Smoothed IV quantile regression and quantile Euler equations

David M. Kaplan
University of Missouri

Coauthors: Luciano de Castro (Iowa), Antonio F. Galvao (Arizona), Xin Liu (Missouri), Yixiao Sun (UCSD)

Outline

Standard consumption Euler equation

- Expected utility maximization, $U(C) = C^{1-\gamma}/(1-\gamma)$:

$$0 = \mathbb{E}[\beta(1+R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t],$$

  Ω_t: information set at time t
  R_{t+1}: real rate of return of asset
  C_t: real consumption at time t
  β: discount factor (e.g., $\beta = 0.99$)
  $1/\gamma$: elasticity of intertemporal substitution (EIS)
- Estimation: use variables in Ω_t as instruments (inflation, etc.); run GMM, or IV/2SLS after log-linearization

Standard consumption Euler equation

- Expected utility maximization, $U(C) = C^{1-\gamma}/(1-\gamma)$:

$$0 = \mathbb{E}[\beta(1+R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t],$$

Ω_t: information set at time t
R_{t+1}: real rate of return of asset
C_t: real consumption at time t
β: discount factor (e.g., $\beta = 0.99$)
$1/\gamma$: elasticity of intertemporal substitution (EIS)

- Estimation: use variables in Ω_t as instruments (inflation, etc.); run GMM, or IV/2SLS after log-linearization
- Drawback: no separation of EIS ($1/\gamma$) and risk aversion (γ)
- Drawback: approximation error from log-linearization

Quantile Euler equation?

- Standard: $0 = \mathbb{E}[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$
- Replace $\mathbb{E}[\cdot \mid \Omega_t]$ with conditional τ-quantile $Q_\tau[\cdot \mid \Omega_t]$:

$$0 = Q_\tau[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$$

Quantile Euler equation?

- Standard: $0 = \mathbb{E}[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$
- Replace $\mathbb{E}[\cdot \mid \Omega_t]$ with conditional τ-quantile $Q_\tau[\cdot \mid \Omega_t]$:

$$0 = Q_\tau[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$$

- Advantage: $1/\gamma$ is EIS, but both τ and γ capture risk attitude
- Advantage: $\ln(Q_\tau(W)) = Q_\tau(\ln(W))$, no error
- Advantage: robust to fat tails in consumption
- Application: economically reasonable estimates even when 2SLS unreasonable

Quantile Euler equation?

- Standard: $0 = \mathbb{E}[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$
- Replace $\mathbb{E}[\cdot \mid \Omega_t]$ with conditional τ-quantile $Q_\tau[\cdot \mid \Omega_t]$:

$$0 = Q_\tau[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$$

- Advantage: $1/\gamma$ is EIS, but both τ and γ capture risk attitude
- Advantage: $\ln(Q_\tau(W)) = Q_\tau(\ln(W))$, no error
- Advantage: robust to fat tails in consumption
- Application: economically reasonable estimates even when 2SLS unreasonable
- Grounded in decision theory? (next slide)
- Practical to estimate? (SIVQR)

Quantile Euler equation: decision theory

- Quantile utility maximization, static setting: Manski (1988), Chambers (2009), and Rostek (2010) (axiomatization)
- Dynamic setting: de Castro and Galvao (2017) show dynamic consistency and derive Euler equation

Quantile Euler equation: estimation

- Can write as $Q_\tau[\epsilon_{t+1} \mid \Omega_t] = 1$, $\epsilon_{t+1} \equiv \beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma}$
- Since $\ln(\cdot)$ is strictly increasing, $Q_\tau[\ln(W)] = \ln(Q_\tau[W])$
- In contrast, $\mathbb{E}[\ln(W)] \leq \ln(\mathbb{E}[W])$ (Jensen's); approx error

Quantile Euler equation: estimation

- Can write as $Q_\tau[\epsilon_{t+1} \mid \Omega_t] = 1$, $\epsilon_{t+1} \equiv \beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma}$

$$\ln(\epsilon_{t+1}) = \ln(\beta) + \ln(1 + R_{t+1}) - \gamma \ln(C_{t+1}/C_t),$$
$$\ln(C_{t+1}/C_t) = \gamma^{-1} \ln(\beta) + \gamma^{-1} \ln(1 + R_{t+1}) - \gamma^{-1} \ln(\epsilon_{t+1})$$

