Smoothed Estimating Equations for Instrumental Variables Quantile Regression

David M. Kaplan Yixiao Sun

University of Missouri UC San Diego

13 November 2015

Kansas State University Economics Seminar

Outline

Big picture (IV-QR)

- Chernozhukov and Hansen (2005)

Big picture (IV-QR)

- Chernozhukov and Hansen (2005)
- Simplest: continuous Y (e.g. earnings), binary endogenous treatment indicator X (e.g. JTPA training)
- Endogeneity from self-selection into treatment (based on unobservables)

Big picture (IV-QR)

- Chernozhukov and Hansen (2005)
- Simplest: continuous Y (e.g. earnings), binary endogenous treatment indicator X (e.g. JTPA training)
- Endogeneity from self-selection into treatment (based on unobservables)
- Quantile function of potential outcomes: $h(x, q)$ for $x \in \{0, 1\}$, $q \in (0, 1)$

Big picture (IV-QR)

- Chernozhukov and Hansen (2005)
- Simplest: continuous Y (e.g. earnings), binary endogenous treatment indicator X (e.g. JTPA training)
- Endogeneity from self-selection into treatment (based on unobservables)
- Quantile function of potential outcomes: $h(x, q)$ for $x \in \{0, 1\}$, $q \in (0, 1)$
  - $h(1, 0.5)$ is the median treated potential outcome
  - $h(1, 0.5) - h(0, 0.5)$ is the "median treatment effect"
  - $h(1, q) - h(0, q)$ is the q-quantile treatment effect (q-QTE), as in Lehmann (1975); Doksum (1974)
  - $\int_0^1 [h(1, q) - h(0, q)] dq$ is the ATE

Big picture (IV-QR)

- Continuous Y, binary endogenous X
- Quantile function of potential outcomes: $h(x, q)$

Big picture (IV-QR)

- Continuous Y, binary endogenous X
- Quantile function of potential outcomes: $h(x, q)$
- Potential outcomes: $Y_d = h(x, U_d)$ for $d = 0, 1$,
  $U_d \sim \mathrm{Unif}(0, 1)$ is “rank variable” (e.g. “ability”)

Big picture (IV-QR)

- Continuous Y, binary endogenous X
- Quantile function of potential outcomes: $h(x,q)$
- Potential outcomes: $Y_d = h(x, U_d)$ for $d = 0, 1$, $U_d \sim \mathrm{Unif}(0,1)$ is "rank variable" (e.g. "ability")
- Instrument Z (e.g. randomized training offer): $X = g(Z, V)$, $Z \perp\!\!\!\perp (U_1, U_0)$

Big picture (IV-QR)

- Continuous Y, binary endogenous X
- Quantile function of potential outcomes: $h(x, q)$
- Potential outcomes: $Y_d = h(x, U_d)$ for $d = 0, 1$, $U_d \sim \text{Unif}(0, 1)$ is "rank variable" (e.g. "ability")
- Instrument Z (e.g. randomized training offer): $X = g(Z, V)$, $Z \perp\!\!\!\perp (U_1, U_0)$
- $U_1 = U_0$ ("rank invariance")

Big picture (IV-QR)

- Continuous Y, binary endogenous X
- Quantile function of potential outcomes: $h(x,q)$
- Potential outcomes: $Y_d = h(x, U_d)$ for $d = 0, 1$, $U_d \sim \mathrm{Unif}(0,1)$ is "rank variable" (e.g. "ability")
- Instrument Z (e.g. randomized training offer): $X = g(Z,V)$, $Z \perp\!\!\!\perp (U_1, U_0)$
- $U_1 = U_0$ ("rank invariance")
- Or, U_0 and U_1 have same distribution conditional on (X,Z,V) ("rank similarity")
- Observable data: X, Z, and $Y = h(X, U_X)$

Big picture (IV-QR)

- Continuous Y, binary endogenous X
- Quantile function of potential outcomes: $h(x, q)$
- Potential outcomes: $Y_d = h(x, U_d)$ for $d = 0, 1$, $U_d \sim \text{Unif}(0, 1)$ is "rank variable" (e.g. "ability")
- Instrument Z (e.g. randomized training offer): $X = g(Z, V)$, $Z \perp\!\!\!\perp (U_1, U_0)$
- $U_1 = U_0$ ("rank invariance")
- Or, U_0 and U_1 have same distribution conditional on (X, Z, V) ("rank similarity")
- Observable data: X, Z, and $Y = h(X, U_X)$
- If $U_1 = U_0 = U$, then $Y < h(X, q) \Leftrightarrow U < q$; since $U \perp\!\!\!\perp Z$, $P(Y < h(X, q) \mid Z) = P(U < q \mid Z) = P(U < q) = q$; compare: $E(\epsilon \mid Z) = 0$

Big picture (IV-QR)

- Continuous Y, binary endogenous X, instrument Z, rank variable U_d (unobs'd)
- Potential outcomes: $Y_d = h(x, U_d)$ for $d = 0, 1$
- $P(Y < h(X, q) \mid Z) = q$
- Analog of (conditional) moment conditions from IV exogeneity

Big picture (IV-QR)

- Continuous Y, binary endogenous X, instrument Z, rank variable U_d (unobs'd)
- Potential outcomes: $Y_d = h(x, U_d)$ for $d = 0, 1$
- $P(Y < h(X, q) \mid Z) = q$
- Analog of (conditional) moment conditions from IV exogeneity
- Also need analog of rank condition/relevance: e.g., if $Z \perp\!\!\!\perp X$, then $h(\cdot, \cdot)$ is not uniquely determined since the unconditional quantile function $Q_Y(q)$ also satisfies $P(Y < Q_Y(q) \mid Z) = q$
- Generalizes to discrete/continuous/multiple endogenous regressors, additional exogenous regressors

