Unconditional Quantile Regression with High-Dimensional Data

Yuya Sasaki[1] Takuya Ura[2] Yichong Zhang[3]

[1]Vanderbilt University

[2]University of California, Davis

[3]Singapore Management University

ES North American Winter Meeting
January 7, 2022

- This paper provides a method of estimation and inference for unconditional quantile partial effects (UQPE) with high-dimensional covariates (X).

- Before talking about this paper, I will spend a minute on the background (UQPE & high-dimensional covariates).

UQPE

- Unconditional quantile regression (Firpo, Fortin, and Lemieux, 2009) measures heterogeneous counterfactual effects.
- Unconditional quantile partial effect (UQPE) in the unconditional quantile regression.
    - “[A] simple way of performing detailed decompositions” (Fortin, Lemieux, and Firpo, 2011, p. 76).
    - Popular.

Unconditional quantile regressions

Authors	Sergio Firpo, Nicole M Fortin, Thomas Lemieux
Publication date	2009/5
Journal	Econometrica
Volume	77
Issue	3
Pages	953-973
Publisher	Blackwell Publishing Ltd
Description	We propose a new regression method to evaluate the impact of changes in the distribution of the explanatory variables on quantiles of the unconditional (marginal) distribution of an outcome variable. The proposed method consists of running a regression of the (recentered) influence function (RIF) of the unconditional quantile on the explanatory variables. The influence function, a widely used tool in robust estimation, is easily computed for quantiles, as well as for other distributional statistics. Our approach, thus, can be readily generalized to other distributional statistics.
Total citations	Cited by 2241

Scholar articles
Unconditional quantile regressions
S Firpo, NM Fortin, T Lemieux - Econometrica, 2009
Cited by 2241 Related articles All 34 versions

High-Dimensional Covariates

- Oaxaca-Blinder decomposition of F_Y into:
  - structural heterogeneity ($F_{Y|X}$) +
  - distributional heterogeneity (F_X)

  see Fortin et al. (2011).

- Causal interpretation under conditional exogeneity (Chernozhukov, Fernández-Val, and Melly, 2013, Sec. 2.3).

  We want to use high-dimensional covariates X.

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

UQPE

- Y = observed outcome.
- X = observed covariates.
- The counterfactual distribution of Y when X_1 changed by ε is

$$F_Y^{\varepsilon}(y) = \int F_{Y|X=(x_1+\varepsilon, x_{-1})} dF_X(x).$$

- The UQPE with respect to the first coordinate, X_1, of X is

$$UQPE(\tau) = \frac{\partial}{\partial \varepsilon} (F_Y^{\varepsilon})^{-1}(\tau)\Big|_{\varepsilon=0}.$$

UQPE with High-Dimensional Covariates

- Three steps of the traditional estimation technique:
  1. Estimate unconditional quantiles
  2. Estimate the Re-centered Influence Function (RIF) regression
  3. Integrate the RIF regression estimates
- To allow for high-dimensional covariates, X, we need some regularized estimator (a.k.a. machine learner) in the 2^{nd} step.

$$\Downarrow$$

The traditional method to incorporate estimation errors from the 2^{nd} step to the 3^{rd} step no longer applies.

$$\Downarrow$$

We propose a locally robust score.

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

Traditional Score

- Following Firpo et al. (2009), rewrite the UQPE in terms of identifiable parameters:

$$UQPE(\tau) = -\frac{\theta(\tau)}{f_Y(q_\tau)},$$

where q_τ is the τ-th quantile of Y and

$$\theta(\tau) = \int \frac{\partial F_{Y|X=x}(q_\tau)}{\partial x_1} dF_X(x).$$

$$\Downarrow$$

- Large bias & variance in estimation if X is high-dimensional.

Doubly and Locally Robust Score

- We therefore estimate another representation:

$$\theta(\tau) = \int \frac{\partial F_{Y|X=x}(q_\tau)}{\partial x_1} dF_X(x) - \underbrace{\int \omega(x)(1\{y \leq q_\tau\} - m_0(x, q_\tau))dF_{Y,X}(y, x)}_{\text{Adjustment for estimation of } m_1(x,q_\tau) = \frac{\partial F_{Y|X=x}(q_\tau)}{\partial x_1}}$$

$$= \int m_1(x, q_\tau) - \omega(x)(1\{y \leq q_\tau\} - m_0(x, q_\tau))dF_{Y,X}(y, x),$$

where

  - $\omega(x) = \frac{\partial}{\partial x_1} \log f_{X_1|X_{-1}=x_{-1}}(x_1)$,
  - $m_0(x, q) = F_{Y|X=x}(q)$, and
  - $m_1(x, q) = \partial m_0(x, q)/\partial x_1$.
- Key insight: Newey (1994).

