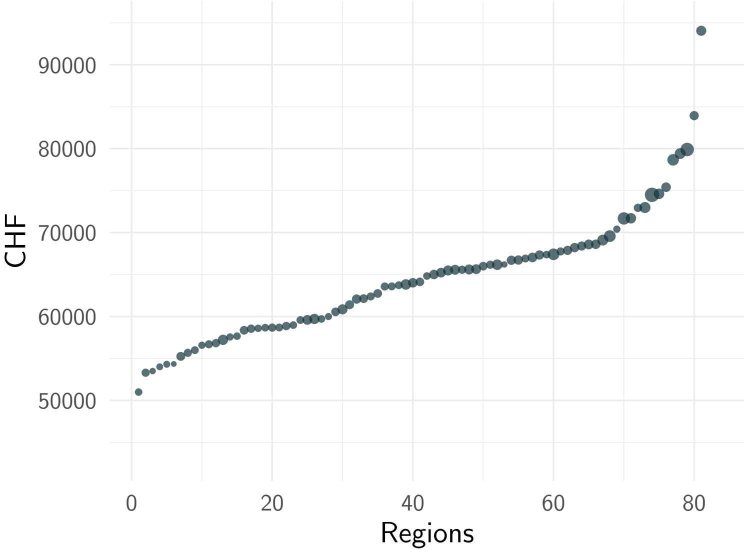
Melly, B. and M. Pons (2023): "Minimum Distance Estimation of Quantile Panel Data Models," *Working Paper*.

Volgushev, S., S.-K. Chao, and G. Cheng (2019): "Distributed inference for quantile regression processes," *The Annals of Statistics*, 47, 1634–1662.

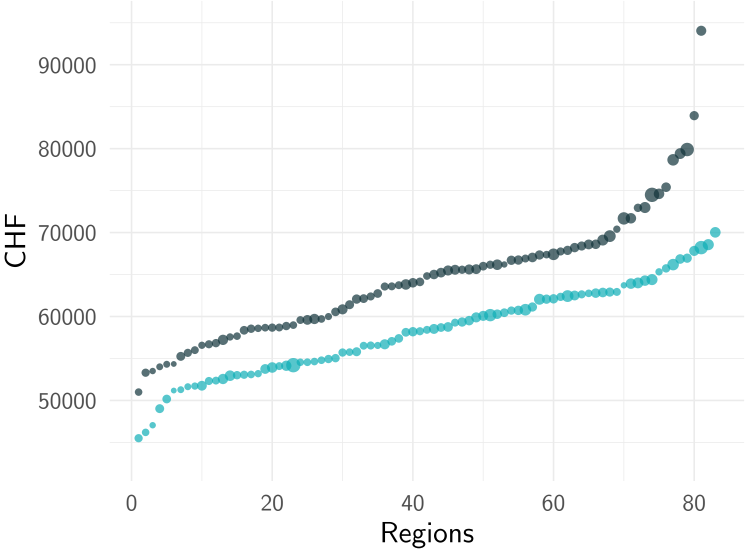
Descriptive Example - Income heterogeneity within and between regions

- 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)

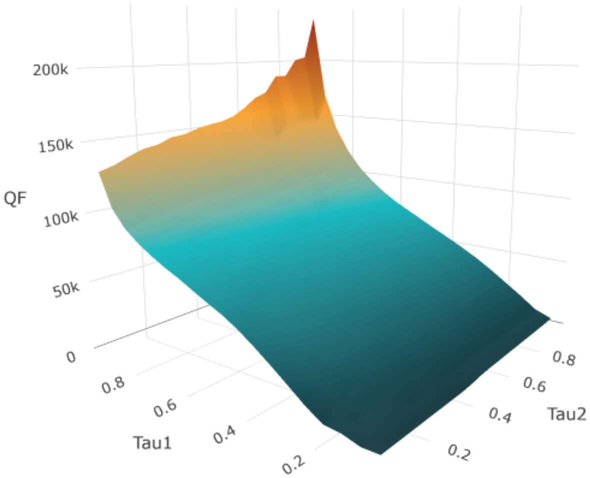
Average Income by Region



Average and Median Income by Region



Two Dimensional Quantile Function



Optimal Treatment Assignment

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

$$W = \int \int y_{ij} \cdot w(\text{Rank}(y_{ij})) di dj.$$

- With rank invariance of individuals and groups over treatment states, treatment effects for individuals at given quantiles are identified.
- Individuals rank implies some welfare weights for each individual and consequently of the group.