Big picture (estimating equations)

Common: moment conditions or "estimating equations"

(linear mean IV)

(linear IV-QR)

Big picture (estimating equations)

Common: moment conditions or "estimating equations"

$$Y = h(X) + U, E(U \mid Z) = 0, h(X) = X'\beta_0 \implies$$

$$0 = E[Z(Y - X'\beta_0)], \beta_0 = [E(ZX')]^{-1}E(ZY) \quad \text{(linear mean IV)}$$

(linear IV-QR)

Big picture (estimating equations)

Common: moment conditions or "estimating equations"

$$Y = h(X) + U, E(U \mid Z) = 0, h(X) = X'\beta_0 \implies$$
$$0 = E[Z(Y - X'\beta_0)], \beta_0 = [E(ZX')]^{-1}E(ZY) \quad \text{(linear mean IV)}$$
$$q = P(Y < h(X) \mid Z), h(X) = X'\beta_0, P(\cdot) = E(1\{\cdot\}) \implies$$
$$0 = E\Big[Z\big(1\{Y - X'\beta_0 < 0\} - q\big)\Big] \quad \text{(linear IV-QR)}$$

Big picture (estimating equations)

Common: moment conditions or "estimating equations"

$$Y = h(X) + U, E(U \mid Z) = 0, h(X) = X'\beta_0 \implies$$

$$0 = E[Z(Y - X'\beta_0)], \beta_0 = [E(ZX')]^{-1}E(ZY) \quad \text{(linear mean IV)}$$

$$q = P(Y < h(X) \mid Z), h(X) = X'\beta_0, P(\cdot) = E(1\{\cdot\}) \implies$$

$$0 = E\Big[Z\Big(1\{Y - X'\beta_0 < 0\} - q\Big)\Big] \quad \text{(linear IV-QR)}$$

Computational difficulty solving IV-QR sample analog
One (prior) idea: smooth $1\{\cdot\}$

Big picture (estimating equations)

Common: moment conditions or "estimating equations"

$$Y = h(X) + U, E(U \mid Z) = 0, h(X) = X'\beta_0 \implies$$

$$0 = E[Z(Y - X'\beta_0)], \beta_0 = [E(ZX')]^{-1}E(ZY) \quad \text{(linear mean IV)}$$

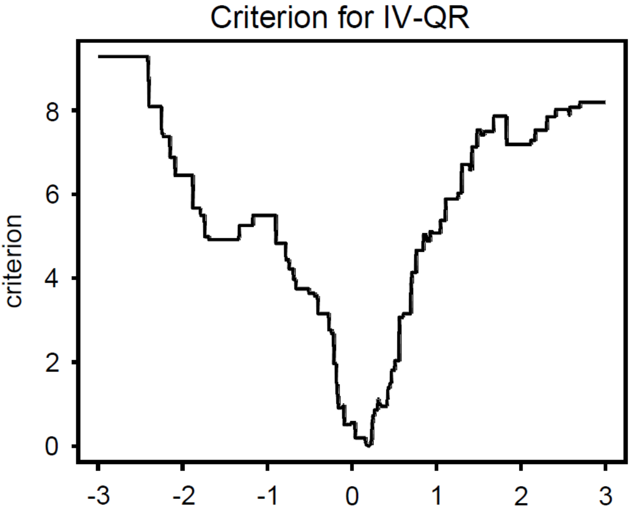
$$q = P(Y < h(X) \mid Z), h(X) = X'\beta_0, P(\cdot) = E(1\{\cdot\}) \implies$$

$$0 = E\Big[Z\big(1\{Y - X'\beta_0 < 0\} - q\big)\Big] \quad \text{(linear IV-QR)}$$

Computational difficulty solving IV-QR sample analog
One (prior) idea: smooth $1\{\cdot\}$
This paper: consequences? optimal amount of smoothing?

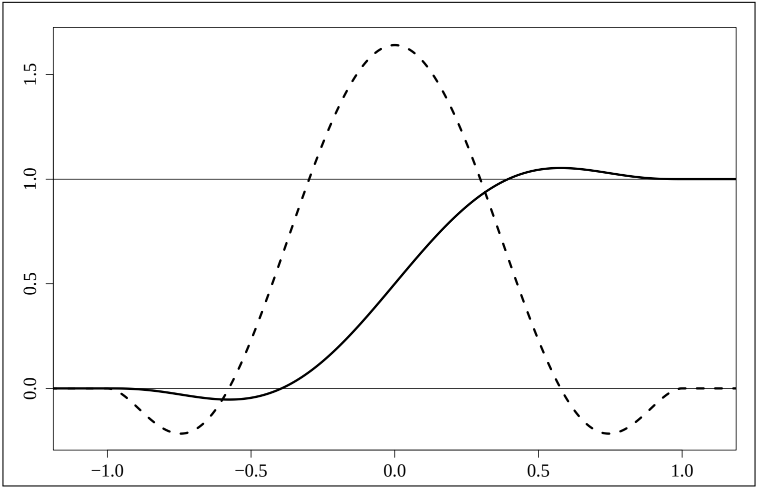
Actual picture

Chernozhukov and Hong (2003), Figure 1(a)