Double Robustness

- This representation

$$\theta(\tau) = \int \left(m_1(x, q_\tau) - \omega(x)(1\{y \leq q_\tau\} - m_0(x, q_\tau))\right) dF_{Y,X}(y, x)$$

is doubly robust in the sense that

$$\theta(\tau) = \int \left(\tilde{m}_1(x, q_\tau) - \omega(x)(1\{y \leq q_\tau\} - \tilde{m}_0(x, q_\tau))\right) dF_{Y,X}(y, x)$$

and

$$\theta(\tau) = \int \left(m_1(x, q_\tau) - \tilde{\omega}(x)(1\{y \leq q_\tau\} - m_0(x, q_\tau))\right) dF_{Y,X}(y, x)$$

hold for a set of values that $(\tilde{\omega}, \tilde{m}_0, \tilde{m}_1)$ takes.

Proof Sketch

Local Robustness

- This representation

$$\theta(\tau) = \int (m_1(x, q_\tau) - \omega(x)(1\{y \leq q_\tau\} - m_0(x, q_\tau)))\, dF_{Y,X}(y, x)$$

is also locally robust in the sense that

$$\theta(\tau) \approx \int (\tilde{m}_1(x, q_\tau) - \tilde{\omega}(x)(1\{y \leq q_\tau\} - \tilde{m}_0(x, q_\tau)))\, dF_{Y,X}(y, x)$$

holds for a set of values that $(\tilde{\omega}, \tilde{m}_0, \tilde{m}_1)$ takes under conditions to be statd ahead.

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

Estimation Procedure

1. Let $\hat{q}_\tau$ = the τ-th empirical quantile of Y.

2. $\hat{f}_Y(y) = \sum_{i=1}^{N} \frac{1}{Nh_1} K_1\left(\frac{Y_i - y}{h_1}\right)$.

3. Construct an estimator $(\hat{\omega}(x), \hat{m}_0(x, q), \hat{m}_1(x, q))$ – *see below.*

4. $\hat{\theta}(\tau) = \frac{1}{N} \sum_{i \in [N]} \left[\hat{m}_1(X_i, \hat{q}_\tau) - \hat{\omega}(X_i)(1\{Y_i \leq \hat{q}_\tau\} - \hat{m}_0(X_i, \hat{q}_\tau))\right]$.

5. $\widehat{UQPE}(\tau) = -\frac{\hat{\theta}(\tau)}{\hat{f}_Y(\hat{q}_\tau)}$.

Estimation: $\hat{m}_0$ and $\hat{m}_1$

- Follow the Lasso logistic regression (Belloni, Chernozhukov, Fernández-Val, and Hansen, 2017).
- Approximately sparse logistic regression model

$$m_0(X, q) = \Lambda(b(X)^\top \beta_q) + r_m(X, q).$$

1. Lasso penalized logistic regression

$$\tilde{\beta}_q = \arg\min_{\beta} \frac{1}{N} \sum_{i \in [N]} M(1\{Y_i \leq q\}, b(X_i); \beta) + \frac{\lambda}{N} ||\Psi_q \beta||_1,$$

Estimation: $\hat{m}_0$ and $\hat{m}_1$, Continued

2. Post-Lasso estimator for β_q defined by

$$\hat{\beta}_q = \operatorname*{arg\,min}_{\beta \in \mathbb{R}^p : \text{Supp}(\beta) \subset \text{Supp}(\tilde{\beta}_q) \cup S_1} \frac{1}{N} \sum_{i \in [N]} M(1\{Y_i \leq q\}, b(X_i); \beta).$$

3. $\hat{m}_0(x, q) = \Lambda(b(x)^\top \hat{\beta}_q)$.

4. $\hat{m}_1(x, q) = \frac{\partial}{\partial x_1} \hat{m}_0(x, q)$.

Estimation: $\hat{\omega}$

- Follow the Riesz representer approach (Chernozhukov, Newey, and Singh, 2021a,b)
- Suppose

$$\omega(x) = h(x)^\top \overline{\rho} + \text{(approximation error)}.$$

- Since $\omega(x) = \frac{\partial}{\partial x_1} \log f_{X_1 | X_{-1} = x_{-1}}(x_1)$, the integration by parts yields

$$\mathbb{E} h(X) \omega(X) = -\mathbb{E} \partial_{x_1} h(X).$$

- Approximating $\omega(x)$ by $h(x)^\top \overline{\rho}$, we have

$$\mathbb{E}[h(X) h(X)^\top] \overline{\rho} = -\mathbb{E} \partial_{x_1} h(X) + \text{(approximation error)}.$$

Estimation: $\hat{\omega}$, Continued

- From the previous slide, we have

$$\mathbb{E}[h(X)h(X)^\top]\overline{\rho} = -\mathbb{E}\partial_{x_1}h(X) + (\text{approximation error}).$$

- $\bar{\rho} \approx \arg\min_\rho \left(-2M^\top \rho + \rho G \rho + 2\lambda_L |\rho|_1\right)$,
  where $G = \mathbb{P}h(X)h(X)^\top$ and $M = -\mathbb{P}\partial_{x_1}h(X)$.

