1 Quantile treatment effects

Quantile treatment effects (QTE) are point identified very similarly to $LATE$, local average treatment effects, when a binary instrument exists (see Abadie, Angrist, and Imbens [1998]). In addition to standard identification conditions for $LATE$ (potential outcomes are independent of instruments ν, treatment z is meaningfully related to the instruments, and treatment adoption is uniformly increasing in the instruments; see chapter 3), $QTE(\theta)$ uniqueness can only be assured if the θ-quantiles for $Y(0)$ and $Y(1)$ conditional on $z_1 - z_0 = 1$ (defined below) are unique. Some explanation is in order.

Quantiles are typically defined by the distribution function,

$$F(y) = \sum_{i=y_l}^{y_u} \Pr(y_i)$$

where y_l is the lower bound of support for y and y_u is the upper bound of support for y. However, if we define another function, say

$$G(y) = \sum_{i=y_u}^{y_l} \Pr(y_i),$$

the quantile is unambiguous or unique if $F^{-1}(\theta) = G^{-1}(1-\theta)$. This statement is always true for continuous random variables but may fail for random variables with discrete support. An example helps clarify.

Suppose the data generating process is uniform$\{1, 2, 3, 4\}$. Then,

y	$F(y)$	$G(y)$
1	0.25	1.0
2	0.50	0.75
3	0.75	0.50
4	1.0	0.25

First, second, and third quartiles are ambiguous but immediate surrounding quantiles are not.

θ	$F^{-1}(\theta)$	$G^{-1}(1-\theta)$
0.24	1	1
0.25	1	2
0.26	2	2
0.49	2	2
0.5	2	3
0.51	3	3
0.74	3	3
0.75	3	4
0.76	4	4

Point identification of $QTE(\theta)$ may fail if the $\theta-$quantile for $Y(0)$ or $Y(1)$ conditional on $z_1 - z_0 = 1$ is ambiguous. When θ-quantiles for $Y(0)$ or $Y(1)$

conditional on $z_1 - z_0 = 1$ are unambiguous, the $\theta-$quantile treatment effect conditional on X is

$$\begin{aligned} QTE\left(\theta \mid X = x, z_1 - z_0 = 1\right) &\equiv Q_\theta\left(Y\left(1\right) \mid X = x, z_1 - z_0 = 1\right) \\ &\quad -Q_\theta\left(Y\left(0\right) \mid X = x, z_1 - z_0 = 1\right) \\ &= \alpha \end{aligned}$$

where $Q_\theta\left(\cdot\right)$ refers to $\theta-$quantile of the random variable, α is the conditional quantile treatment effect from a quantile regression with treatment $z = 0, 1$ and covariates X,

$$Y \equiv zY\left(1\right) + \left(1 - z\right)Y\left(0\right) = \alpha z + X\beta + \varepsilon$$

$x\beta$ is the $\theta-$quantile for $Y\left(0\right)$ conditional on $X = x_i$, z_1 is treatment when the instrument $\nu = 1$ and z_0 is treatment when $\nu = 0$. Hence, $z_1 - z_0 = 1$ refers to the target subpopulation of compliers. That is, those individuals who adopt treatment when the instrument is manipulated from zero to one.

Next, we explore two instrumental variable strategies to identify conditional (on the covariates) quantile treatment effects and compare them with exogenous treatment. In addition to the approach outlined above focusing on the subpopulation of compliers proposed by Abadie, Angrist, and Imbens (AAI), we consider an approach developed by Chernozhukov and Hansen (CH) that seeks to identify QTE for the entire population. Then, we explore identification of marginal (not conditional on the covariates) quantile treatment effects proposed by Firpo (F).

1.1 Identification

First, consider conventional quantile regression. Parameters are identified via

$$\underset{b}{argmin} \quad E\left[\left(Y - Xb\right) \cdot \left(\theta - 1\left\{Y - Xb < 0\right\}\right)\right]$$

The intuition for this result is that $\widehat{Y} = Xb$ is chosen to minimize the expected linear loss where the loss function is

$$C\left(\widehat{Y}, Y \mid X\right) = \begin{array}{ll} c_1\left|\widehat{Y} - Y\right| & \widehat{Y} \leq Y \\ c_2\left|\widehat{Y} - Y\right| & \widehat{Y} > Y \end{array}$$

and $\theta = \frac{c_1}{c_1 + c_2}$. Suppose $c_1 = \alpha\theta$ then $\theta = \frac{\alpha\theta}{\alpha\theta + c_2}$ or $1 = \frac{\alpha}{\alpha\theta + c_2}$ which gives $c_2 = \alpha\left(1 - \theta\right)$. In other words, c_1 is proportional to θ as c_2 is proportional to $1 - \theta$. The QTE identification strategies are similar in that they employ variations (typically, different weighting schemes) on the above conventional quantile regression.

1.1.1 AAI identification strategy

The AAI conditions

(1) (independence) for $Y_{\nu z}$ and z_ν, Y_{10}, Y_{11}, Y_{00},Y_{01}, z_1, and z_0 are independent of the instrument ν given X,

(2) (exclusion) $Y_z = Y_{0z} = Y_{1z}$,

(3) (non-trivial assignment) $\Pr(z = 1 \mid X) \in (0, 1)$,

(4) (first-stage) $E[z_1 \mid X] \neq E[z_0 \mid X]$, and

(5) (monotonicity or uniformity) $\Pr(z_1 \geq z_0 \mid X) = 1$,

lead to

$$\underset{a,b}{argmin} \quad E\left[\begin{array}{c} \left(1 - \frac{(1-\nu)z}{\Pr(\nu=0|X)} - \frac{(1-z)\nu}{\Pr(\nu=1|X)}\right) \\ \times (Y - az - Xb)(\theta - 1\{Y - az - Xb < 0\}) \end{array} \right]$$

equals α, β. Hence, the conditional $\theta-$quantile treatment effect, α, is point identified for the (unidentified) subpopulation of compliers and the variation from conventional quantile regression is that $\kappa = \left(1 - \frac{(1-\nu)z}{\Pr(\nu=0|X)} - \frac{(1-z)\nu}{\Pr(\nu=1|X)}\right)$ replaces weight equal to one.