Optimal Treatment Assignment

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

$$W = \int \int y_{ij} \cdot w(\text{Rank}(y_{ij}))didj.$$

- 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.
- Here, the goal is to select a treatment rule that assigns individuals depending on their ranks (u_{ij}, v_j).
- Exploit treatment effect heterogeneity within and between groups to allocate the treatment more efficiently.

Welfare Function

Welfare is a weighted average of the outcomes with weights that depend on both within and between rank:

$$W = \int_0^1 \int_0^1 q(\tau_1, \tau_2) \cdot w(\tau_1, \tau_2) d\tau_2 d\tau_1$$

let $w(\tau_1, \tau_2) = w(\tau_1)$. The welfare function simplifies to a weighted average of the expectation of the group quantiles:

$$W = \int_0^1 E[q(\tau_1, v_j)] w(\tau_1) d\tau_1,$$

if $w(\tau_1, \tau_2) = w(\tau_2)$. The welfare function simplifies to the weighted average of the conditional expectation in each group:

$$W = E\left[\int_0^1 q(\tau_1, v_j) d\tau_1 w(v_j)\right] = E_j\left[E_{i|j}[y_{ij}] w(v_j)\right].$$

The welfare function assigns different weights to different groups. The outcome distribution within the group does not matter.

Asymptotics - Intuition

If the first stage parameter vector $\beta_0(\tau_1)$ was known, the true parameter vector $\delta_0(\beta_0, \tau)$ of the second stage quantile regression uniquely satisfies:

$$\mathbb{E}[m_{ij}(\delta_0, \beta_0, \tau)] = 0 \tag{4}$$

with $m_{ij}(\delta, \beta, \tau) = x_{ij}'[\tau_2 - 1(\tilde{x}_{ij}'\beta_j(\tau_1) \leq x_{ij}'\delta(\beta(\tau_1), \tau_2))]$.

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 –** The observations $(y_{ij}, x_{ij})_{i=1,\dots,n,\ j=1,\dots,m}$ are i.i.d across i and j.
2. **Covariates –** (i) For all $j = 1, \dots, m$ and all $i = 1, \dots, n$, $\|x_{ij}\| \leq C$ almost surely. (ii) The eigenvalues of $\mathbb{E}_{i|j}[x_{1ij}x'_{1ij}]$ and $\mathbb{E}[x_{ij}x'_{ij}]$ are bounded away from zero and infinity uniformly across j. (iii) as $m \to \infty$,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{j=1}^{m} \mathbb{E}_{i|j} \left[f_{Q(\tau_1, y_{ij}|\nu_j)|x_{ij}}(x'_{ij}\delta_0|x) x_{ij} \tilde{x}'_{ij} \right] = \mathbb{E}[f_{Q(\tau_1, y_{ij}|\nu_j)|x_{ij}}(x'_{ij}\delta_0|x) x_{ij} \tilde{x}'_{ij}]$$

where the eigenvalues of $\mathbb{E}[f_{Q(\tau_1, y_{ij}|\nu_j)|x_{ij}}(x'_{ij}\delta_0|x) x_{ij} x'_{ij}]$ are abounded from below and above.

Assumptions II

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

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

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}

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

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

Assumptions III

5. **Group level heterogeneity–** The conditional distribution $F_{Q(\tau_1, y_{ij}|x_{ij}, v_j)|x_{ij}}(q|x)$ is twice continuously differentiable w.r.t. q, with the corresponding derivatives $f_{Q(\tau_1, y_{ij}|x_{ij}, v_j)|x_{ij}}(q|x)$ and $f'_{Q(\tau_1, y_{ij}|x_{ij}, v_j)|x_{ij}}(q|x)$. Further, assume that

$$f_Q^{max} := \sup_{q \in \mathbb{R}, x \in \mathcal{X}} |f_{Q(\tau_1, y_{ij}|x_{ij}, v_j)|x_{ij}}(q|x)| < \infty$$

and

$$\bar{f}'_Q := \sup_{q \in \mathbb{R}, x \in \mathcal{X}} |f'_{Q(\tau_1, y_{ij}|x_{ij}, v_j)|x_{ij}}(q|x)| < \infty.$$

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

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

$$0 < f_{min} \leq \inf_{\tau_2 \in \mathcal{T}} \inf_{x \in \mathcal{X}} f_{Q(\tau_1, y_{ij}|x_{ij}, v_j)|x_{ij}}(x'_{ij}\delta_0(\tau)|x).$$

Assumptions IV

7. **Compact parameter space** – For all τ, $\beta_{j,0}(\tau_1) \in \text{int}(\mathcal{B}_j)$ and $\delta_0(\beta_0, \tau) \in \text{int}(\mathcal{D})$, where $\mathcal{B}_j$ and $\mathcal{D}$ are compact subsets of $\mathbb{R}^{K_1}$ and $\mathbb{R}^{K}$, respectively.
8. **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

Empirical Application

Specification:

$$Q(\tau_2, Q(\tau_1, y_{ijt}|D_{ij}, S_{ij}, W_{5,ij}, v_j)|D_{ij}, S_{ij}, W_{5,ijt}) = \beta_1(\tau)\cdot D_{ij} + \beta_2(\tau)\cdot S_{ij} + \beta_3(\tau)\cdot W_{5,ijt},$$

- y_{ijt}: outcome of firm i operating in market j in wave t
- 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).
- $W_{5,ijt}$: indicator variable for the last wave.