- $\gamma > 0 \implies -\gamma^{-1} \ln(\epsilon)$ strictly ↓ in ϵ:

$$0 = \ln(1) = \ln\big(Q_\tau[\epsilon_{t+1} \mid \Omega_t]\big) = Q_\tau[\ln(\epsilon_{t+1}) \mid \Omega_t]$$
$$= Q_{1-\tau}[-\ln(\epsilon_{t+1}) \mid \Omega_t] = Q_{1-\tau}[-\gamma^{-1} \ln(\epsilon_{t+1}) \mid \Omega_t]$$

- Parameters for τ-quantile maximization correspond to the $1 - \tau$ IV quantile regression of $\ln(C_{t+1}/C_t)$ on a constant and $\ln(1 + R_{t+1})$

IV quantile regression (IVQR)

- So we can just run IVQR; but. . .

IV quantile regression (IVQR)

- So we can just run IVQR; but. . .
- Very difficult to compute IVQR estimator numerically

IV quantile regression (IVQR)

- So we can just run IVQR; but. . .
- Very difficult to compute IVQR estimator numerically
- Chernozhukov and Hansen (2006): iid, linear model, only 1 or 2 endogenous regressors (so can't have many interactions, polynomial terms, etc.)

IV quantile regression (IVQR)

- So we can just run IVQR; but. . .
- Very difficult to compute IVQR estimator numerically
- Chernozhukov and Hansen (2006): iid, linear model, only 1 or 2 endogenous regressors (so can't have many interactions, polynomial terms, etc.)
- Other methods (also iid): compliers/LQTE (Abadie, Angrist, and Imbens, 2002), MCMC (Chernozhukov and Hong, 2003; Lancaster and Jun, 2010), triangular system (Lee, 2007, and others), very slow MIQP (Chen and Lee, 2017)

IV quantile regression (IVQR)

- So we can just run IVQR; but. . .
- Very difficult to compute IVQR estimator numerically
- Kaplan and Sun (2017), smoothing: iid, linear, but fast and allows many endogenous regressors; also high-order MSE improvement, connection to 2SLS

IV quantile regression (IVQR)

- So we can just run IVQR; but. . .
- Very difficult to compute IVQR estimator numerically
- Kaplan and Sun (2017), smoothing: iid, linear, but fast and allows many endogenous regressors; also high-order MSE improvement, connection to 2SLS
- de Castro, Galvao, and Kaplan (2017): dependent data, nonlinear model; fast and robust computation, consistency and asymptotic normality
- New results underway from Xin Liu: smoothed two-step GMM, computation (code) and asymptotic theory

Outline

Smoothed IVQR (SIVQR): benefits

- Approach: smooth the moment conditions (estimating equations)
- For now: use just-identified system for numerical robustness; if over-identified, just take linear combination of instruments

Smoothed IVQR (SIVQR): benefits

- Approach: smooth the moment conditions (estimating equations)
- For now: use just-identified system for numerical robustness; if over-identified, just take linear combination of instruments
- Benefit #1: computation is feasible, fast, scalable (many endogenous regressors), and numerically robust
- Benefit #2: often improves MSE (Kaplan and Sun, 2017)
- Benefit #3: important first step toward true IV quantile GMM (in progress with Xin Liu)

Smoothed QR (not IV): literature

- Horowitz (1998): smooths criterion fn instead of moments; Studentized bootstrap refinement
- Whang (2006): same moment smoothing used here, but for empirical likelihood QR; also in Otsu (2008)
- Fernandes, Guerre, and Horta (2017): kernel-smoothed QR criterion; FOC same as smoothed moments above
- MaCurdy and Hong (1999): original IVQR smoothing? (unpub'd notes)

What does IVQR estimate?

- Chernozhukov and Hansen (2005): identification of (conditional) τ-quantile treatment effects, or "structural quantile effects"

What does IVQR estimate?