Now, let's explore the derivation of the weights

$$\kappa = \left(1 - \frac{(1-\nu)z}{\Pr(\nu = 0 \mid X)} - \frac{(1-z)\nu}{\Pr(\nu = 1 \mid X)}\right)$$

in the above expression

$$\begin{array}{rl} & E[\kappa(Y - az - Xb) \cdot (\theta - 1\{Y - az - Xb < 0\})] \\ = & E\left[\begin{array}{c} \left(1 - \frac{(1-\nu)z}{\Pr(\nu=0|X)} - \frac{(1-z)\nu}{\Pr(\nu=1|X)}\right) \\ \times (Y - az - Xb)(\theta - 1\{Y - az - Xb < 0\}) \end{array} \right] \end{array}$$

For any real function ψ of (Y, X, Z), $\psi(Y, X, Z)$ along with independence condition (1),

$$\frac{E[\kappa\psi(Y, X, Z)]}{\Pr(z_1 > z_0)} = E[\psi(Y, X, Z) \mid z_1 > z_0]$$

To see this, write via Bayes and monotonicity or uniformity (precludes defiers)

$$\begin{array}{rcl} E[\psi \mid X] & = & E[\psi \mid X, z_1 > z_0] \Pr(z_1 > z_0 \mid X) \\ & & +E[\psi \mid X, z_1 = z_0 = 1] \Pr(z_1 = z_0 = 1 \mid X) \\ & & +E[\psi \mid X, z_1 = z_0 = 0] \Pr(z_1 = z_0 = 0 \mid X) \end{array}$$

Rearranging gives

$$\begin{array}{rl} & E[\psi \mid X, z_1 > z_0] \\ = & \frac{1}{\Pr(z_1 > z_0 \mid X)} \\ & \times \left\{ \begin{array}{c} E[\psi \mid X] - E[\psi \mid X, z_1 = z_0 = 1] \Pr(z_1 = z_0 = 1 \mid X) \\ -E[\psi \mid X, z_1 = z_0 = 0] \Pr(z_1 = z_0 = 0 \mid X) \end{array} \right\} \end{array}$$

Monotonicity means $\nu = 1$ and $z = 0$ are never-takers which implies

$$\begin{aligned} E\left[\psi \mid X, z_1 = z_0 = 0\right] &= E\left[\psi \mid X, \nu = 1, z = 0\right] \\ &= E\left[\frac{\psi\nu\left(1-z\right)}{\Pr\left(\nu = 1, z = 0 \mid X\right)} \mid X\right] \\ &= E\left[\frac{\psi\nu\left(1-z\right)}{\Pr\left(\nu = 1 \mid X, z = 0\right)\Pr\left(z = 0 \mid X\right)} \mid X\right] \\ &= \frac{1}{\Pr\left(\nu = 1 \mid X, z = 0\right)} E\left[\frac{\psi\nu\left(1-z\right)}{\Pr\left(z = 0 \mid X\right)} \mid X\right] \end{aligned}$$

Uniformity and independence imply

$$\Pr\left(\nu = 1 \mid X, z = 0\right) = \Pr\left(z_1 = z_0 = 0 \mid X\right)$$

Hence,

$$E\left[\psi \mid X, z_1 = z_0 = 0\right]\Pr\left(z_1 = z_0 = 0 \mid X\right) = E\left[\frac{\psi\nu\left(1-z\right)}{\Pr\left(z = 0 \mid X\right)} \mid X\right]$$

Likewise, $\nu = 0$ and $z = 1$ are always-takers and

$$\begin{aligned} E\left[\psi \mid X, z_1 = z_0 = 1\right] &= E\left[\psi \mid X, \nu = 0, z = 1\right] \\ &= E\left[\frac{\psi z\left(1-\nu\right)}{\Pr\left(\nu = 0, z = 1 \mid X\right)} \mid X\right] \\ &= E\left[\frac{\psi z\left(1-\nu\right)}{\Pr\left(\nu = 0 \mid X, z = 1\right)\Pr\left(z = 1 \mid X\right)} \mid X\right] \\ &= \frac{1}{\Pr\left(\nu = 0 \mid X, z = 1\right)} E\left[\frac{\psi z\left(1-\nu\right)}{\Pr\left(z = 1 \mid X\right)} \mid X\right] \end{aligned}$$

Again, uniformity and independence imply

$$\Pr\left(\nu = 0 \mid X, z = 1\right) = \Pr\left(z_1 = z_0 = 1 \mid X\right)$$

Hence,

$$E\left[\psi \mid X, z_1 = z_0 = 1\right]\Pr\left(z_1 = z_0 = 1 \mid X\right) = E\left[\frac{\psi z\left(1-\nu\right)}{\Pr\left(z = 1 \mid X\right)} \mid X\right]$$

Substitution into the expression above yields

$$E\left[\psi \mid X, z_1 > z_0\right] = \frac{E\left[\psi \mid X\right] - E\left[\frac{\psi z(1-\nu)}{\Pr(z=1|X)} \mid X\right] - E\left[\frac{\psi\nu(1-z)}{\Pr(z=0|X)} \mid X\right]}{\Pr\left(z_1 > z_0 \mid X\right)}$$

Iterated expectations produces the result

$$\begin{aligned} E_X\left[E\left[\psi \mid X, z_1 > z_0\right]\right] &= \frac{E\left[\psi\right] - E\left[\frac{\psi z(1-\nu)}{\Pr(z=1|X)}\right] - E\left[\frac{\psi\nu(1-z)}{\Pr(z=0|X)}\right]}{\Pr\left(z_1 > z_0\right)} \\ E\left[\psi \mid z_1 > z_0\right] &= \frac{E\left[\kappa\psi\right]}{\Pr\left(z_1 > z_0\right)} \end{aligned}$$

Let $\psi = (Y - az - Xb)(\theta - 1\{Y - az - Xb < 0\})$ and the $\theta-$quantile AAI instrumental variable identification strategy for the subpopulation of compliers is complete.

1.1.2 CH identification strategy

The CH identification strategy revolves around defining a quantile response function, $Y_z \equiv q(z, x, U_z)$ where Y_z is potential outcome with treatment z, x is potential values of covariates, and U_z is unobservable outcome. For $U_z \sim U(0,1)$ and $U_z = \theta$, $q(z, x, U_z)$ identifies the θ-th quantile for potential outcome Y_z. CH identification conditions include

(1) (potential outcomes) $Y_z = q(z, x, U_z)$, where $q(z, x, \theta)$ is strictly increasing in θ and $U_z \sim U(0,1)$,

(2) (independence) conditional on X, U_z is independent of ν,

(3) (selection) $Z \equiv \delta(\nu, X, V)$ is some unknown function δ and random vector V,

(4) (rank invariance or rank similarity) conditional on $X = x$ and $\nu = \nu$, $U_z = U_{z'}$, or $U_z \sim U_{z'}$ conditional on V, and

(5) (observables) observable variables consist of $Y_z \equiv q(z, x, U_z), Z, X$, and ν.

The principle follows from conventional quantile regression proposed by Koenker and Bassett [1978]

$$Q_{Y|X}(\theta) \in \arg\min_{f \in \mathcal{F}} E[\rho_\theta(Y - f(X))]$$

where $\rho_\theta(u) = (\theta - 1(u < 0))u$, the asymmetric least absolute deviation loss. The implication of the CH identification conditions is 0 is the τ-th quantile of $Y - q(Z, X, \theta)$ conditional on covariates and instruments (X, ν).

$$0 = Q_{Y - q(Z,X,\theta)}(\theta \mid X, \nu)$$

Thus, CH pose the problem as finding a function that satisfies

$$0 \in \arg\min_{f \in \mathcal{F}} E[\rho_\theta(Y - q(Z, X, \theta) - f(X, \nu))]$$

1.1.3 F identification strategy

As with the other identification strategies outcome is continuous (to avoid ambiguity in the quantiles — the examples violate this but we accommodate this via partial identification), in addition the F (Firpo) identification strategy draws on

(1) (strong ignorability) Conditional on X, $Y(1), Y(0)$ are independent of treatment, Z.