Actual picture

$G(\cdot)$: solid line; $G'(\cdot)$: broken line



Outline

Setup

- Instrumental variables quantile regression (IV-QR):
  - $Y_i = X_i'\beta_0 + U_i$, $P(U_i \le 0 \mid Z_i) = q$,
    $0 = E\Big[Z(1\{Y - X'\beta_0 < 0\} - q)\Big]$
  - $E(Z_i X_i')$ full rank
  - (X_i', Z_i', Y_i) iid across i, $X_i, Z_i \in \mathbb{R}^d$
    (Overidentified: project X onto original $\tilde{Z}$ to get Z)
    (Benefit: know when numerical method returns correct $\hat{\beta}$)

Setup

- Instrumental variables quantile regression (IV-QR):
  - $Y_i = X_i'\beta_0 + U_i$, $P(U_i \le 0 \mid Z_i) = q$,
    $0 = E\Big[Z(1\{Y - X'\beta_0 < 0\} - q)\Big]$
  - $E(Z_i X_i')$ full rank
  - (X_i', Z_i', Y_i) iid across i, $X_i, Z_i \in \mathbb{R}^d$
    (Overidentified: project X onto original $\tilde{Z}$ to get Z)
    (Benefit: know when numerical method returns correct $\hat{\beta}$)
- Smoothed estimating equations (SEE):

$$0 = n^{-1} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i'\hat{\beta} - Y_i}{h} \right) - q \right]$$

  - Same as Whang (2006) (for empirical likelihood)

Setup

- Instrumental variables quantile regression (IV-QR):
  - $Y_i = X_i'\beta_0 + U_i$, $P(U_i \le 0 \mid Z_i) = q$,
    $0 = E\Big[Z(1\{Y - X'\beta_0 < 0\} - q)\Big]$
  - $E(Z_i X_i')$ full rank
  - (X_i', Z_i', Y_i) iid across i, $X_i, Z_i \in \mathbb{R}^d$
    (Overidentified: project X onto original $\tilde{Z}$ to get Z)
    (Benefit: know when numerical method returns correct $\hat{\beta}$)
- Smoothed estimating equations (SEE):

$$0 = n^{-1} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i'\hat{\beta} - Y_i}{h} \right) - q \right]$$

  - Same as Whang (2006) (for empirical likelihood)
  - ☆ $G'(\cdot)$: rth order kernel on $[-1, 1]$; $f_{U|Z}$: $\ge r$ derivatives wrt u

Advantages

$$0 = n^{-1} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \hat{\beta} - Y_i}{h} \right) - q \right] \qquad \text{(IVQR-SEE)}$$

- Technical (Horowitz, 1998)

Advantages

$$0 = n^{-1} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \hat{\beta} - Y_i}{h} \right) - q \right] \qquad \text{(IVQR-SEE)}$$

- Technical (Horowitz, 1998)
- Computational (esp. IV-QR: Chernozhukov and Hansen, 2006)

Advantages

$$0 = n^{-1} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \hat{\beta} - Y_i}{h} \right) - q \right] \qquad \text{(IVQR-SEE)}$$

- Technical (Horowitz, 1998)
- Computational (esp. IV-QR: Chernozhukov and Hansen, 2006)
- Statistical

IV-QR	IVQR-SEE	mean IV
$h = 0$	$h > 0$	$h \to \infty$

Advantages

$$0 = n^{-1} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \hat{\beta} - Y_i}{h} \right) - q \right] \qquad \text{(IVQR-SEE)}$$

- Technical (Horowitz, 1998)
- Computational (esp. IV-QR: Chernozhukov and Hansen, 2006)
- Statistical

IV-QR	IVQR-SEE	mean IV
$h = 0$	$h > 0$	$h \to \infty$

Special case: $X_i = Z_i = 1$, $G'(u) = 1\{-1 \le u \le 1\}$, $q = 0.5 \implies$ Winsorized mean

Advantages

$$0 = n^{-1} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \hat{\beta} - Y_i}{h} \right) - q \right] \qquad \text{(IVQR-SEE)}$$

- Technical (Horowitz, 1998)
- Computational (esp. IV-QR: Chernozhukov and Hansen, 2006)
- Statistical

  IV-QR — IVQR-SEE — mean IV

  $h = 0$ — $h > 0$ — $h \to \infty$

  Special case: $X_i = Z_i = 1$, $G'(u) = 1\{-1 \le u \le 1\}$, $q = 0.5 \implies$ Winsorized mean
- Econometric (MSE, power curves)

Outline

1 Motivation

2 Setup

3 Mean squared error and optimal bandwidth

4 Empirical application (JTPA)

5 Simulations

6 Conclusion

Why MSE of SEE?

- Ultimately, care more about MSE of $\hat{\beta}$
- Large statistics literature on optimal estimating equations leading to optimal point estimation for unbiased EE; here: biased
- Later connect MSE of SEE directly with MSE of $\hat{\beta}$
- SEE: can compute finite-sample bias/variance; $\hat{\beta}$: asy. approx.
- SEE also used for inference; robust to strength of IV (unlike Wald)

MSE of SEE

$$m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \beta_0 - Y_i}{h} \right) - q \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i$$

MSE of SEE

$$m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \beta_0 - Y_i}{h} \right) - q \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i$$

$$E(W_i) = \frac{h^r}{r!} \left(\int G'(v) v^r dv \right) E\left[f_{U|Z}^{(r-1)}(0 \mid Z_j) Z_j \right] + o(h^r)$$

