Smoothed Estimating Equations for Instrumental Variables Quantile Regression

David M. Kaplan Yixiao Sun

University of Missouri UC San Diego

21 August 2015

Outline

1 Motivation

2 Setup

3 Optimal bandwidth

4 Simulations

5 Conclusion

Big picture

Common: moment conditions or “estimating equations”

Big picture

Common: moment conditions or "estimating equations"

$$0 = E[Z(Y - X'\beta_0)] \qquad \text{(mean IV)}$$

Big picture

Common: moment conditions or "estimating equations"

$$0 = E[Z(Y - X'\beta_0)] \qquad \text{(mean IV)}$$

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

Big picture

Common: moment conditions or "estimating equations"

$$0 = E[Z(Y - X'\beta_0)] \qquad \text{(mean IV)}$$

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

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

Big picture

Common: moment conditions or "estimating equations"

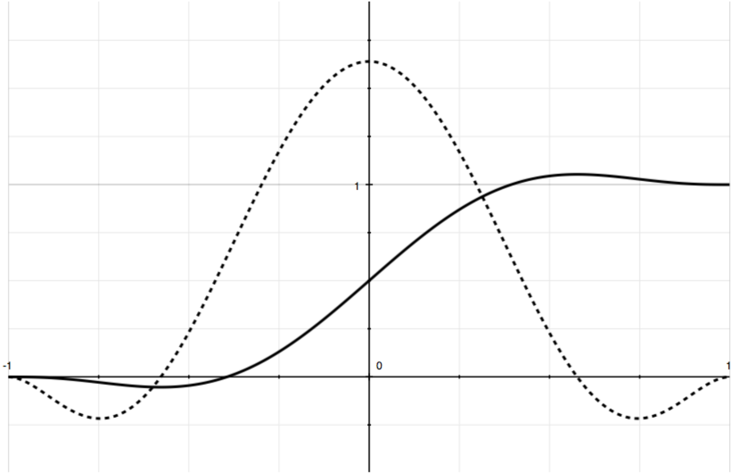
$$0 = E[Z(Y - X'\beta_0)] \qquad \text{(mean IV)}$$

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

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

Actual picture

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



Outline

Setup

- Instrumental variables quantile regression (IV-QR):
  - $Y_i = X_i'\beta + U_i$
  - $P(U_i \le 0 \mid Z_i) = q$
  - $E(Z_i X_i')$ 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):
  - $Y_i = X_i'\beta + U_i$
  - $P(U_i \le 0 \mid Z_i) = q$
  - $E(Z_i X_i')$ 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):

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

  - Whang (2006)

Setup

- Instrumental variables quantile regression (IV-QR):
  - $Y_i = X_i'\beta + U_i$
  - $P(U_i \le 0 \mid Z_i) = q$
  - $E(Z_i X_i')$ 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):

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

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

Advantages

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

- Technical (Horowitz 1998)

Advantages

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

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

Advantages

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

- 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

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

- 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

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

- 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

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

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

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

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

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

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

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

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:

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

- $\implies h^*_{\mathrm{SEE}}$ minimizes highlighted term (which is negative)

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 $h^*_{\mathrm{SEE}}\ldots$

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}...
- ... but $\nexists$: depends on combination of components