(2) (common support) Let $p(x) = \Pr(Z = 1 \mid X = x)$, then $c \leq p(x) \leq 1 - c$ for $c > 0$.

1.2 Estimation

Conventional quantile regression has a linear programming (LP) formulation.

$$\begin{aligned} \min_{\tau \geq 0} \quad & c^T \tau \\ s.t. \quad & A\tau = Y \end{aligned}$$

where $c = (o, o, \theta \cdot \iota, (1-\theta) \cdot \iota)^T$, $\tau = (a^+, b^+, a^-, b^-, u^+, u^-)^T$, $u = Y - za - Xb$, $A = [z, X, -z, -X, I, -I]$, o is an $h+1$ element vector of zeroes, ι is an n element vector of ones, X is an $n \times h$ matrix of covariates, I is an $n \times n$ identity matrix, b has h elements, e^+ denotes the positive part and e^- denotes the negative part of real number e.

1.2.1 AAI estimation

AAI estimation of $QTE(\theta)$ involves a variation on the above where c is re-defined as $(o, o, \theta \cdot K, (1-\theta) \cdot K)^T$ and $K = \kappa_1, \ldots, \kappa_n$, an n element vector composed of the sample analog of $\left(1 - \frac{(1-\nu)z}{\Pr(\nu=0|X)} - \frac{(1-z)\nu}{\Pr(\nu=1|X)}\right)$. However, when κ_i is negative (for instance, $\nu = 1$ and $z = 0$) the LP is unbounded. This necessitates further modification. Two additional constraints and one additional parameter, s_i, are added for each instance where κ_i is negative.

$$\begin{aligned} u_i^+ &\leq M s_i \\ u_i^- &\leq M(1 - s_i) \end{aligned}$$

where M is a large (nonbinding) constant, and $s_i \in \{0, 1\}$, an integer. In other words, we now have a mixed integer linear program ($MILP$) formulation for QTE estimation.

1.2.2 CH estimation

CH focus on the basic linear-in-parameters model

$$q(z, x, \theta) = z\alpha(\theta) + x'\beta(\theta)$$

for the binary treatment case. Define the finite-sample analog weighted quantile regression objective function as

$$Q_n(\alpha, \beta, \gamma) = \frac{1}{n} \sum_{i=1}^{n} \rho_\theta \left(Y_i - Z_i\alpha - X_i'\beta - \widehat{\Phi}_i(\theta)\gamma\right) \widehat{V}_i(\theta)$$

where $\widehat{\Phi}_i(\theta)$ is a transformation of the instruments (a likely practical candidate is the projection of Z onto (X, ν)) and $\widehat{V}_i(\theta)$ is a positive weight function (practical candidates include $\widehat{V}(\theta) = I$ or alternatively $\widehat{V}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \widehat{\Phi}_i(\theta)\left(\widehat{\Phi}_i(\theta)\right)^T$.

CH propose a simple practical implementation.

(1) Construct a grid of values $\alpha_j, j = 1, \ldots, J$. Employ conventional quantile regression of $Y_i - Z_i\alpha$ on X_i and $\widehat{\Phi}_i(\theta)$ to obtain estimates $\widehat{\beta}(\alpha_j, \theta)$ and $\widehat{\gamma}(\alpha_j, \theta)$.

(2) Choose $\widehat{\alpha}(\theta) = \widehat{QTE}(\theta)$ from the α_j such that $\widehat{\gamma}(\alpha_j, \theta)^T \widehat{V}(\theta) \widehat{\gamma}(\alpha_j, \theta)$ is zero or as near as possible to zero.

1.2.3 F estimation

Let $\widehat{p}(x)$ denote the nonparametric estimator for the propensity score. Then, the estimator for the marginal quantile treatment effect is

$$mQTE(\theta) = \widehat{Q}_1(\theta) - \widehat{Q}_0(\theta)$$

where, for $j = 0, 1$,

$$\widehat{Q}_j(\theta) = \arg\min_q \sum_{i=1}^{n} \omega_{j,i} \rho_\theta (Y_i - q)$$

$$\rho_\theta(u) = u(\theta - 1(u \leq 0))$$

$$\omega_{1,i} = \frac{Z_i}{N \cdot \widehat{p}(X_i)}$$

and

$$\omega_{0,i} = \frac{1 - Z_i}{N \cdot (1 - \widehat{p}(X_i))}$$

The examples below involve no covariates except for the F strategy to make conditional and marginal QTE comparable. Also, $\omega_{j,i} = \frac{0}{0}$ is set to zero. It's time for some examples.

1.3 Examples

Each example explores four experimental designs: exogenous treatment design for conditional QTE (conventional quantile regression, CQR), AAI instrumental variable design aimed at conditional QTE for the subpopulation of compliers, CH instrumental variable design targeting conditional QTE for the entire (commonly supported) population, and F design targeting marginal QTE for the population with common support. The first example illustrates unconfounded quantile treatment effects as outcomes are independent of treatment. In other words, treatment serves as an instrument and the entire population is composed of compliers. As there are no covariates $X = \iota$.

Example 1 (unconfounded *QTE*) *Suppose the DGP is*

Y	$Y(1)$	$Y(0)$	$TE = Y(1) - Y(0)$	z	ν	X^*	K	$p(x)$	Φ
0	2	0	2	0	0	0	1	0.5	0
2	2	0	2	1	1	0	1	0.5	1
2	4	2	2	0	0	1	1	0.5	0
4	4	2	2	1	1	1	1	0.5	1
3	5	3	2	0	0	2	1	0.5	0
5	5	3	2	1	1	2	1	0.5	1
4	6	4	2	0	0	3	1	0.5	0
6	6	4	2	1	1	3	1	0.5	1
6	8	6	2	0	0	4	1	0.5	0
8	8	6	2	1	1	4	1	0.5	1

Exogenous treatment DGP

$^*X$ *is only utilized for strategy F*

All quantile treatment effects except $\theta = 0.2, 0.4, 0.6, 0.8$ *(where* $Q_\theta[Y(0)]$ *and* $Q_\theta[Y(1)]$ *are not unique) are point identified. Some quantile treatment effects,* α*, along with quantiles for* $Y(0)$*,* β*, are tabulated below. Partially identified (non-unique) quantities are indicated by intervals within which the objective function value is constant and minimized.*[1]

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE(\theta)$
0.1	2	0	2	2
0.2	$(0, 4)$	$(0, 2)$	$(2, 4)$	$(0, 4)$
0.3	2	2	4	2
0.4	$(1, 3)$	$(2, 3)$	$(4, 5)$	$(1, 3)$
0.5	2	3	5	2
0.6	$(1, 3)$	$(3, 4)$	$(5, 6)$	$(1, 3)$
0.7	2	4	6	2
0.8	$(0, 4)$	$(4, 6)$	$(6, 8)$	$(0, 4)$
0.9	2	6	8	2