MSE of SEE

$$m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \beta_0 - Y_i}{h} \right) - q \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i$$

$$E(W_i) = \frac{h^r}{r!} \left(\int G'(v) v^r dv \right) E\left[f_{U|Z}^{(r-1)}(0 \mid Z_j) Z_j \right] + o(h^r)$$

$$V \equiv \lim_{n \to \infty} \operatorname{Var}(W_i) = q(1-q) E(Z_i Z_i')$$

MSE of SEE

$$m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \beta_0 - Y_i}{h} \right) - q \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i$$

$$E(W_i) = \frac{h^r}{r!} \left(\int G'(v) v^r dv \right) E\left[f_{U|Z}^{(r-1)}(0 \mid Z_j) Z_j \right] + o(h^r)$$

$$V \equiv \lim_{n \to \infty} \operatorname{Var}(W_i) = q(1-q) E(Z_i Z_i')$$

$$E(W_i W_i') = V - h\left[1 - \int_{-1}^{1} [G(u)]^2 du \right] E\{ f_{U|Z}(0 \mid Z_i) Z_i Z_i' \} + o(h)$$

MSE of SEE

$$m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[G\left(\frac{X_i' \beta_0 - Y_i}{h} \right) - q \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i$$

$$E(W_i) = \frac{h^r}{r!} \left(\int G'(v) v^r dv \right) E\left[f_{U|Z}^{(r-1)}(0 \mid Z_j) Z_j \right] + o(h^r)$$

$$V \equiv \lim_{n\to\infty} \mathrm{Var}(W_i) = q(1-q) E(Z_i Z_i')$$

$$E(W_i W_i') = V - h\left[1 - \int_{-1}^{1} [G(u)]^2 du\right] E\{f_{U|Z}(0 \mid Z_i) Z_i Z_i'\} + o(h)$$

- $\uparrow$ bias $\implies$ $\downdownarrows$ var

MSE of SEE

$$h^*_{\text{SEE}} \equiv \arg\min_h E\{m_n' V^{-1} m_n\}$$

MSE of SEE

$$h^*_{\text{SEE}} \equiv \arg\min_h E\{m_n' V^{-1} m_n\}$$

$$E\{m_n' V^{-1} m_n\} = n^{-1} \sum_{j=1}^{n} \left(E\{W_j' V^{-1} W_j\} + \sum_{i \neq j} E\{W_i V^{-1} W_j\} \right)$$

MSE of SEE

$$h^*_{\text{SEE}} \equiv \arg\min_h E\{m_n' V^{-1} m_n\}$$

$$E\{m_n' V^{-1} m_n\} = n^{-1} \sum_{j=1}^{n} \left(E\{W_j' V^{-1} W_j\} + \sum_{i \neq j} E\{W_i V^{-1} W_j\} \right)$$

$$h^*_{\text{SEE}} = n^{-1/(2r-1)} (\cdot)^{\frac{1}{2r-1}}$$

MSE of SEE

$$h^*_{\text{SEE}} \equiv \arg\min_h E\{m_n' V^{-1} m_n\}$$

$$E\{m_n' V^{-1} m_n\} = n^{-1} \sum_{j=1}^{n} \left(E\{W_j' V^{-1} W_j\} + \sum_{i \neq j} E\{W_i V^{-1} W_j\} \right)$$

$$h^*_{\text{SEE}} = n^{-1/(2r-1)} (\cdot)^{\frac{1}{2r-1}}$$

$$h^*_{\text{SEE}} = \left(\frac{(r!)^2 \left[1 - \int_{-1}^{1} G^2(u) du \right] f_U(0)}{2r(\int G'(v) v^r dv)^2 \left[f_U^{(r-1)}(0) \right]^2} \frac{d}{n} \right)^{\frac{1}{2r-1}} \quad \text{if } U \perp\!\!\!\perp Z$$

Type I error

- χ^2 test: $m_n' V^{-1} m_n \xrightarrow{d} \chi^2_d$, use $\hat{V} = q(1-q)n^{-1} \sum_{i=1}^{n} Z_i Z_i'$

Type I error

- χ^2 test: $m_n' V^{-1} m_n \xrightarrow{d} \chi^2_d$, use $\hat{V} = q(1-q)n^{-1}\sum_{i=1}^{n} Z_i Z_i'$
- $S_n \equiv m_n' \hat{V}^{-1} m_n = S_n^L + O_p(n^{-1/2} + h^2)$
- Characteristic function, Fourier–Stieltjes inverse:

$$P(S_n^L < x) \doteq \mathcal{G}_d(x) + d^{-1}\mathcal{G}_d'(x)x\left\{hE(A'A) - nE(W_j')V_n^{-1}E(W_j)\right\}$$

$\mathcal{G}_d(\cdot)$ is χ^2_d CDF

- $\implies h^*_{\text{SEE}}$ maximizes highlighted term (which is positive)

Type I error

- χ^2 test: $m_n' V^{-1} m_n \xrightarrow{d} \chi^2_d$, use $\hat{V} = q(1-q)n^{-1}\sum_{i=1}^n Z_i Z_i'$
- $S_n \equiv m_n' \hat{V}^{-1} m_n = S_n^L + O_p(n^{-1/2} + h^2)$
- Characteristic function, Fourier–Stieltjes inverse:

$$P(S_n^L < x) \doteq \mathcal{G}_d(x) + d^{-1}\mathcal{G}_d'(x)x\left\{hE(A'A) - nE(W_j')V_n^{-1}E(W_j)\right\}$$

$\mathcal{G}_d(\cdot)$ is χ^2_d CDF

- $\implies h^*_{\text{SEE}}$ maximizes highlighted term (which is positive)
- Why? h^* makes m_n more precise than $h = 0$: lower probability S_n deviates beyond fixed critical value, and smaller cv to achieve same type I error probability, i.e. steeper power curve