Outline

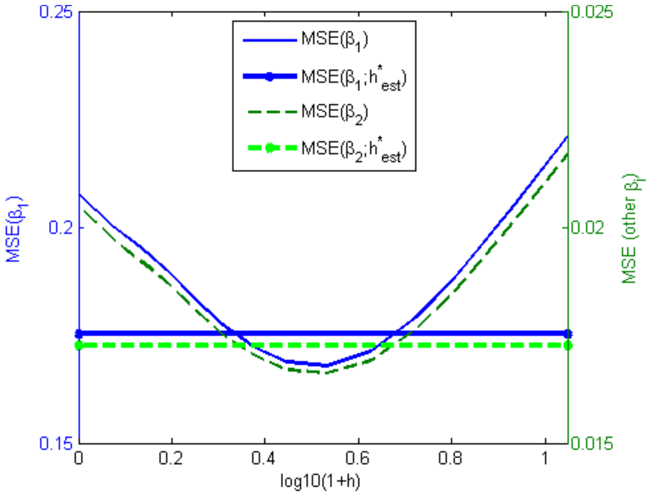
Simulation setup

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

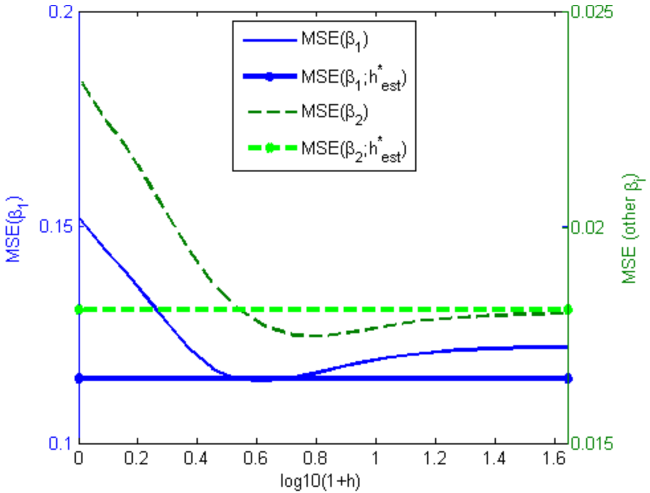
Simulation setup

- 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)

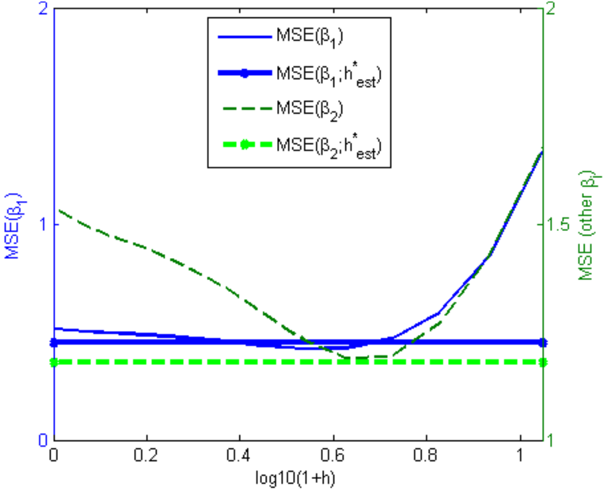
$q = 0.5$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(1,5)$, $U \sim t_3$ w/ $\sigma_U^2 = 2$ (Horowitz 1998)



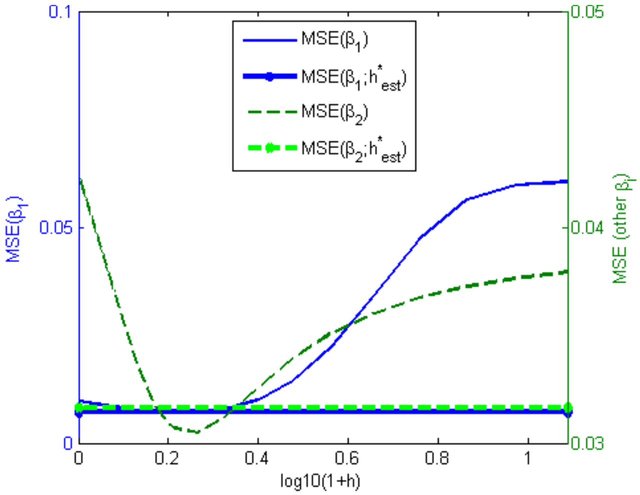
$q = 0.5$, $n = 50$, $\beta_0 = (1, 1)'$, $X_{(2)} \sim U(1, 5)$, $U = V(1 + X_{(2)})/4$, $V \sim N(0, 1)$ (Horowitz 1998)



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

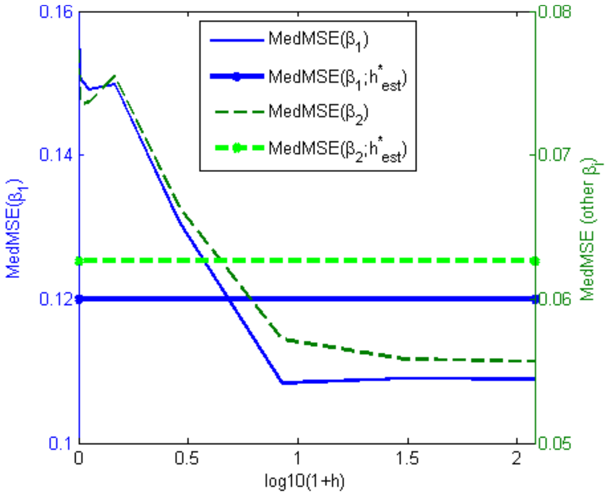
Are there results like Kinal (1980) for (non-)existence of moments of IVQR estimator?