Quantities identified by all four strategies (CQR, AAI, and F)

$$QTE(\theta) = (\min Q_\theta[Y(1)] - \max Q_\theta[Y(0)], \max Q_\theta[Y(1)] - \min Q_\theta[Y(0)])$$

The CH strategy requires separate treatment as it appears to be sensitive to the discrete nature of the DGP (CH and F are specifically engineered for continuous

[1] Recall, quantiles of discrete distributions can be ambiguous when $F^{-1}(\theta) \neq G^{-1}(1-\theta)$. This occurs at discrete mass points. Let θ_m denote such mass points. Now, define quantile θ_m by the interval $\left(F^{-1}(\theta_m), F^{-1}(\theta_m + \varepsilon)\right)$. Accordingly, the quantile treatment effect is

$$QTE(\theta_m) = \left(F^{-1}_{Y(1)}(\theta_m) - F^{-1}_{Y(0)}(\theta_m + \varepsilon), F^{-1}_{Y(1)}(\theta_m + \varepsilon) - F^{-1}_{Y(0)}(\theta_m)\right)$$

or

$$QTE(\theta) = (\min Q_\theta[Y(1)] - \max Q_\theta[Y(0)], \max Q_\theta[Y(1)] - \min Q_\theta[Y(0)])$$

That is, $QTE(\theta)$ is an interval when $\theta = \theta_m$ and a point when $\theta \neq \theta_m$.

outcomes).

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	γ	$QTE(\theta)$
0.1	2	0	0	2
0.2	2	0	0	$(0,4)$
0.25	2	2	0	2
0.4	2	2	0	$(1,3)$
0.5	2	3	0	2
0.6	$(1,3)$	$(3,4)$	0	$(1,3)$
0.75	2	4	0	2
0.8	4	4	0	$(0,4)$
0.9	2	6	0	2

Quantities identified by the CH strategy for the entire population

While the CH strategy identifies QTE, the intervals deviate from expected. In particular, $\theta = 0.2, 0.4, 0.8$ are expected to be partially identified but are point identified by the CH strategy with $\theta = 0.2, 0.4$ in the middle with $QTE(\theta = 0.2) = 2$ but $\theta = 0.8$ identified at the boundary $QTE(\theta = 0.8) = 4$ by CH. Only $QTE(\theta = 0.6) = (1,3)$ matches partially identified expectations for the CH strategy but CH is a point identification strategy (as is F). Outcomes are homogeneous and since $\Pr(z_0 = 1) = 0$, $QTE(\theta)$ for the compliers equals $QTT(\theta)$, the quantile treatment effect for the treated. Likewise, as $\Pr(z_1 = 1) = 1$, $QTE(\theta)$ for the compliers equals $QTUT(\theta)$, the quantile treatment effect for the untreated. This is a case of unconfounded treatment as treatment adopted serves the role of an instrument.

Example 2 (*QTE* for subsample of compliers) *Suppose the DGP is a slight variation of example 1.*

Y	$Y(1)$	$Y(0)$	$TE = Y(1) - Y(0)$	z	ν	X^*	K	$p(x)$	Φ
0	2	0	2	0	0	0	1	0.0	0
0	2	0	2	0	1	0	-1	0.0	0.8
2	4	2	2	0	0	1	1	0.5	0
4	4	2	2	1	1	1	1	0.5	0.8
3	5	3	2	0	0	2	1	0.5	0
5	5	3	2	1	1	2	1	0.5	0.8
4	6	4	2	0	0	3	1	0.5	0
6	6	4	2	1	1	3	1	0.5	0.8
6	8	6	2	0	0	4	1	0.5	0
8	8	6	2	1	1	4	1	0.5	0.8

Homogeneous DGP for complier subpopulation

$^*X$ is only utilized for strategy F

Conventional quantile regression involves exogenous treatment. CQR identified

quantities are

	identified quantities			*target quantities*	
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE^*(\theta)$	$QTE(\theta)$
0.1	4	0	4	2	2
0.2	4	0	4	2	$(0, 4)$
0.25	$(4, 5)$	0	$(4, 5)$	$(1, 3)$	2
0.3	5	0	5	2	2
0.4	3	2	5	2	$(1, 3)$
0.5	$(2, 4)$	$(2, 3)$	$(5, 6)$	$(1, 3)$	2
0.6	3	3	6	2	$(1, 3)$
0.7	2	4	6	2	2
0.75	$(2, 4)$	4	$(6, 8)$	$(0, 4)$	2
0.8	4	4	8	2	$(0, 4)$
0.9	2	6	8	2	2

Quantities identified by the CQR strategy

Even though QTE is homogeneous across quantiles, exogenous treatment poorly identifies QTE at many quantiles. Interestingly, quartiles are non-unique rather than quintiles as the population suggests. Apparently, this reflects quantiles with common support (rows 3 through 10). If CQR utilized covariate X to identify marginal (not conditional) QTE for the subpopulation with common support then CQR would identify the same quantities as AAI and F (reported below). For the AAI strategy, compliers are represented by rows 3 through 10. All quantile treatment effects except $\theta = 0.25, 0.5, 0.75$ (where $Q_\theta[Y(0)]$ and $Q_\theta[Y(1)]$ are not unique) are point identified for the subpopulation of compliers. Some quantile treatment effects for the compliers, α, along with quantiles for $Y(0)$, β, are tabulated below. Partially identified (non-unique) quantities are indicated by intervals. The F strategy requires covariates to identify where common support

is lacking.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE^c(\theta) = mQTE(\theta)$
0.1	2	2	4	2
0.2	2	2	4	2
0.25	$(1,3)$	$(2,3)$	$(4,5)$	$(1,3)$
0.4	2	3	5	2
0.5	$(1,3)$	$(3,4)$	$(5,6)$	$(1,3)$
0.6	2	4	6	2
0.75	$(0,4)$	$(4,6)$	$(6,8)$	$(0,4)$
0.8	2	6	8	2
0.9	2	6	8	2

Quantities identified by the AAI strategy for the subpopulation of compliers including $QTE^c(\theta)$ as well as $mQTE(\theta)$ for the F strategy subpopulation where common support is satisfied

Outcomes are again homogeneous and since $\Pr(z_0 = 1) = 0$*, $QTE(\theta)$ for the compliers equals $QTT(\theta)$, the quantile treatment effect for the treated. The AAI strategy effectively identifies QTE for the various quantiles associated with the subpopulation of compliers. Notice, even though quartiles are uniquely defined for the population that is not the case for the subpopulation of compliers. The CH strategy deviates from AAI or CQR in that ambiguity reflects population quantiles even though lack of common support leads to identification of wide intervals for treatment effect quantiles less than* 0.20*.*

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	γ	$QTE(\theta)$
0.1	$(-\infty, 4)$	0	0	2
0.2	$(0,5)$	0	0	$(0,4)$
0.25	2	2	0	2
0.4	$(1,3)$	$(2,3)$	0	$(1,3)$
0.5	2	3	0	2
0.6	$(0,3)$	$(3,4)$	0	$(1,3)$
0.75	2	4	0	2
0.8	$(0,4)$	$(4,6)$	0	$(0,4)$
0.9	2	6	0	2

Quantities identified by the CH strategy for the entire population

Not surprisingly, no identification strategy is very effective where common support is lacking.