MSE of estimator

- Expand $0 = n^{-1/2} m_n(\hat{\beta})$ around $\beta_0 \implies$

$$\sqrt{n}\big(\hat{\beta} - \beta_0\big) = -\left\{E \frac{\partial}{\partial \beta'} m_n / \sqrt{n}\right\}^{-1} m_n + O_p(1/\sqrt{nh}),$$

$$E \frac{\partial}{\partial \beta'} m_n / \sqrt{n} = E\big[Z_j X_j' f_{U|Z,X}(0 \mid Z_j, X_j)\big] + O(h^r)$$

MSE of estimator

- Expand $0 = n^{-1/2} m_n(\hat{\beta})$ around $\beta_0 \implies$

$$\sqrt{n}\big(\hat{\beta} - \beta_0\big) = -\left\{E \frac{\partial}{\partial \beta'} m_n / \sqrt{n}\right\}^{-1} m_n + O_p(1/\sqrt{nh}),$$

$$E \frac{\partial}{\partial \beta'} m_n / \sqrt{n} = E\big[Z_j X_j' f_{U|Z,X}(0 \mid Z_j, X_j)\big] + O(h^r)$$

- If $\exists\, h^*$ universal, then same as h^*_{SEE}...

MSE of estimator

- Expand $0 = n^{-1/2} m_n(\hat{\beta})$ around $\beta_0 \implies$

$$\sqrt{n}\big(\hat{\beta} - \beta_0\big) = -\left\{E \frac{\partial}{\partial \beta'} m_n / \sqrt{n}\right\}^{-1} m_n + O_p(1/\sqrt{nh}),$$

$$E \frac{\partial}{\partial \beta'} m_n / \sqrt{n} = E\big[Z_j X_j' f_{U|Z,X}(0 \mid Z_j, X_j)\big] + O(h^r)$$

- If $\exists\, h^*$ universal, then same as $h^*_{\mathrm{SEE}} \ldots$
- ... but $\nexists$: depends on combination of components

Outline

JTPA: context

- Abadie et al. (2002), 5102 adult men
- Randomized offer of services to individuals (Z_i), around 60% uptake ($P(D_i = 1 \mid Z_i = 1) = 0.62$), other regressors (age, race, etc.)
- Endogeneity from self-selection into treatment; OLS estimate over twice 2SLS
- Y_i: 30-month earnings (US dollars) in "after" period

JTPA: results

		Quantile				
Regressor	Method	0.15	0.25	0.50	0.75	0.85
Training	AAI	121	702	1544	3131	3378
Training	SEE ($\hat{h}$)	57	381	1080	2630	2744
Training	CH	−125	341	385	2557	3137
Training	tiny h	−129	500	381	2760	3114
Training	huge h	1579	1584	1593	1602	1607
Training	2SLS			1593		
Married	AAI	1564	3190	7683	9509	10 185
Married	SEE ($\hat{h}$)	1132	2357	7163	10 174	10 431
Married	CH	504	2396	7722	10 463	10 484
Married	tiny h	504	2358	7696	10 465	10 439
Married	huge h	6611	6624	6647	6670	6683
Married	2SLS			6647		

JTPA: results

Replace age dummies w/ quartic in age; add continuous baseline measures (wage, weekly hrs worked)

Still computes in one second or less; tiny h takes around 10 seconds

JTPA: results

Regressor	Method	Quantile 0.15	0.25	0.50	0.75	0.85
				Original controls		
Training	SEE ($\hat{h}$)	57	381	1080	2630	2744
Training	CH	−125	341	385	2557	3137
Training	tiny h	−129	500	381	2760	3114
Training	2SLS			1593		
				Modified controls		
Training	SEE ($\hat{h}$)	74	398	1045	2748	2974
Training	CH	−20	451	911	2577	3415
Training	tiny h	−50	416	721	2706	3555
Training	huge h	1568	1573	1582	1590	1595
Training	2SLS			1582		

Outline

JTPA DGP 1

- Same variables used in the original analysis in Abadie et al. (2002), drawn roughly from joint distribution in sample
- 1000 simulation replications, $n = 5102$
- "Robust RMSE" is square root of: squared median bias plus $(\text{IQR}/1.35)^2$
  - Matches usual RMSE if normal sampling distribution of estimator

JTPA DGP 1 Results

q	Unsmoothed	SEE ($\hat{h}$)	2SLS
	Robust RMSE		
0.15	2796	2083	1349
0.25	1746	1375	1200
0.50	985	883	922
0.75	866	877	872
0.85	1082	970	927
	Median Bias		
0.15	−238	9	1041
0.25	−122	16	841
0.50	24	−9	341
0.75	−6	−48	−159
0.85	50	−18	−359

JTPA DGP 2

- Additional: second endogenous regressor (and instrument) and four exogenous regressors, all with normal distributions; 20 regressors altogether
- To make the asymptotic bias of 2SLS relatively more important, the sample size is increased to $n = 50\,000$

JTPA DGP 2 Results (binary regressor)

q	$h = 400$	SEE ($\hat{h}$)	2SLS
	Robust RMSE		
0.15	884	735	1035
0.25	550	477	845
0.50	318	310	413
0.75	293	301	356
0.85	384	346	499
	Median Bias		
0.15	−36	11	994
0.25	−19	18	794
0.50	−15	−22	294
0.75	−10	−23	−206
0.85	−16	−17	−406