1. $\hat{\rho} = \arg\min_\rho \left(-2\hat{M}_l^\top \rho + \rho^\top \hat{G}_l \rho + 2\lambda_L |\rho|_1\right)$, where
   $\hat{G} = \frac{1}{N}\sum_{i\in[N]} h(X_i)h(X_i)^\top$ and $\hat{M} = -\frac{1}{N}\sum_{i\in[N]} \partial_{x_1}h(X_i)$.
2. $\hat{\omega}(x) = h(x)^\top \hat{\rho}$

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

Bootstrap Procedure

- Steps to compute the boostrap estimator $\widehat{UQPE}^*(\tau)$:

1. Draw $\eta_i \overset{\text{i.i.d.}}{\sim} N(0,1)$ independently from data.
2. r_N^* = the integer part of $1 + \sum_{i=1}^N (\tau + \eta_i(\tau - \mathbf{1}\{Y_i \leq \hat{q}_\tau\}))$.
3. $\hat{q}_\tau^*$ = the r_N^*-th order statistic of Y_i.
4. $\hat{f}_Y^*(y) = \sum_{i=1}^N \frac{(\eta_i+1)}{\sum_{i=1}^N(\eta_i+1)} \frac{1}{h_1} K_1\left(\frac{Y_i-y}{h_1}\right)$.

Bootstrap Procedure, Cotinued

5. $\hat{\theta}^*(\tau)$:

$$\hat{\theta}^*(\tau) = \frac{1}{\sum_{i \in [N]} (\eta_i + 1)} \sum_{i \in [N]} (\eta_i + 1) \times \left[\hat{m}_1(X_i, \hat{q}^*_\tau) - \hat{\omega}(X_i)(1\{Y_i \le \hat{q}^*_\tau\} - \hat{m}_0(X_i, \hat{q}^*_\tau)) \right].$$

6. $\widehat{UQPE}^*(\tau)$:

$$\widehat{UQPE}^*(\tau) = -\frac{\hat{\theta}^*(\tau)}{\hat{f}^*_Y(\hat{q}^*_\tau)}.$$

Inference

- Uniform confidence band for $\{UQPE(\tau) : \tau \in \Upsilon\}$:
  - Let $c_{\Upsilon}(1-\alpha)$ = the $(1-\alpha)$ quantile of

$$\sup_{\tau \in \Upsilon} \left| (\widehat{UQPE}^{*}(\tau) - \widehat{UQPE}(\tau))/\hat{\sigma}(\tau) \right|$$

    where

$$\hat{\sigma}(\tau) = \frac{Q_{\widehat{UQPE}^{*}(\tau)}(0.75) - Q_{\widehat{UQPE}^{*}(\tau)}(0.25)}{1.34896}.$$

  - Bounds of CB_{Υ} on Υ are $\widehat{UQPE}(\tau) \pm \hat{\sigma}(\tau) c_{\Upsilon}(1-\alpha)$, $\tau \in \Upsilon$.

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

Assumption (Primitive Model Restrictions)

1. For every $\tau \in \Upsilon$, $F_Y^{\varepsilon}(q)$ is differentiable with respect to ε in a neighborhood of zero for every q in a neighborhood of q_τ, and $Q_\tau(F_Y^{\varepsilon})$ is well defined and is differentiable with respect to ε in a neighborhood of zero.
2. $\int \left[\sup_{q \in \mathcal{Q}^\delta} |m_1(x,q)|\right]^{2+d} dF_X(x)$ and $\int |\omega(x)|^{2+d} dF_X(x)$ are finite for some constant $d > 0$.
3. For every x_{-1} in the support of X_{-1}, the conditional distribution of X_1 given $X_{-1} = x_{-1}$ has a probability density function, denoted by $f_{X_1|X_{-1}}$, which is continuously differentiable everywhere and is zero on the boundary of the support of the conditional distribution of X_1.
4. $m_1(x,q)$ and $m_0(x,q)$ are differentiable with respect to q for $q \in \mathcal{Q}^\delta$, and the derivatives are bounded in absolute value uniformly over $x \in \text{Supp}(X)$ and $q \in \mathcal{Q}^\delta$.
5. $f_Y(y)$ is three times differentiable on $\mathcal{Q}^\delta$ with all the derivatives uniformly bounded. $f_Y(q_\tau) > 0$ for every $\tau \in \Upsilon$.