Example 3 (more variation in QTE) *Suppose the DGP involves more variation than example 2.*

Y	$Y(1)$	$Y(0)$	$TE = Y(1) - Y(0)$	z	ν	X^*	K	$p(x)$	Φ
0	2	0	2	0	0	0	1	0.0	0
0	2	0	2	0	1	0	-1	0.0	0.8
2	4	2	2	0	0	1	1	0.5	0
4	4	2	2	1	1	1	1	0.5	0.8
5	5	5	0	0	0	2	1	0.5	0
5	5	5	0	1	1	2	1	0.5	0.8
5	6	5	1	0	0	3	1	0.5	0
6	6	5	1	1	1	3	1	0.5	0.8
6	8	6	2	0	0	4	1	0.5	0
8	8	6	2	1	1	4	1	0.5	0.8

Heterogeneous DGP for complier subpopulation

$^*X$ *is only utilized for strategy F*

Conventional quantile regression involves exogenous treatment. CQR identified quantities are

	identified quantities			*target quantities*	
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE^*(\theta)$	$QTE(\theta)$
0.1	4	0	4	2	2
0.2	4	0	4	2	$(0, 4)$
0.25	$(4, 5)$	0	$(4, 5)$	$(-1, 3)$	2
0.3	5	0	5	0	0
0.4	3	2	5	0	$(-1, 3)$
0.5	$(0, 4)$	$(2, 5)$	$(5, 6)$	$(0, 1)$	0
0.6	1	5	6	1	$(0, 1)$
0.7	1	5	6	1	1
0.75	$(1, 3)$	5	$(6, 8)$	$(0, 3)$	1
0.8	3	5	8	2	$(0, 3)$
0.9	2	6	8	2	2

Quantities identified by the CQR strategy

$QTE^*(\theta)$ *based on overlapping support data*

$QTE(\theta)$ *based on population DGP*

Again, CQR fails to effectively identify QTE. Even though the data are heterogeneous, CQR restricted to the common support subpopulation (as indicated by X) effectively identifies QTE (as is the case for AAI — focusing on the subpopulation of compliers, and F — based on common X support). For the AAI strategy, compliers are represented by rows 3 through 10. Again, all quantile treatment effects except $\theta = 0.25, 0.5, 0.75$ *(where* $Q_\theta[Y(0)]$ *and* $Q_\theta[Y(1)]$ *are not unique) are point identified for the subpopulation of compliers. Some quantile treatment effects for the compliers,* α*, along with quantiles for* $Y(0)$*,* β*, are tabulated below. Partially identified (non-unique) quantities are indicated*

by intervals. Again, the F strategy utilizes covariates to identify where common support is lacking.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE^c(\theta) = mQTE(\theta)$
0.1	2	2	4	2
0.2	2	2	4	2
0.25	$(-1, 3)$	$(2, 5)$	$(4, 5)$	$(-1, 3)$
0.4	0	5	5	0
0.5	$(0, 1)$	5	$(5, 6)$	$(0, 1)$
0.6	1	5	6	1
0.75	$(0, 3)$	$(5, 6)$	$(6, 8)$	$(0, 3)$
0.8	2	6	8	2
0.9	2	6	8	2

Quantities identified by the AAI strategy for the subpopulation of compliers including $QTE^c(\theta)$ as well as $mQTE(\theta)$ for the F strategy where common support is satisfied

Outcomes are heterogeneous but since $\Pr(z_0 = 1) = 0$, $QTE(\theta)$ for the compliers equals $QTT(\theta)$, the quantile treatment effect for the treated. For the CH strategy once again the lack of common support yields exceptionally wide intervals for unsupported quantiles ($\theta \leq 0.2$).

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	γ	$QTE(\theta)$
0.1	$(-\infty, 4)$	0	0	2
0.2	$(0, 5)$	0	0	$(0, 4)$
0.25	2	2	0	2
0.4	$(-1, 3)$	$(2, 5)$	0	$(-1, 3)$
0.5	0	5	0	0
0.6	$(0, 1)$	5	0	$(0, 1)$
0.75	1	5	0	1
0.8	$(0, 3)$	$(5, 6)$	0	$(0, 3)$
0.9	2	6	0	2

Quantities identified by the CH strategy for the entire population

As in example 2, quintiles are non-unique for the CH identification strategy while quartiles are ambiguous for the AAI identification strategy.

Example 4 (AAI identification strategy) *Suppose the DGP is*

Y	$Y(1)$	$Y(0)$	$TE = Y(1) - Y(0)$	z	ν	X^*	K	$p(x)$	Φ
0	2	0	2	0	0	1	1	$\frac{1}{3}$	0.2
2	2	0	2	1	1	1	1	$\frac{1}{3}$	0.8
2	3	2	1	0	0	1	1	$\frac{1}{3}$	0.2
2	3	2	1	0	1	1	-1	$\frac{1}{3}$	0.8
2	3	2	1	0	0	1	1	$\frac{1}{3}$	0.2
3	3	2	1	1	1	1	1	$\frac{1}{3}$	0.8
5	5	2	3	0	0	2	1	0.75	0.2
2	5	2	3	1	1	2	1	0.75	0.8
5	5	2	3	1	0	2	-1	0.75	0.2
5	5	2	3	1	1	2	1	0.75	0.8

Heterogeneous DGP with complier subpopulation

$^*X$ *is only utilized for strategy F*

Conventional quantile regression involves exogenous treatment. CQR identified quantities are

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE(\theta)$
0.1	2	0	2	2
0.2	$(0, 3)$	$(0, 2)$	$(2, 3)$	$(0, 3)$
0.25	1	2	3	1
0.3	1	2	3	1
0.4	1	2	3	1
0.5	3	2	5	1
0.6	3	2	5	$(1, 3)$
0.7	3	2	5	3
0.75	3	2	5	3
0.8	3	2	5	3
0.9	3	2	5	3

Quantities identified by the CQR strategy

CQR fails to effectively identify QTE for the median and mistakenly appears to point identify $\theta = 0.6$. For the AAI strategy, compliers reside in rows 1,2,5,6,7,8, never takers in rows 3,4, always takers in rows 9,10 and no defiers. Hence, AAI effectively identifies QTE for the subpopulation of compliers

(non-uniqueness for the subpopulation of compliers occurs at $\theta = \frac{1}{3}$ and $\frac{2}{3}$).