Seems to be similar phenomenon; instead of $\mathrm{MSE} = \mathrm{Bias}^2 + \mathrm{Var}$, next graphs show

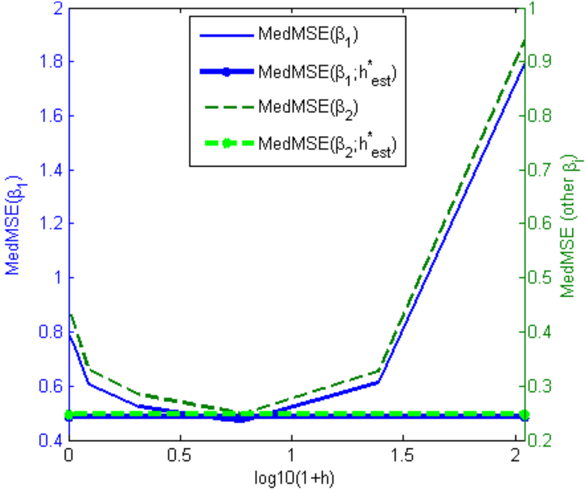
$$\mathrm{MedMSE} = M^2 + (\mathrm{IQR}/1.35)^2,$$

M=median bias, IQR=interquatile range.

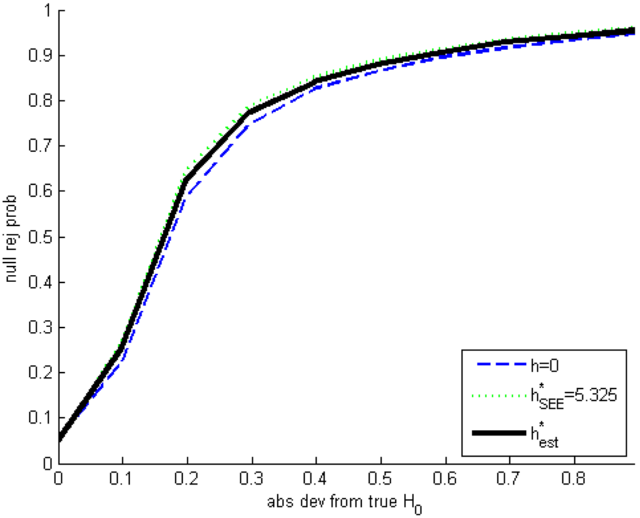
$$y_1 = 1 + 0.5z + v_1,\ (v_1, v_2) \sim (\epsilon_1, \sqrt{1 - 0.5^2}\epsilon_2 + \rho\epsilon_1),$$
$$y_2 = 1 + 0.5z + v_2,\ \epsilon \sim N(0,1),\ q = 0.5,\ n = 100,\ z \sim N(0,1)$$



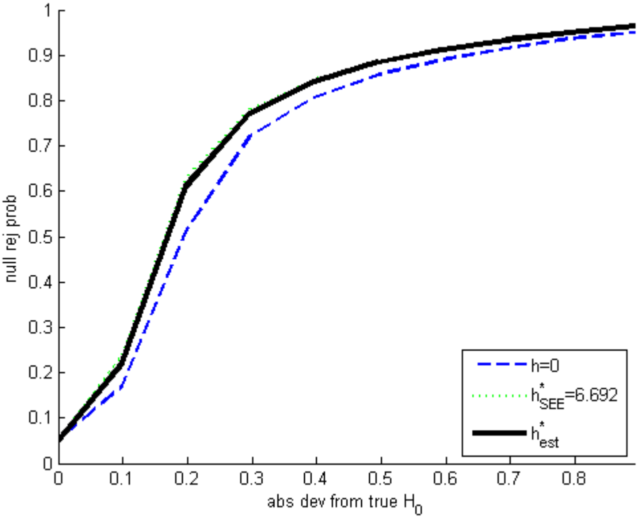
Similar to previous IV but Cauchy errors, $n = 250$