Conditions for $\hat{m}_0$ and $\hat{m}_1$

Assumption (Boundedness)

6. The following statements hold for positive constants $\delta, \bar{c}, \underline{c}$:

   (i) $\underline{c} \leq \mathbb{E} b_j(X)^2 \leq \bar{c}$ for every $j = 1, \ldots, p$.

   (ii) $\sup_{x \in \mathrm{Supp}(X), q \in \mathcal{Q}^\delta} |m_1(x, q)| \leq \bar{c}$.

   (iii) $\sup_{q \in \mathcal{Q}^\delta} || \frac{\partial}{\partial x_1} b(X)^\top \beta_q ||_{\mathbb{P}, \infty} \leq \bar{c}$.

Conditions for $\hat{m}_0$ and $\hat{m}_1$

Assumption (Restricted Eigenvalue Condition)

7. There are positive constants $\bar{c}, \underline{c}$ and a sequence $m_N \to \infty$ such that, with probability approaching one,

$$\underline{c} \leq \inf_{\beta \neq 0, ||\beta||_0 \leq m_N} \frac{||b(X)^\top \beta||_{\mathbb{P}_n,2}}{||\beta||_2} \leq \sup_{\beta \neq 0, ||\beta||_0 \leq m_N} \frac{||b(X)^\top \beta||_{\mathbb{P}_n,2}}{||\beta||_2} \leq \bar{c},$$

$$\underline{c} \leq \inf_{\beta \neq 0, ||\beta||_0 \leq m_N} \frac{||\frac{\partial}{\partial x_1} b(X)^\top \beta||_{\mathbb{P}_n,2}}{||\beta||_2} \leq \sup_{\beta \neq 0, ||\beta||_0 \leq m_N} \frac{||\frac{\partial}{\partial x_1} b(X)^\top \beta||_{\mathbb{P}_n,2}}{||\beta||_2} \leq \bar{c},$$

$$\sup_{\beta \neq 0, ||\beta||_0 \leq m_N} \left| \frac{||\frac{\partial}{\partial x_1} b(X)^\top \beta||_{\mathbb{P}_n,2}}{||\frac{\partial}{\partial x_1} b(X)^\top \beta||_{\mathbb{P},2}} - 1 \right| + \sup_{\beta \neq 0, ||\beta||_0 \leq m_N} \left| \frac{||\frac{\partial}{\partial x_1} b(X)^\top \beta||_{\mathbb{P}_n,2}}{||\frac{\partial}{\partial x_1} b(X)^\top \beta||_{\mathbb{P},2}} - 1 \right| = o_P(1)$$

where $||v||_0$ denotes the the number of nonzero coordinates of vector v.

Conditions for $\hat{m}_0$ and $\hat{m}_1$

Assumption (Sparsity)

8. $\sup_{q \in \mathcal{Q}^\delta} ||\beta_q||_0 \leq s_b$ for a sequence $s_b = s_{m.N}$ satisfying $s_b = o(m_N)$, $s_b \log(p_b) = o(N)$, where

$$\zeta_N = \max \left(|| \max_{j=1,\ldots,p_b} |b_j(X)| ||_{\mathbb{P},\infty}, || \max_{j=1,\ldots,p_b} |\frac{\partial}{\partial x_1} b_j(X)| ||_{\mathbb{P},\infty} \right).$$

Assumption (Approximation Error)

9.

$$\sup_{q \in \mathcal{Q}^\delta} \left\| \frac{\partial}{\partial x_1} r_m(X, q) \right\|_{\mathbb{P},2} = O((s_b \log(p_b)/N)^{1/2})$$

$$\sup_{q \in \mathcal{Q}^\delta} || \frac{\partial}{\partial x_1} r_m(X, q) ||_{\mathbb{P},\infty} = O((\log(p_b) s_b^2 \zeta_N^2 / N)^{1/2}).$$

Asymptotic Properties of $\hat{m}_0$ and $\hat{m}_1$

Theorem

For $j = 0$ *and* 1,

$$\sup_{q \in \mathcal{Q}^\delta} \int \left| \hat{m}_j(x,q) - m_j(x,q) \right|^2 dF_X(x) = O_P\left(\frac{s_b \log(p_b)}{N} \right),$$

$$\sup_{q \in \mathcal{Q}^\delta, x \in Supp(X)} \left| \hat{m}_j(x,q) - m_j(x,q) \right| = O_P\left(\zeta_N s_b \sqrt{\frac{\log(p_b)}{N}} \right).$$

Conditions for $\hat{\omega}$

Assumption (Boundedness)

6. There is C such that with probability one, $\max_{1\leq j\leq p_h} |h_j(X)| \leq C$.

Assumption (Approximation Error)

7. Suppose $|\hat{G}_l - G|_\infty + |\hat{M}_l - M|_\infty = O_p\left(\sqrt{\frac{\log(p_h)}{N}}\right)$ where $|A|_\infty = \max_{ij} |A_{ij}|$ for a matrix A.