JTPA DGP 2 Results (continuous regressor)

q	$h = 400$	SEE ($\hat{h}$)	2SLS
	Robust RMSE		
0.15	97	87	107
0.25	103	97	107
0.50	118	111	107
0.75	168	160	107
0.85	209	194	107
	Median Bias		
0.15	−3	−5	−7
0.25	−3	−4	−7
0.50	−6	−10	−7
0.75	−12	−18	−7
0.85	−17	−18	−7

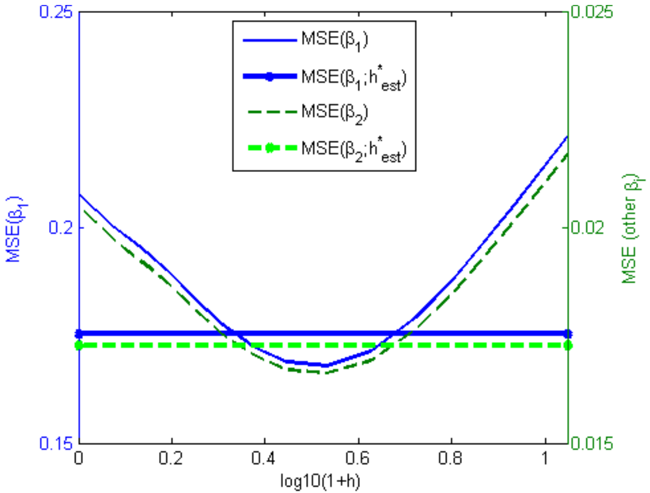
Additional simulations: setup

- h: range incl. $h = 0$
- $\hat{h}_{\mathsf{SEE}}$: $r = 4$, $G(\cdot)$ pictured, $U \perp\!\!\!\perp Z$, $\hat{f}_U(0)$, $\hat{f}_U^{(r-1)}(0)$

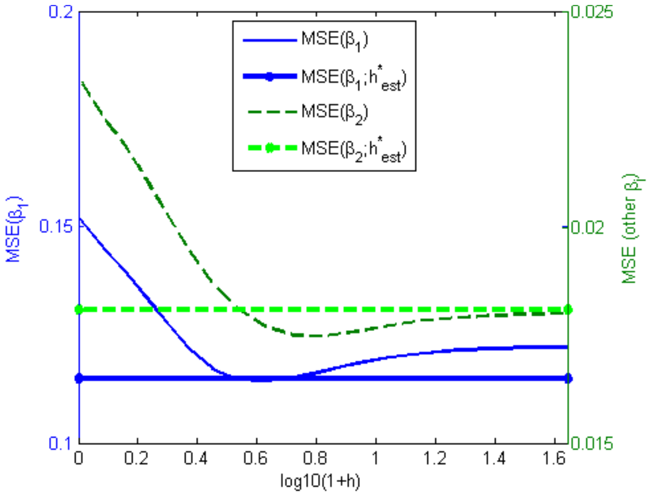
Additional simulations: setup

- h: range incl. $h = 0$
- $\hat{h}_{\mathsf{SEE}}$: $r = 4$, $G(\cdot)$ pictured, $U \perp\!\!\!\perp Z$, $\hat{f}_U(0)$, $\hat{f}_U^{(r-1)}(0)$
- Vary: exog/IV, homo/heteroskedastic, error distribution, q
- (estimator and simulation code at my website)

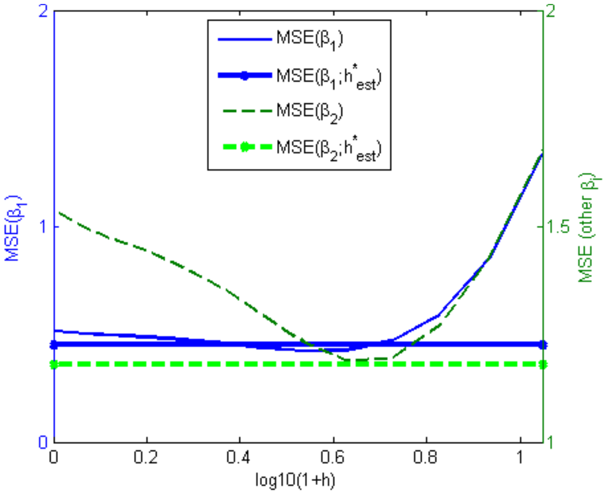
$q = 0.5$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(1,5)$, $U \sim t_3$ w/ $\sigma_U^2 = 2$ (Horowitz, 1998)



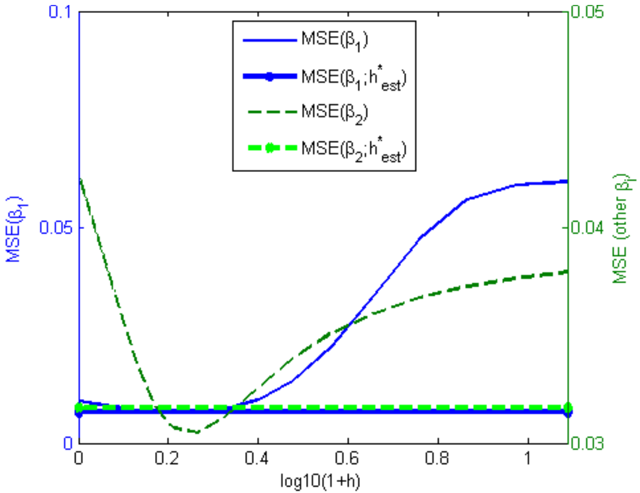
$q = 0.5$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(1,5)$, $U = V(1 + X_{(2)})/4$, $V \sim N(0,1)$ (Horowitz, 1998)