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE^c(\theta)$
0.1	2	0	2	2
0.2	2	0	2	2
0.25	2	0	2	2
0.4	1	2	3	1
0.5	1	2	3	1
0.6	1	2	3	1
0.75	3	2	5	3
0.8	3	2	5	3
0.9	3	2	5	3

Quantities identified by the AAI strategy for the subpopulation of compliers including $QTE^c(\theta)$

The CH strategy fails as the rank condition is not satisfied. In particular, U_0 and U_1 are not equal identically or in distribution.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	γ	$QTE(\theta)$
0.05	2	0	0	2
0.1	2	0	0	2
0.2	2	0	0	$(0,3)$
0.25	$(0,1)$	2	0	1
0.4	$(0,1)$	2	0	1
0.5	$(1,3)$	2	0	1
0.6	$(1,3)$	2	0	$(1,3)$
0.75	3	2	0	3
0.8	$(-\infty,3)$	$(2,\infty)$	0	3
0.9	$(-\infty,3)$	$(2,\infty)$	0	3
0.95	$(-\infty,3)$	$(2,\infty)$	0	3

Quantities identified by the CH strategy for the entire population

The F strategy fails to effectively identify QTE for $0.2 \leq \theta \leq 0.3$ as ignorability

is not satisfied.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE(\theta)$
0.05	2	0	2	2
0.1	2	0	2	2
0.2	0	2	2	$(0,3)$
0.25	0	2	2	1
0.3	$(0,1)$	2	$(2,3)$	1
0.4	1	2	3	1
0.5	1	2	3	1
0.6	$(1,3)$	2	$(3,5)$	$(1,3)$
0.7	3	2	5	3
0.75	3	2	5	3
0.8	3	2	5	3
0.9	3	2	5	3
0.95	3	2	5	3

Quantities identified by the F strategy

Example 5 (CH identification strategy) *Suppose the DGP is*

Y	$Y(1)$	$Y(0)$	$TE = Y(1) - Y(0)$	z	ν	X^*	K	$p(x)$	Φ
0	2	0	2	0	0	1	1	0.5	0.4
2	2	0	2	1	1	1	1	0.5	0.6
2	3	1	2	0	0	1	1	0.25	0.4
2	3	1	2	0	1	1	-1	0.25	0.6
2	3	1	2	0	0	1	1	0.25	0.4
3	3	1	2	1	1	1	1	0.25	0.6
5	5	3	2	1	0	2	-1	0.75	0.4
2	5	3	2	0	1	2	-1	0.75	0.6
5	5	3	2	1	0	2	-1	0.75	0.4
5	5	3	2	1	1	2	1	0.75	0.6

Heterogeneous DGP for complier subpopulation

$^*X$ *is only utilized for strategy F*

Conventional quantile regression involves exogenous treatment. CQR identified

quantities are

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE(\theta)$
0.1	2	0	2	2
0.2	$(1,3)$	$(0,1)$	$(2,3)$	$(1,3)$
0.25	2	1	3	2
0.3	2	1	3	2
0.4	$(2,4)$	1	$(3,5)$	2
0.5	4	1	5	2
0.6	4	1	5	$(0,4)$
0.7	4	1	5	2
0.75	4	1	5	2
0.8	$(2,4)$	$(1,3)$	5	2
0.9	3	2	5	2

Quantities identified by the CQR strategy

CQR fails to effectively identify QTE for quantiles $0.4 < \theta < 0.8$. *For the AAI strategy, compliers reside in rows 1,2,5,6, never takers in rows 3,4 and always takers in rows 9,10 but defiers are represented by rows 7 and 8 causing AAI identification failure. Surprisingly, AAI failure is not apparent for quantiles except the median appears to be point identified while the DGP indicates an interval for the median QTE.*

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE^c(\theta)$
0.1	2	0	2	2
0.2	2	0	2	2
0.25	2	0	2	2
0.4	2	0	2	2
0.5	2	0	2	$(1,3)$
0.6	2	0	2	2
0.75	2	0	2	2
0.8	2	0	2	2
0.9	2	0	2	2

Quantities identified by the AAI strategy for the subpopulation of compliers including $QTE^c(\theta)$

The CH strategy is effective in identifying QTE as the rank condition is satisfied, however, some intervals differ and/or are wider than expected. Even though, U_0

and U_1 are identically equal the strategy is not very useful for quantiles $\theta > 0.6$.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	γ	$QTE(\theta)$
0.05	2	0	0	2
0.1	2	0	0	2
0.2	2	0	0	$(1,3)$
0.25	$(1,2)$	1	0	2
0.4	$(1,2)$	1	0	2
0.5	$(2,5)$	$(0,1)$	0	2
0.6	$(3,4)$	1	0	$(0,4)$
0.75	$(2,\infty)$	$(1,3)$	0	2
0.8	$(-\infty,\infty)$	$(1,\infty)$	0	2
0.9	$(-\infty,2)$	$(3,\infty)$	0	2
0.95	$(-\infty,2)$	$(3,\infty)$	0	2

Quantities identified by the CH strategy for the entire population

Results for the AAI and CH strategies are difficult to reconcile with the DGP. The F strategy fails to effectively identify all QTE as ignorability is not satisfied.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE(\theta)$
0.05	2	0	2	2
0.1	2	0	2	2
0.2	1	1	2	$(1,3)$
0.25	2	1	3	2
0.3	$(1,2)$	1	$(2,3)$	2
0.4	2	1	3	2
0.5	2	1	3	2
0.6	$(0,4)$	$(1,3)$	$(3,5)$	$(2,5)$
0.7	2	3	5	2
0.75	2	3	5	2
0.8	2	3	5	2
0.9	2	3	5	2
0.95	2	3	5	2

Quantities identified by the F strategy

Remarkably, the AAI and F strategies perform at least as well as the CH strategy even though the DGP suggests the CH strategy would be most effective at identifying QTE. Perhaps, this reflects the idea that ν is a weak instrument (that is, it only weakly predicts treatment Z) along with discrete rather than continuous potential outcomes, whereas in example 1 through 3 the instrument is strongly related to treatment and outcomes are closer to continuous.[2]

[2] Additional experimentation suggests the CH strategy is sensitive to both outcome continuity and quality of instruments.

Example 6 (F identification strategy) *Suppose the DGP is*

Y	$Y(1)$	$Y(0)$	$TE = Y(1) - Y(0)$	z	ν	X^*	K	$p(x)$	Φ
0	2	0	2	0	0	0	1	0.5	0.4
2	2	0	2	1	1	0	1	0.5	0.6
2	3	2	1	0	0	1	1	0.25	0.4
2	3	2	1	0	1	1	-1	0.25	0.6
2	3	2	1	0	0	1	1	0.25	0.4
3	3	2	1	1	1	1	1	0.25	0.6
5	5	2	3	1	0	2	-1	0.75	0.4
2	5	2	3	0	1	2	-1	0.75	0.6
5	5	2	3	1	0	2	-1	0.75	0.4
5	5	2	3	1	1	2	1	0.75	0.6

Heterogeneous DGP for complier subpopulation

$^*X$ is only utilized for strategy F

Conventional quantile regression involves exogenous treatment. CQR identified quantities are