Assumption (Sparsity)

8. There exists $C > 1$, $\xi \geq 1/2$ such that for $s_h = C\left(\frac{\log(p_h)}{N}\right)^{-1/(1+2\xi)}$ there is $\overline{\rho}$ with $||\overline{\rho}||_0 \leq s_h$ such that

$$\left(\int (\omega(x) - h(x)^\top \overline{\rho})^2 dF_X(x)\right)^{1/2} \leq C(s_h)^{-\xi} \quad \text{and} \quad ||\omega(x) - h(x)^\top \overline{\rho}||_{\mathbb{P},\infty} = o(1).$$

Conditions for $\hat{\omega}$

Assumption (Restricted Eigenvalue Condition)

9. Suppose G is nonsingular and both G and $\hat{G}$'s eigenvalues are uniformly bounded in n, with probability approaching one. Also, there is $\kappa > 3$ such that for $\rho = \rho^*$ and $\overline{\rho}$,

$$\inf_{\Delta:\Delta\neq 0, \sum_{j\in\mathcal{J}_\rho^c}|\Delta_j|\leq\kappa\sum_{j\in\mathcal{J}_\rho}|\Delta_j|} \frac{\Delta' G \Delta}{\sum_{j\in\mathcal{J}_\rho}\Delta_j^2} > 0.$$

Assumption (Tuning Parameter)

10. $\lambda_L = \ell_n \sqrt{\log(p_h)/N}$, where $\ell_n = \log(\log(N))$.

Asymptotic Properties of $\hat{\omega}$

Theorem

$$\int [\hat{\omega}(x) - \omega(x)]^2 \, dF_X(x) = O_p(\ell_n^2 s_h \log(p_h)/N)$$

and

$$\sup_{x \in Supp(X)} |\hat{\omega}(x) - \omega(x)| = o_p(1).$$

Conditions for $\hat{f}_Y$

Assumption

1. $K_1(\cdot)$ is a second-order symmetric kernel function with a compact support.
2. $h_1 = c_1 N^{-H}$ for some positive constant c_1 and some $1/2 > H \geq 1/5$.

Asymptotic Properties of $\widehat{UQPE}$

Theorem

$$\widehat{UQPE}(\tau) - UQPE(\tau) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{IF}_i(\tau) + \frac{\theta(\tau) f_Y^{(2)}(q_\tau)(\int u^2 K_1(u)du) h_1^2}{2 f_Y^2(q_\tau)} + R(\tau)$$

$$\widehat{UQPE}^*(\tau) - \widehat{UQPE}(\tau) = \frac{1}{N}\sum_{i=1}^{N} \eta_i \cdot \mathrm{IF}_i(\tau) + R^*(\tau),$$

where the residuals are $o_P((\log(N) N h_1)^{-1/2})$ *uniformly in* τ.

The influence function is $\mathrm{IF}_i(\tau) = \frac{\theta(\tau)}{f_Y^2(q_\tau) h_1} K_1\left(\frac{Y_i - q_\tau}{h_1}\right)$.

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

Simulation Setting

- The outcome variable:

$$Y \mid X \sim N\left(g(X_1) + \sum_{j=2}^{p} \alpha_j X_j, \;\; 1\right),$$

where

$$\begin{aligned}
g(x) &= x && \text{in DGP 1,} \\
g(x) &= x - 0.10 \cdot x^2 && \text{in DGP 2, and} \\
g(x) &= x - 0.10 \cdot x^2 + 0.01 \cdot x^3 && \text{in DGP 3.}
\end{aligned}$$

Simulation Setting, Continued

- The high-dimensional controls $(X_1, ..., X_p)^T$:

$$X_1 \mid (X_2, ..., X_p) \sim N\left(\sum_{j=2}^{p} \gamma_j X_j, 1\right) \quad \text{and}$$
$$(X_2, ..., X_p) \sim N(0, \Sigma_{p-1}),$$

where Σ_{p-1} is the $(p-1) \times (p-1)$ variance-covariance matrix whose (r, c)-element is $0.5^{2(|r-c|+1)}$.