Rank Correlation

Table: Correlation of Ranks over τ_1

	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.2	1							
0.3	0.81	1						
0.4	0.67	0.81	1					
0.5	0.61	0.76	0.87	1				
0.6	0.53	0.64	0.78	0.86	1			
0.7	0.45	0.54	0.69	0.72	0.82	1		
0.8	0.35	0.44	0.58	0.62	0.69	0.8	1	
0.9	0.22	0.32	0.48	0.5	0.56	0.6	0.73	1

Note:
The table shows the correlation matrix of the ranks at different values of τ_1.

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.
  - $\beta(\tau_1, \tau_2) = 1 + 0.1 \cdot F^{-1}(\tau_1) + 0.1 \cdot F^{-1}(\tau_2)$
  - $\gamma(\tau_1, \tau_2) = 1 - 0.1 \cdot F^{-1}(\tau_1) - 0.1 \cdot F^{-1}(\tau_2)$.

- $(m, n) = \{(25, 25), (200, 25), (25, 200), (200, 200)\}$
- $\mathcal{T} = \{0.25, 0.5, 0.75\}$
- 2'000 Monte Carlo simulations.

back

Simulations

Table: Bias and Standard Deviation

$\tau_1 \setminus \tau_2$	β 0.25	β 0.5	β 0.75	γ 0.25	γ 0.5	γ 0.75
(m, n) = (25,25)						
0.25	-0.026	0.001	0.031	-0.019	0.004	0.028
	(0.114)	(0.109)	(0.118)	(0.237)	(0.223)	(0.243)
0.5	-0.027	-0.004	0.023	-0.020	0.000	0.023
	(0.112)	(0.104)	(0.109)	(0.241)	(0.219)	(0.240)
0.75	-0.034	-0.006	0.022	-0.020	-0.002	0.026
	(0.116)	(0.109)	(0.114)	(0.239)	(0.220)	(0.241)
(m, n) = (25,200)						
0.25	-0.010	-0.001	0.007	-0.008	0.000	0.007
	(0.074)	(0.066)	(0.072)	(0.234)	(0.219)	(0.230)
0.5	-0.008	-0.002	0.004	-0.008	-0.001	0.004
	(0.072)	(0.065)	(0.069)	(0.234)	(0.221)	(0.231)
0.75	-0.012	-0.003	0.005	-0.010	-0.002	0.004
	(0.074)	(0.067)	(0.070)	(0.235)	(0.219)	(0.231)

back

Simulations

Table: Bias and Standard Deviation

	β			γ		
$\tau_1 \setminus \tau_2$	0.25	0.5	0.75	0.25	0.5	0.75
(m, n) = (200,25)						
0.25	-0.022	0.006	0.031	-0.022	-0.001	0.021
	(0.042)	(0.038)	(0.042)	(0.079)	(0.073)	(0.079)
0.5	-0.025	-0.001	0.023	-0.020	-0.002	0.017
	(0.041)	(0.037)	(0.039)	(0.078)	(0.073)	(0.078)
0.75	-0.033	-0.007	0.020	-0.023	-0.003	0.018
	(0.042)	(0.038)	(0.041)	(0.079)	(0.074)	(0.081)
(m, n) = (200,200)						
0.25	-0.005	0.002	0.007	-0.004	0.000	0.006
	(0.028)	(0.026)	(0.028)	(0.076)	(0.070)	(0.078)
0.5	-0.004	0.001	0.005	-0.003	0.000	0.006
	(0.027)	(0.025)	(0.028)	(0.075)	(0.070)	(0.079)
0.75	-0.006	0.000	0.006	-0.004	0.000	0.006
	(0.027)	(0.026)	(0.028)	(0.076)	(0.070)	(0.078)

back

Simulations

Table: Bootstrap Standard Errors relative to Standard Deviation

	β			γ		
$\tau_1 \setminus \tau_2$	0.25	0.5	0.75	0.25	0.5	0.75
			(m, n) = (25,25)			
0.25	1.203	1.114	1.204	1.114	1.088	1.264
0.5	1.207	1.140	1.202	1.138	1.085	1.295
0.75	1.184	1.115	1.229	1.127	1.077	1.267
			(m, n) = (25,200)			
0.25	1.249	1.206	1.350	1.251	1.122	1.553
0.5	1.314	1.216	1.439	1.292	1.126	1.651
0.75	1.330	1.172	1.386	1.324	1.119	1.593