$q = 0.3$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(0,1)$, U ~Cauchy



$q = 0.35$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(0,1)$, $U = (1 + X_{(2)})V$, $V \sim Beta(2,5)$



IV simulations

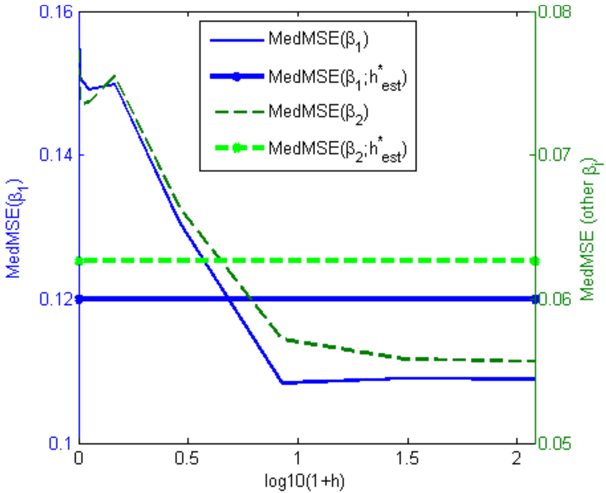
Are there results like Kinal (1980) for (non-)existence of moments of IVQR estimator?

Next graphs show "robust MSE" again:

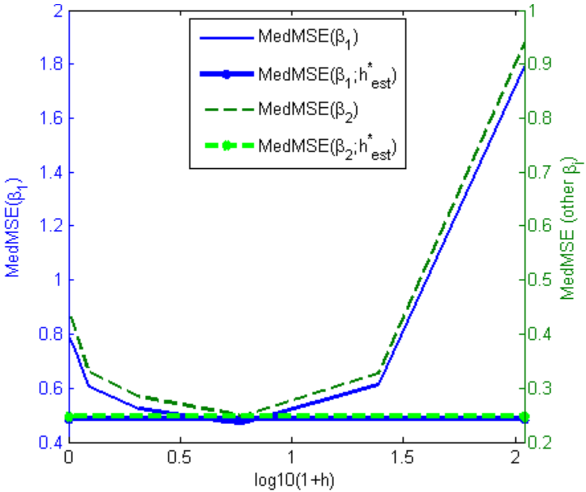
$$\text{MedMSE} = M^2 + (\text{IQR}/1.35)^2,$$

M=median bias, IQR=interquatile range.

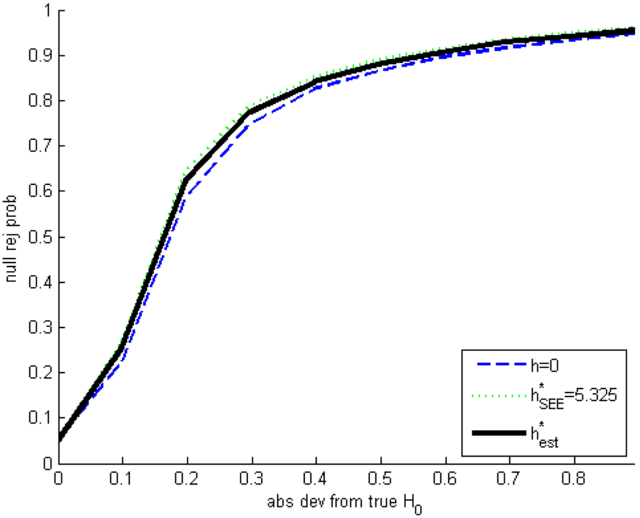
$$y_1 = 1 + 0.5z + v_1,\ (v_1, v_2) \sim (\epsilon_1, \sqrt{1 - 0.5^2}\epsilon_2 + \rho\epsilon_1),$$
$$y_2 = 1 + 0.5z + v_2,\ \epsilon \sim N(0,1),\ q = 0.5,\ n = 100,\ z \sim N(0,1)$$



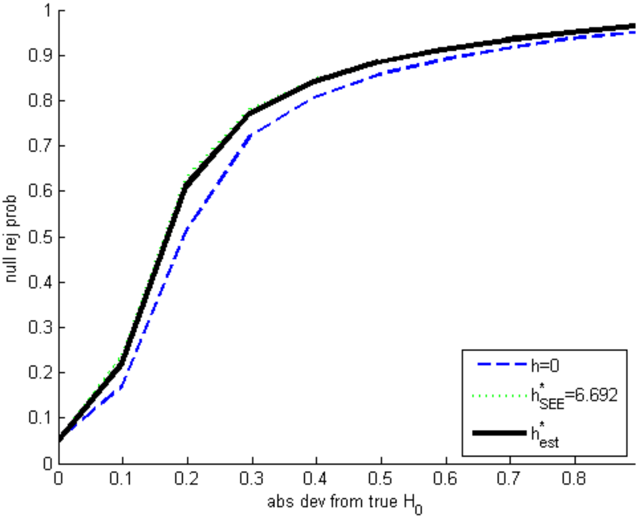
Similar to previous IV but Cauchy errors, $n = 250$