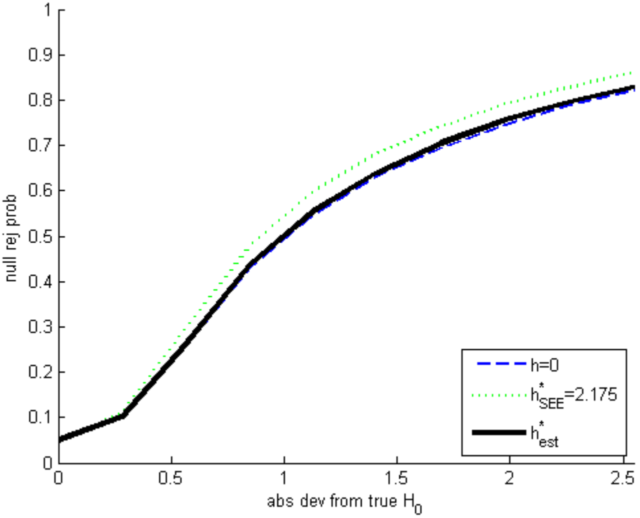
$q = 0.5$, $n = 50$, $\beta_0 = (1,1)'$, $X_{(2)} \sim U(1,5)$, $U \sim t_3$ w/ $\sigma_U^2 = 2$ (Horowitz 1998)



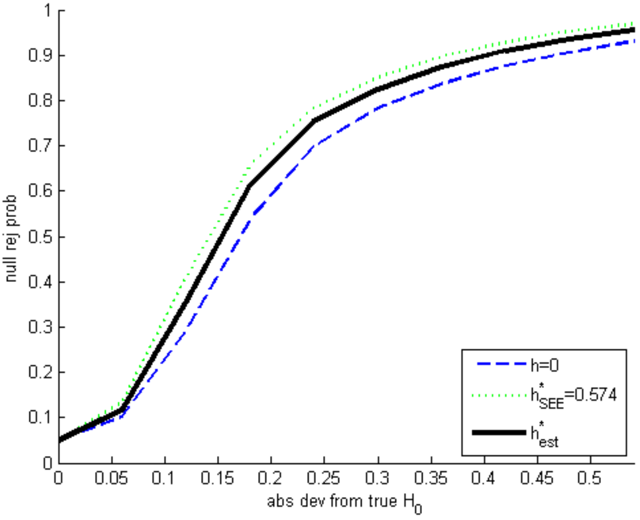
$q = 0.5$, $n = 50$, $\beta_0 = (1, 1)'$, $X_{(2)} \sim U(1, 5)$, $U = V(1 + X_{(2)})/4$, $V \sim N(0, 1)$ (Horowitz 1998)



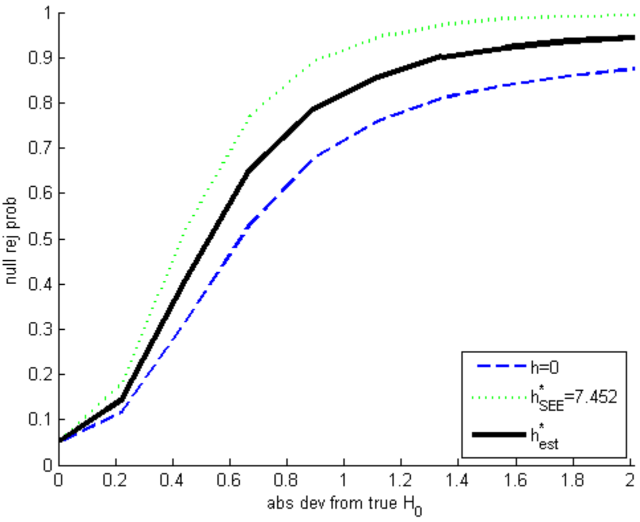
$q = 0.3,\ n = 50,\ \beta_0 = (1,1)',\ X_{(2)} \sim U(0,1),\ U \sim$Cauchy



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



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



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)