- Chernozhukov and Hansen (2005): identification of (conditional) τ-quantile treatment effects, or "structural quantile effects"
- Simplistic example: linear structural random coefficient model, $Y = \mathbf{X}'\beta(U)$, assume $\mathbf{X}'\beta(U)$ monotonic in unobserved $U \sim \mathrm{Unif}(0,1)$
- If Y is wage, U is "ability": $\mathbf{X}'\beta(0.5)$ traces out potential wage outcomes (given different $\mathbf{X}$) for individual with median ability $(\mathrm{P}(U \leq 0.5) = 0.5)$

What does IVQR estimate?

- Chernozhukov and Hansen (2005): identification of (conditional) τ-quantile treatment effects, or "structural quantile effects"
- Simplistic example: linear structural random coefficient model, $Y = \mathbf{X}'\beta(U)$, assume $\mathbf{X}'\beta(U)$ monotonic in unobserved $U \sim \mathrm{Unif}(0,1)$
- If Y is wage, U is "ability": $\mathbf{X}'\beta(0.5)$ traces out potential wage outcomes (given different $\mathbf{X}$) for individual with median ability ($\mathrm{P}(U \le 0.5) = 0.5$)
- If instrument vector $\mathbf{Z} \perp\!\!\!\perp U$, then $\mathrm{P}(Y \le \mathbf{X}'\beta(\tau) \mid \mathbf{Z}) = \mathrm{P}(U \le \tau \mid \mathbf{Z}) = \mathrm{P}(U \le \tau) = \tau$: a conditional quantile restriction on the observables Y, $\mathbf{X}$, and $\mathbf{Z}$, and the parameter vector $\beta(\tau)$

Alternative quantile models with endogeneity

- Triangular system: compared in Chernozhukov, Hansen, and Wüthrich (2017, §2.5); like Chesher (2003), Lee (2007), et al.
- LQTE (like LATE): compared in Melly and Wüthrich (2017, §§5–6); like Abadie et al. (2002), Frölich and Melly (2013), et al.

IVQR moment conditions

(linear IV)

$$\tau = \mathrm{P}(Y \le h(\mathbf{X},\tau) \mid \mathbf{Z})$$

(linear IVQR)

IVQR moment conditions

(linear IV)

$$\tau = \mathrm{P}(Y \leq h(\mathbf{X},\tau) \mid \mathbf{Z}),\ h(\mathbf{X},\tau) = \mathbf{X}'\boldsymbol{\beta}_{0\tau},\ \mathrm{P}(\cdot) = \mathbb{E}(\mathbb{1}\{\cdot\}) \implies$$

$$\mathbf{0} = \mathbb{E}\Big[\mathbf{Z}\Big(\mathbb{1}\Big\{Y - \mathbf{X}'\boldsymbol{\beta}_{0\tau} \leq 0\Big\} - \tau\Big)\Big]$$ (linear IVQR)

IVQR moment conditions

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

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

$$\tau = \mathrm{P}(Y \le h(\mathbf{X},\tau) \mid \mathbf{Z}),\ h(\mathbf{X},\tau) = \mathbf{X}'\beta_{0\tau},\ \mathrm{P}(\cdot) = \mathbb{E}(\mathbb{1}\{\cdot\}) \implies$$

$$\mathbf{0} = \mathbb{E}\left[\mathbf{Z}\left(\mathbb{1}\left\{Y - \mathbf{X}'\beta_{0\tau} \le 0\right\} - \tau\right)\right] \qquad \text{(linear IVQR)}$$

IVQR moment conditions

$$Y = h(\mathbf{X}) + U,\ \mathbb{E}(U \mid \mathbf{Z}) = 0,\ h(\mathbf{X}) = \mathbf{X}'\boldsymbol{\beta}_0 \implies$$