Simulation Setting, Continued

- High-dimensional parameter vectors:

$$
\begin{aligned}
&\text{(i)} \quad (\alpha_2, \ldots, \alpha_p)^\top = (\gamma_2, \ldots, \gamma_p)^\top = (0.5^2, 0.5^3, ..., 0.5^p)^\top \\
&\text{(ii)} \quad (\alpha_2, \ldots, \alpha_p)^\top = (\gamma_2, \ldots, \gamma_p)^\top = (0.5^2, 0.5^{5/2}, ..., 0.5^{(p+2)/2})^\top \\
&\text{(iii)} \quad (\alpha_2, \ldots, \alpha_p)^\top = (\gamma_2, \ldots, \gamma_p)^\top = (0.5^2, 0.5^{7/3}, ..., 0.5^{(p+4)/3})^\top \\
&\text{(iv)} \quad (\alpha_2, \ldots, \alpha_p)^\top = (\gamma_2, \ldots, \gamma_p)^\top = (0.5^2, 0.5^{9/4}, ..., 0.5^{(p+6)/4})^\top
\end{aligned}
$$

- Focus on (i) in this presentation.

Additional Details on Simulations

- Lasso preliminary estimators.
- $h(x) = (x^T, (x^2)^T, (x^3)^T)^T$ for estimation of ω_l.
- $b(x) = (x^T, (x^2)^T, (x^3)^T)^T$ for estimation of m_0 and m_1.
- $h_1 = 1.06\hat{\sigma}(Y)N^{-1/5-0.01}$
  (the rule-of-thumb optimal choice undersmoothed)
- $\lambda_L = \log(\log(N))\sqrt{\log(\dim(h(X))/N)}$
  (following our assumption)
- $N = 500$.
- $p = 100$.

Simulation Results 1 – Our Proposed Method

				True	Estimates			95% Cover	
DGP	N	p	τ	UQPE	Mean	Bias	RMSE	Point	Unif.
1 (i)	500	100	0.20	1.00	1.03	0.03	0.16	0.948	0.956
			0.40	1.00	1.02	0.02	0.13	0.948	
			0.60	1.00	1.03	0.03	0.14	0.954	
			0.80	1.00	0.99	-0.01	0.16	0.948	
2 (i)	500	100	0.20	1.12	1.14	0.02	0.18	0.952	0.956
			0.40	1.03	1.05	0.02	0.13	0.946	
			0.60	0.95	0.98	0.03	0.13	0.950	
			0.80	0.87	0.88	0.00	0.15	0.950	
3 (i)	500	100	0.20	1.14	1.17	0.03	0.18	0.950	0.950
			0.40	1.04	1.06	0.02	0.13	0.942	
			0.60	0.97	1.00	0.03	0.13	0.944	
			0.80	0.91	0.90	0.00	0.13	0.952	

Simulation Results 2 – Conventional RIF-Logit

DGP	N	p	τ	True UQPE	Estimates			95% Cover	
					Mean	Bias	RMSE	Point	Unif.
1 (i)	500	25	0.20	1.00	1.05	0.05	0.17	0.872	0.902
			0.40	1.00	1.03	0.03	0.13	0.892	
			0.60	1.00	1.03	0.03	0.13	0.898	
			0.80	1.00	1.05	0.05	0.17	0.886	
	500	50	0.20	1.00	0.18	-0.82	1.32	0.008	0.000
			0.40	1.00	1.33	0.33	0.63	0.500	
			0.60	1.00	1.33	0.33	0.57	0.474	
			0.80	1.00	0.15	-0.85	1.06	0.010	
2 (i)	500	25	0.20	1.12	1.19	0.07	0.20	0.876	0.906
			0.40	1.03	1.06	0.04	0.14	0.884	
			0.60	0.96	0.99	0.03	0.12	0.900	
			0.80	0.88	0.92	0.04	0.15	0.878	
	500	50	0.20	1.12	0.08	-1.03	1.28	0.006	0.000
			0.40	1.03	1.36	0.34	0.58	0.464	
			0.60	0.95	1.24	0.29	0.49	0.528	
			0.80	0.87	0.34	-0.53	1.04	0.020	
3 (i)	500	25	0.20	1.14	1.22	0.07	0.21	0.886	0.912
			0.40	1.04	1.08	0.04	0.14	0.892	
			0.60	0.97	1.01	0.03	0.13	0.900	
			0.80	0.90	0.95	0.04	0.15	0.874	
	500	50	0.20	1.14	0.04	-1.10	1.15	0.006	0.000
			0.40	1.04	1.41	0.37	0.64	0.444	
			0.60	0.97	1.28	0.31	0.52	0.502	
			0.80	0.90	0.26	-0.65	0.98	0.016	