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE(\theta)$
0.1	2	0	2	2
0.2	$(0, 3)$	$(0, 2)$	$(2, 3)$	$(0, 3)$
0.25	1	2	3	1
0.3	1	2	3	1
0.4	1	2	3	1
0.5	3	2	5	1
0.6	3	2	5	$(1, 3)$
0.7	3	2	5	3
0.75	3	2	5	3
0.8	3	2	5	3
0.9	3	2	5	3

Quantities identified by the CQR strategy

CQR fails to effectively identify QTE for the median and mistakenly appears to point identify $\theta = 0.6$. For the AAI strategy, compliers reside in rows 1,2,5,6, never takers in rows 3,4 and always takers in rows 9,10 but defiers are represented by rows 7 and 8 causing AAI identification failure. AAI failure is

apparent for quantiles above the median.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE^c(\theta)$
0.1	2	0	2	2
0.2	2	0	2	2
0.25	2	0	2	2
0.4	2	0	2	2
0.5	$(0,3)$	0	$(0,3)$	$(0,3)$
0.6	2	0	2	1
0.75	2	0	2	1
0.8	2	0	2	1
0.9	2	0	2	1

Quantities identified by the AAI strategy for the subpopulation of compliers including $QTE^c(\theta)$

The CH strategy fails as the rank condition is not satisfied with many quantiles involving very wide intervals. In particular, U_0 *and* U_1 *are not equal identically or in distribution.*

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	γ	$QTE(\theta)$
0.05	2	0	0	2
0.1	2	0	0	2
0.2	2	0	0	$(0,3)$
0.25	$(-\infty,1)$	2	0	1
0.4	$(-\infty,1)$	2	0	1
0.5	$(0,5)$	$(0,2)$	0	1
0.6	$(0,3)$	2	0	$(1,3)$
0.75	$(3,\infty)$	2	0	3
0.8	$(-\infty,\infty)$	$(2,\infty)$	0	3
0.9	$(-\infty,\infty)$	$(2,\infty)$	0	3
0.95	$(-\infty,\infty)$	$(2,\infty)$	0	3

Quantities identified by the CH strategy for the entire population

The F strategy effectively identifies QTE as ignorability is satisfied.

	identified quantities			*target*
θ	$Q_\theta[Y(1)] - Q_\theta[Y(0)]$	$Q_\theta[Y(0)]$	$Q_\theta[Y(1)]$	$QTE(\theta)$
0.05	2	0	2	2
0.1	2	0	2	2
0.2	(0, 3)	(0, 2)	(2, 3)	(0, 3)
0.25	1	2	3	1
0.3	1	2	3	1
0.4	1	2	3	1
0.5	3	2	3	1
0.6	(1, 3)	2	(3, 5)	(1, 3)
0.7	3	2	5	3
0.75	3	2	5	3
0.8	3	2	5	3
0.9	3	2	5	3
0.95	3	2	5	3

Quantities identified by the F strategy

Next, we revisit monotone treatment response (MTR) and explore partial identification of QTE.

1.4 *MTR* and partial identification of *QTE*

MTR says if treatment $t > s$, then $y_j(t) > y_j(s)$ for all individuals j. MTR bounds for outcome quantity D that respects stochastic dominance are

$$D[y_0(t)] \leq D[y(t)] \leq D[y_1(t)]$$

where

$$y_{0j}(t) \equiv \begin{array}{ll} y_j & t \geq z_j \\ y_0 & \text{otherwise} \end{array}$$

$$y_{1j}(t) \equiv \begin{array}{ll} y_j & t \leq z_j \\ y_1 & \text{otherwise} \end{array}$$

and z_j is individual j's adopted treatment.

Partial identification bounds for MTR quantiles are

$$\begin{array}{rcl} 0 < \theta \leq \Pr(t < z) & \Rightarrow & y_0 \leq Q_\theta[y(t)] \leq Q_{\lambda_1}(y \mid t \leq z) \\ \Pr(t < z) < \theta \leq \Pr(t \leq z) & \Rightarrow & Q_{\lambda_0}(y \mid t \geq z) \leq Q_\theta[y(t)] \leq Q_{\lambda_1}(y \mid t \leq z) \\ \Pr(t \leq z) < \theta < 1 & \Rightarrow & Q_{\lambda_0}(y \mid t \geq z) \leq Q_\theta[y(t)] \leq y_1 \end{array}$$

where

$$\begin{array}{rcl} \lambda_1 & \equiv & \dfrac{\theta}{\Pr(t \leq z)} \\ \lambda_0 & \equiv & \dfrac{\theta - \Pr(t < z)}{\Pr(t \geq z)} \end{array}$$

In the binary treatment setting, $t = 0, 1$, the *MTR* quantile bounds are

$$\begin{array}{cc} 0 < \theta \leq \Pr(t < z) & \\ t = 0 & \Rightarrow \quad y_0 \leq Q_\theta[y(0)] \leq Q_\theta(y) \\ \Pr(t < z) < \theta \leq \Pr(t \leq z) & \\ t = 0 & \Rightarrow \quad Q_\theta(y \mid z = 0) \leq Q_\theta[y(0)] \leq Q_\theta(y) \\ t = 1 & \Rightarrow \quad Q_\theta(y) \leq Q_\theta[y(1)] \leq Q_\theta(y \mid z = 1) \\ \Pr(t \leq z) < \theta < 1 & \\ t = 1 & \Rightarrow \quad Q_\theta(y) \leq Q_\theta[y(1)] \leq y_1 \end{array}$$

Then, the *MTR* treatment effect for any quantity D that respects stochastic dominance (e.g.,. means or quantiles) has bounds

$$0 \leq D[y(t)] - D[y(s)] \leq D[y_1(t)] - D[y_0(s)]$$

To appreciate this result consider the bounds on the following exhaustive monotone treatment response cases.

$$\begin{array}{lcll} s < t < z_j & \Rightarrow & y_0 \leq y_j(s) \leq y_j(t) \leq y_j & (1) \\ s < t = z_j & \Rightarrow & y_0 \leq y_j(s) \leq y_j(t) = y_j & (2) \\ s < z_j < t & \Rightarrow & y_0 \leq y_j(s) \leq y_j \leq y_j(t) \leq y_1 & (3) \\ s = z_j < t & \Rightarrow & y_j = y_j(s) \leq y_j(t) \leq y_1 & (4) \\ z_j < s < t & \Rightarrow & y_j \leq y_j(s) \leq y_j(t) \leq y_1 & (5) \end{array}$$

For simplicity, consider the implications for quantile treatment effect bounds with binary treatment, $s = 0$ and $t = 1$. Only cases (2) and (4) apply.