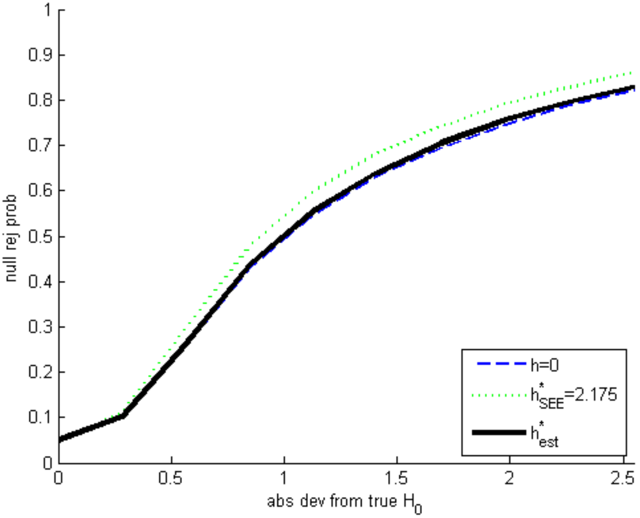
$q = 0.5$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(1,5)$, $U \sim t_3$ w/ $\sigma_U^2 = 2$ (Horowitz, 1998)



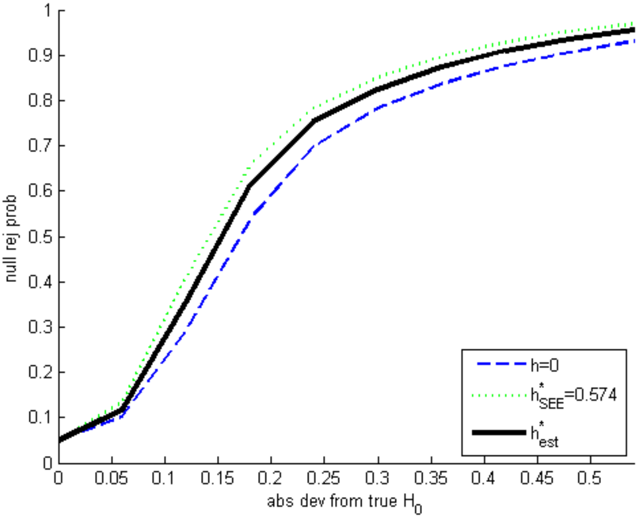
$q = 0.5$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(1,5)$, $U = V(1 + X_{(2)})/4$, $V \sim N(0,1)$ (Horowitz, 1998)



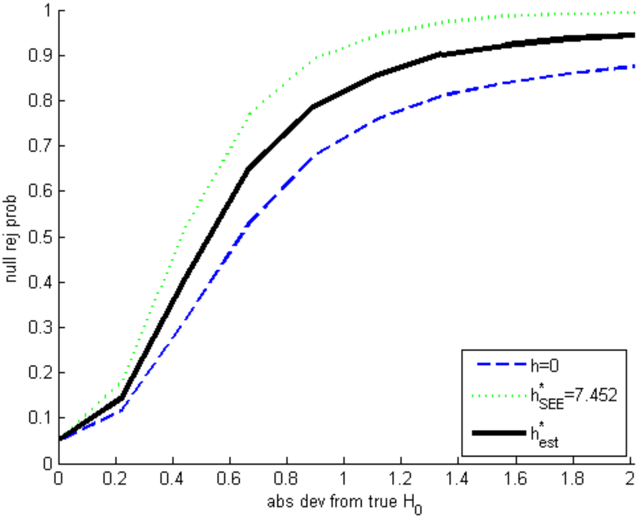
$q = 0.3,\ n = 50,\ \beta_0 = (1, 1)',\ X_{(2)} \sim U(0, 1),\ U$ ~Cauchy



$q = 0.35$, $n = 50$, $\beta_0 = (1, 1)'$, $X_{(2)} \sim U(0, 1)$, $U = (1 + X)V$, $V \sim Beta(2, 5)$



$$y_1 = 1 + 0.5z + v_1,\ (v_1, v_2) \sim (\epsilon_1, \sqrt{1 - 0.5^2}\epsilon_2 + \rho\epsilon_1),$$
$$y_2 = 1 + 0.5z + v_2,\ \epsilon \sim N(0,1),\ q = 0.5,\ n = 100,\ z \sim N(0,1)$$



Outline

Conclusion

- Smooth estimating equations (moment conditions) $\Longrightarrow$
  - $\downarrow$ MSE
  - $\uparrow$ size-adj'd power
  - $\uparrow$ reliable/scalable computation
  - flexible/adaptive between (IV) QR and mean regression
- Thank you!

Conclusion

- Smooth estimating equations (moment conditions) $\Longrightarrow$
  - $\downarrow$ MSE
  - $\uparrow$ size-adj'd power
  - $\uparrow$ reliable/scalable computation
  - flexible/adaptive between (IV) QR and mean regression
- Thank you!
- (And any further questions or comments)

References I

Abadie, A., Angrist, J., and Imbens, G. (2002). Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica*, 70(1):91–117.

Chernozhukov, V. and Hansen, C. B. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Chernozhukov, V. and Hansen, C. B. (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*, 132(2):491–525.

Chernozhukov, V. and Hong, H. (2003). An MCMC approach to classical estimation. *Journal of Econometrics*, 115(2):293–346.

References II

Doksum, K. (1974). Empirical probability plots and statistical inference for nonlinear models in the two-sample case. *Annals of Statistics*, 2(2):267–277.

Horowitz, J. L. (1998). Bootstrap methods for median regression models. *Econometrica*, 66(6):1327–1351.

Kinal, T. W. (1980). The existence of moments of k-class estimators. *Econometrica*, 48(1):241–249.

Lehmann, E. L. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day, San Francisco.

Whang, Y.-J. (2006). Smoothed empirical likelihood methods for quantile regression models. *Econometric Theory*, 22(2):173–205.