Simulation Results 3 – With & Without the Orthogonal Score

With the Doubly Robust Score						Without the Doubly Robust Score					
				95% Cover						95% Cover	
DGP	N	p	τ	Point	Unif.	DGP	N	p	τ	Point	Unif.
1 (i)	500	100	0.20	0.948	0.956	1 (i)	500	100	0.20	0.930	0.912
			0.40	0.948					0.40	0.912	
			0.60	0.954					0.60	0.902	
			0.80	0.948					0.80	0.910	
2 (i)	500	100	0.20	0.952	0.956	2 (i)	500	100	0.20	0.924	0.910
			0.40	0.946					0.40	0.908	
			0.60	0.950					0.60	0.906	
			0.80	0.950					0.80	0.916	
3 (i)	500	100	0.20	0.950	0.950	3 (i)	500	100	0.20	0.932	0.914
			0.40	0.942					0.40	0.910	
			0.60	0.944					0.60	0.908	
			0.80	0.952					0.80	0.922	

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

Job Corps Training

- Largest training program for disadvantaged youth in the U.S.
- Interested in heterogeneous counterfactual marginal effects of the duration of the exposure to the program on hourly wages.
- 42 observed controls (and their powers).
- Different sets of observations are missing across different variables $\Rightarrow$ Take the intersection of non-missing observations.
  - $n = 347$

		25th Percentile	Median	Mean	75th Percentile	Non-Missing
Outcome Y	Hourly wage	4.750	5.340	5.892	6.500	7606
Treatment X_1	Days in Job Corps	54.0	129.0	153.4	237.0	4748
	Days taking classes	41.0	91.0	120.2	179.0	4207
Controls X_{-1}	Age	17.00	18.00	18.43	20.00	14653
	Female	0.000	0.000	0.396	1.000	14653
	White	0.000	0.000	0.303	1.000	14327
	Black	0.000	1.000	0.504	1.000	14327
	Hispanic origin	0.000	0.000	0.184	0.000	14288
	Native language is English	1.000	1.000	0.855	1.000	14327
	Years of education	9.00	10.00	10.24	11.00	14327
	Other job trainings	0.000	0.000	0.339	1.000	13500
	Mother's education	11.00	12.00	11.53	11.53	11599
	Mother worked	1.000	1.000	0.752	1.000	14223
	Father's education	11.00	12.00	11.50	12.00	8774
	Father worked	0.000	1.000	0.665	1.000	12906
	Received welfare	0.000	1.000	0.563	1.000	14327
	Head of household	0.000	0.000	0.123	0.000	14327
	Number of people in household	2.000	3.000	3.890	5.000	14327
	Married	0.000	0.000	0.021	0.000	14327
	Separated	0.000	0.000	0.017	0.000	14327
	Divorced	0.000	0.000	0.007	0.000	14327
	Living with spouse	0.000	0.000	0.014	0.000	14235
	Child	0.000	0.000	0.266	1.000	13500
	Number of children	0.000	0.000	0.347	0.000	13500
	Past work experience	0.000	1.000	0.648	1.000	14327
	Past hours of work per week	0.000	24.00	25.15	40.00	14299
	Past hourly wage	4.250	5.000	5.142	5.500	7884
	Expected wage after training	7.000	9.000	9.910	11.000	6561
	Public housing or subsidy	0.000	0.000	0.200	0.000	14327
	Own house	0.000	0.000	0.411	1.000	11457
	Have contributed to mortgage	0.000	0.000	0.255	1.000	13951
	Past AFDC	0.000	0.000	0.301	1.000	14327
	Past SSI or SSA	0.000	0.000	0.251	1.000	14327
	Past food stamps	0.000	0.000	0.438	1.000	14327
	Past family income $\geq$ \$12K	0.000	1.000	0.576	1.000	14327
	In good health	1.000	1.000	0.871	1.000	14327
	Physical or emotional problem	0.000	0.000	0.049	0.000	14327
	Smoke	0.000	1.000	0.537	1.000	14327
	Alcohol	0.000	1.000	0.584	1.000	14327
	Marijuana or hashish	0.000	0.000	0.369	1.000	14327
	Cocaine	0.000	0.000	0.033	0.000	14327
	Heroin/opium/methadone	0.000	0.000	0.012	0.000	14327
	LSD/peyote/psilocybin	0.000	0.000	0.055	0.000	14327
	Arrested	0.000	0.000	0.266	1.000	14327
	Number of times arrested	0.000	0.000	0.537	1.000	14218