Case (2) identifies quantile bounds as

$$0 < \theta \leq \Pr(t < z) \quad \Rightarrow \quad y_0 \leq Q_\theta[y(0)] \leq Q_\theta(y)$$

and

$$0 = \Pr(t < z) < \theta \leq \Pr(t \leq z) \quad \Rightarrow \quad Q_\theta(y) \leq Q_\theta[y(1)] \leq Q_\theta(y \mid z = 1)$$

Hence, the case (2) quantile treatment effect

$$QTE(\theta) = Q_\theta[y(1)] - Q_\theta[y(0)]$$

has bounds

$$0 = Q_\theta(y) - Q_\theta(y) \leq QTE(\theta) \leq Q_\theta(y \mid z = 1) - y_0$$

Case (4) identifies quantile bounds as

$$\Pr(t < z) < \theta \leq \Pr(t \leq z) \quad \Rightarrow \quad Q_\theta(y \mid z = 0) \leq Q_\theta[y(0)] \leq Q_\theta(y)$$

and

$$\Pr(t \leq z) < \theta < 1 \quad \Rightarrow \quad Q_\theta(y) \leq Q_\theta[y(1)] \leq y_1$$

Hence, the quantile treatment effect for case (4)

$$QTE(\theta) = Q_\theta[y(1)] - Q_\theta[y(0)]$$

has bounds

$$0 = Q_\theta(y) - Q_\theta(y) \leq QTE(\theta) \leq y_1 - Q_\theta(y \mid z = 0)$$

1.5 Quantile treatment effects based on the data alone

On the other hand, quantile treatment effects based on the data alone are wider. From section 11.1.2, the θ-quantile bounds are

$$\begin{array}{rl} & r(\theta, x) \\ \leq & Q_\theta(y \mid x) \leq \\ & s(\theta, x) \end{array}$$

where

$$r(\theta, x) = \begin{array}{ll} Q_{\frac{\theta - \Pr(\text{missing} \mid x)}{\Pr(\text{observed} \mid x)}}(y \mid x, \text{observed}) & \text{if } \Pr(\text{missing} \mid x) < \theta \\ y_0 & \text{otherwise} \end{array}$$

and

$$s(\theta, x) = \begin{array}{ll} Q_{\frac{\theta}{\Pr(\text{observed} \mid x)}}(y \mid x, \text{observed}) & \text{if } \Pr(\text{missing} \mid x) < 1 - \theta \\ y_1 & \text{otherwise} \end{array}$$

To illustrate quantile bounds for treatment effects, consider the binary treatment case. Quantile bounds based on the data alone are

$$\begin{array}{rcl} r(\theta, x, 0) & \leq & Q_\theta[y(0) \mid x] \leq s(\theta, x, 0) \\ r(\theta, x, 1) & \leq & Q_\theta[y(1) \mid x] \leq s(\theta, x, 1) \end{array}$$

so that quantile treatment effect, $Q_\theta[y(1) \mid x] - Q_\theta[y(0) \mid x]$, bounds are

$$r(\theta, x, 1) - s(\theta, x, 0) \leq QTE(\theta \mid x) \leq s(\theta, x, 1) - r(\theta, x, 0)$$

where

$$\begin{array}{rl} r(\theta, x, 0) = Q_{\frac{\theta - \Pr(z=1)}{\Pr(z=0)}}(y \mid x, z = 0) & \text{if } \Pr(z = 1 \mid x) < \theta \\ r(\theta, x, 0) = y_0 & \text{otherwise} \\ s(\theta, x, 0) = Q_{\frac{\theta}{\Pr(z=0)}}(y \mid x, z = 0) & \text{if } \Pr(z = 1 \mid x) < 1 - \theta \\ s(\theta, x, 0) = y_1 & \text{otherwise} \\ r(\theta, x, 1) = Q_{\frac{\theta - \Pr(z=0)}{\Pr(z=1)}}(y \mid x, z = 1) & \text{if } \Pr(z = 0 \mid x) < \theta \\ r(\theta, x, 1) = y_0 & \text{otherwise} \\ s(\theta, x, 1) = Q_{\frac{\theta}{\Pr(z=1)}}(y \mid x, z = 1) & \text{if } \Pr(z = 0 \mid x) < 1 - \theta \\ s(\theta, x, 1) = y_1 & \text{otherwise} \end{array}$$

A binary treatment illustration helps illuminate some of the subtleties associated with quantile treatment effect bounds.

1.6 Example

Example 7 (*MTR* bounds for *QTE*) *Suppose the DGP is the same as example 2.*

Y	$Y(1)$	$Y(0)$	$TE = Y(1) - Y(0)$	z	ν
0	2	0	2	0	0
0	2	0	2	0	1
2	4	2	2	0	0
4	4	2	2	1	1
3	5	3	2	0	0
5	5	3	2	1	1
4	6	4	2	0	0
6	6	4	2	1	1
6	8	6	2	0	0
8	8	6	2	1	1

The median and first and third quartile bounds based on MTR are

θ	$Q_\theta^{lower}[y(0)]$	$Q_\theta^{upper}[y(0)]$	$Q_\theta^{lower}[y(1)]$	$Q_\theta^{upper}[y(1)]$
0.2	0	$(0,2)$	$(0,2)$	4
0.25	0	2	2	$(4,5)$
0.4	0	$(3,4)$	$(3,4)$	5
0.5	$(2,3)$	4	4	8
0.6	3	$(4,5)$	$(4,5)$	8
0.75	4	6	6	8
0.8	4	6	6	8

θ	$QTE^{lower}(\theta \mid MTR)$	$QTE^{upper}(\theta \mid MTR)$
0.2	$(0,2) - (0,2) = 0$	$4 - 0 = 4$
0.25	$2 - 2 = 0$	$5 - 0 = 5$
0.4	$(3,4) - (3,4) = 0$	$5 - 0 = 5$
0.5	$4 - 4 = 0$	$8 - 2 = 6$
0.6	$(4,5) - (4,5) = 0$	$8 - 3 = 5$
0.75	$6 - 6 = 0$	$8 - 4 = 4$
0.8	$6 - 6 = 0$	$8 - 4 = 4$

where non-unique quantiles are indicated by intervals. While these bounds may not seem very tight, MTR (in conjunction with the data) always results in informative bounds. Monotone response implies the treatment effect is never negative but the data alone may not rule out negative treatment effects. The data alone

identify the following substantially wider quantile treatment effect bounds.

θ	$QTE^{lower}(\theta \mid data)$	$QTE^{upper}(\theta \mid data)$
0.2	$0-(0,2)=-2$	$(5,6)-0=6$
0.25	$0-2=-2$	$6-0=6$
0.4	$0-(3,4)=-4$	$8-0=8$
0.5	$0-(4,6)=-6$	$8-0=8$
0.6	$(0,4)-(6,8)=-8$	$8-(0,2)=8$
0.75	$5-8=-3$	$8-3=5$
0.8	$6-8=-2$	$8-(3,4)=5$

As indicated in example 2, treatment effects are homogeneously equal to 2 *for all unique quantiles for this DGP. The existence of a binary instrument leads to the following bounds on* $QTE(\theta)$ *for the subpopulation of compliers (quartile treatment effects are not point identified as quartiles are not unique for the DGP conditional on* $z_1 - z_0 = 1$*).*

θ	$QTE^{lower}(\theta \mid z_1 - z_0 = 1)$	$QTE^{upper}(\theta \mid z_1 - z_0 = 1)$
0.2	2	2
0.25	1	3
0.4	2	2
0.5	1	3
0.6	2	2
0.75	0	4
0.8	2	2