back

Simulations

Table: Bootstrap Standard Errors relative to Standard Deviation

	β			γ		
$\tau_1 \setminus \tau_2$	0.25	0.5	0.75	0.25	0.5	0.75
(m, n) = (200,25)						
0.25	1.054	1.025	1.019	1.035	1.029	1.019
0.5	1.036	1.022	1.015	1.003	1.017	1.025
0.75	1.018	1.012	1.033	1.005	0.998	1.021
(m, n) = (200,200)						
0.25	1.075	1.033	1.059	1.033	1.078	1.053
0.5	1.069	1.065	1.062	1.022	1.078	1.052
0.75	1.067	1.070	1.046	1.030	1.068	1.059

back

Simulations

Table: Coverage Probability of Bootstrap Confidence Interval

	β			γ		
$\tau_1 \setminus \tau_2$	0.25	0.5	0.75	0.25	0.5	0.75
(m, n) = (25,25)						
0.25	0.972	0.968	0.963	0.933	0.950	0.938
0.5	0.976	0.973	0.966	0.931	0.947	0.947
0.75	0.969	0.967	0.969	0.939	0.943	0.935
(m, n) = (25,200)						
0.25	0.987	0.986	0.985	0.946	0.954	0.960
0.5	0.984	0.982	0.986	0.951	0.953	0.959
0.75	0.986	0.986	0.982	0.949	0.952	0.954

back

Simulations

Table: Coverage Probability of Bootstrap Confidence Interval

$\tau_1 \setminus \tau_2$	β 0.25	β 0.5	β 0.75	γ 0.25	γ 0.5	γ 0.75
(m, n) = (200,25)						
0.25	0.925	0.950	0.888	0.940	0.935	0.926
0.5	0.912	0.949	0.904	0.929	0.941	0.936
0.75	0.881	0.944	0.921	0.924	0.929	0.925
(m, n) = (200,200)						
0.25	0.956	0.953	0.947	0.939	0.949	0.943
0.5	0.952	0.962	0.953	0.944	0.945	0.942
0.75	0.946	0.960	0.956	0.945	0.952	0.950

back

Questions

- Convergence Rate ▸ more
- Growth Condition ▸ more
- Degenerate Distribution ▸ more
- Smoothed Quantile Regression ▸ more
- Bias Correction ▸ more

Degenerate Distribution

- In similar settings, Galvao et al. (2020), Melly and Pons (2023) 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 $\sqrt{mn}\frac{1}{m}\sum_{j=1}^{m} \bar{\Gamma}_{2,j}(\delta_0, \beta_0, \tau)\left(\hat{\beta}_j(\tau_1) - \beta_{j,0}(\tau_1)\right) \xrightarrow{d} N(0, \Omega_1(\tau))$.
- 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.
- 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.

▸ Asym. Dist ▸ Questions

Convergence Rate - OLS

Both $\beta(\tau_1, \tau_2)$ and $\gamma(\tau_1, \tau_2)$ converge at the slower rate because I want to allow them to be different between groups.

Consider a dgp:

$$y_{ij} = x_{1ij}\beta + x_{2j}\gamma + \eta_j + \nu_{ij} \quad \text{with}$$

- It is possible to estimate β at the $1/\sqrt{mn}$ by exploiting only the within variation (i.e. fixed effects estimator).
- However, this strategy does not identify heterogeneity over groups.
- γ can only be estimated the a the $1/\sqrt{m}$ rate as x_{2j} varies only between groups.
  - Exception: if there is no group-level heterogeneity ($\eta_j = 0 \ \forall j = 1, \ldots, m$).

back

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

Smoothed Panel Data Quantile Regression

- 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 correction possible (analytical formula of Split panel Jackknife (see Dhaene and Jochmans, 2015). Bias corrected estimator is centered at zero under the same growth condition.
  - Growth rate required for unbiased asymptotic normality of FE QR used to be $m^2(\log m)(\log n)^2/n \to 0$.
  - Galvao et al. (2020) showed that unbiased asymptotic normality continues to hold provided $m(\log n)^2/n \to 0$.
  - The estimators considered in these papers converge at the $1/\sqrt{mn}$ rate.
- Smoothed QR estimator requires stronger conditions on the smoothness of the outcome variable and the choice of a bandwidth that is arbitrary.

back

Bias Correction

-

▸ back