Heterogeneous Counterfactual Marginal Effects

	Outcome	Treatment	τ	$\widehat{UQPE}(\tau)$	Pointwise 95% CI		Uniform 95% CB	
(I)	Hourly wage	Days in Job Corps	0.2	1.16	[0.79	1.54]	[0.30	2.03]
			0.4	1.95	[1.52	2.39]	[0.94	2.97]
			0.6	1.60	[0.26	2.94]	[0.11	3.09]
			0.8	4.56	[2.96	6.16]	[-0.67	9.79]
(II)	Log hourly wage	Days in Job Corps	0.2	0.20	[0.13	0.27]	[0.02	0.38]
			0.4	0.50	[0.30	0.69]	[-0.12	1.11]
			0.6	0.12	[-0.15	0.38]	[-0.20	0.43]
			0.8	0.66	[0.37	0.96]	[-0.04	1.37]
(III)	Hourly wage	Days in Job Corps classes	0.2	2.69	[0.08	5.30]	[-19.06	24.44]
			0.4	2.66	[2.07	3.25]	[-0.48	5.80]
			0.6	1.14	[0.00	2.29]	[-0.58	2.87]
			0.8	5.30	[2.76	7.84]	[-5.64	16.25]
(IV)	Log hourly wage	Days in Job Corps classes	0.2	0.46	[0.01	0.90]	[-4.24	5.15]
			0.4	0.64	[0.38	0.89]	[-1.38	2.65]
			0.6	0.17	[-0.22	0.55]	[-0.25	0.58]
			0.8	0.77	[0.42	1.13]	[-1.29	2.84]

Variables Selected at Different Quantiles

τ	0.2	0.4	0.6	0.8
	Intercept	Intercept	Intercept	Intercept
				Married
			Separated	Separated
				Living with spouse
			Education	Education
	Number of people in household	Number of people in household	Number of people in household	Number of people in household

UQPE and High-Dimensional Covariates

Robust Score and Estimation

Estimation Procedure

Bootstrap Inference

Asymptotic Theory

Simulation Studies

Heterogeneous Counterfactual Effects of Job Corps Training

Summary

Summary

- UQPE (Firpo et al., 2009)
  - *"a simple way of performing detailed decompositions"*
  - Popular (many Google Scholar citations)
- High-dimensional $X \Rightarrow$ unconfoundedness is more plausible.
- Robust score.
- Lasso + Riesz representer.
- Bootstrap inference.
- Simulations
- Heterogeneous counterfactual marginal effects of Job Corps.

Doubly Robust Score: A Proof Sketch

$$\int (\tilde{m}_1(x,q_\tau) - \omega(x)(1\{y \leq q_\tau\} - \tilde{m}_0(x,q_\tau)))\, dF_{Y,X}(y,x)$$
$$= \int \tilde{m}_1(x,q_\tau) dF_X(x) - \iint (m_0(x,q_\tau) - \tilde{m}_0(x,q_\tau)) \left(\frac{\partial}{\partial x_1} f_{X_1|X_{-1}=x_{-1}}(x_1)\right) dx_1 dF_{X_{-1}}(x_{-1})$$
$$= \int \tilde{m}_1(x,q_\tau) dF_X(x) + \iint \left(m_1(x,q_\tau) - \left(\frac{\partial}{\partial x_1}\tilde{m}_0(x,q_\tau)\right)\right) \left(f_{X_1|X_{-1}=x_{-1}}(x_1)\right) dx_1 dF_{X_{-1}}(x_{-1})$$
$$= \int m_1(x,q_\tau) dF_X(x)$$
$$= \theta(\tau)$$

$$\int (m_1(x,q_\tau) - \tilde{\omega}(x)(1\{y \leq q_\tau\} - m_0(x,q_\tau)))\, dF_{Y,X}(y,x)$$
$$= \int m_1(x,q_\tau) dF_X(x) - \iint \tilde{\omega}(x)(m_0(x,q_\tau) - m_0(x,q_\tau)) f_{X_1|X_{-1}=x_{-1}} dx_1 dF_{X_{-1}}(x_{-1})$$
$$= \int m_1(x,q_\tau) dF_X(x)$$
$$= \theta(\tau).$$

◂ Go Back

Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2017): "Program Evaluation with High-dimensional Data," *Econometrica*, 85, 233–298.

Chernozhukov, V., I. Fernández-Val, and B. Melly (2013): "Inference on counterfactual distributions," *Econometrica*, 81, 2205–2268.

Chernozhukov, V., W. K. Newey, and R. Singh (2021a): "Automatic debiased machine learning of causal and structural effects," *Econometrica, forthcoming.*

——— (2021b): "Debiased Machine Learning of Global and Local Parameters Using Regularized Riesz Representers," *Econometrics Journal, forthcoming.*

Firpo, S., N. M. Fortin, and T. Lemieux (2009): "Unconditional quantile regressions," *Econometrica*, 77, 953–973.

Fortin, N., T. Lemieux, and S. Firpo (2011): "Decomposition methods in economics," in *Handbook of labor economics*, Elsevier, vol. 4, 1–102.

Newey, W. K. (1994): "The asymptotic variance of semiparametric estimators," *Econometrica*, 1349–1382.