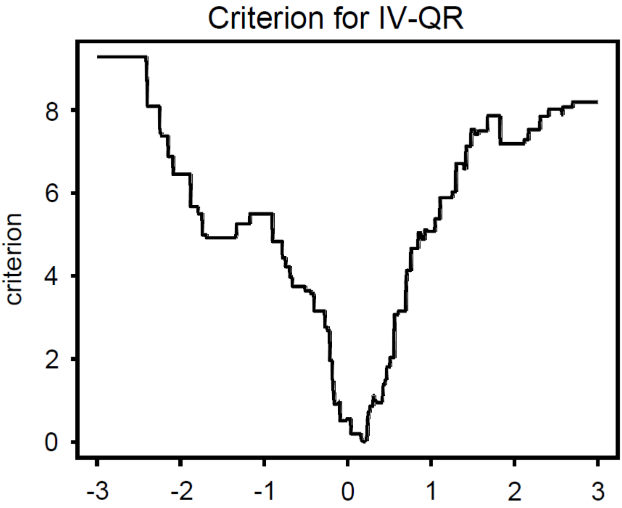
$$\mathbf{0} = \mathbb{E}[\mathbf{Z}(Y - \mathbf{X}'\boldsymbol{\beta}_0)],\ \boldsymbol{\beta}_0 = [\mathbb{E}(\mathbf{Z}\mathbf{X}')]^{-1}\,\mathbb{E}(\mathbf{Z}Y) \qquad \text{(linear IV)}$$

$$\tau = \mathrm{P}(Y \le h(\mathbf{X},\tau) \mid \mathbf{Z}),\ h(\mathbf{X},\tau) = \mathbf{X}'\boldsymbol{\beta}_{0\tau},\ \mathrm{P}(\cdot) = \mathbb{E}(\mathbb{1}\{\cdot\}) \implies$$

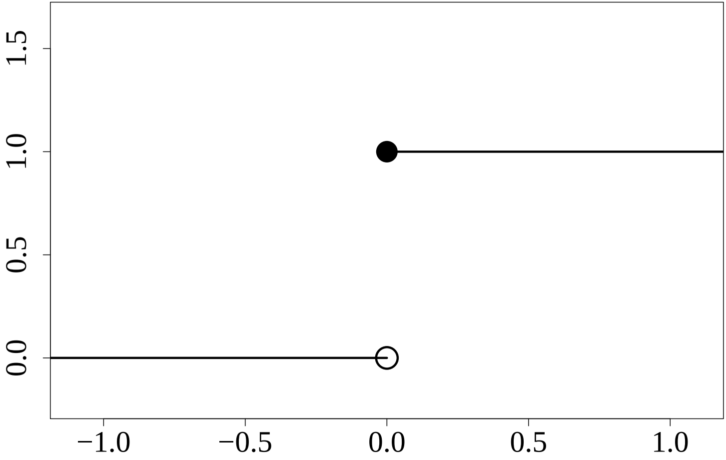
$$\mathbf{0} = \mathbb{E}\left[\mathbf{Z}\left(\mathbb{1}\left\{Y - \mathbf{X}'\boldsymbol{\beta}_{0\tau} \le 0\right\} - \tau\right)\right] \qquad \text{(linear IVQR)}$$

Just run GMM?

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

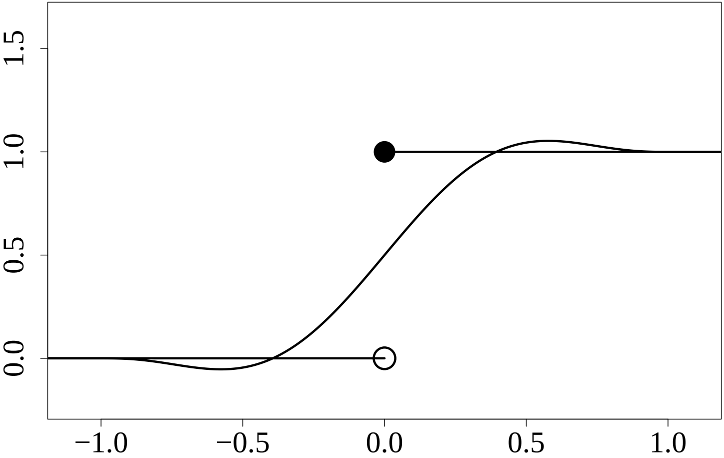


Smoothing the indicator function



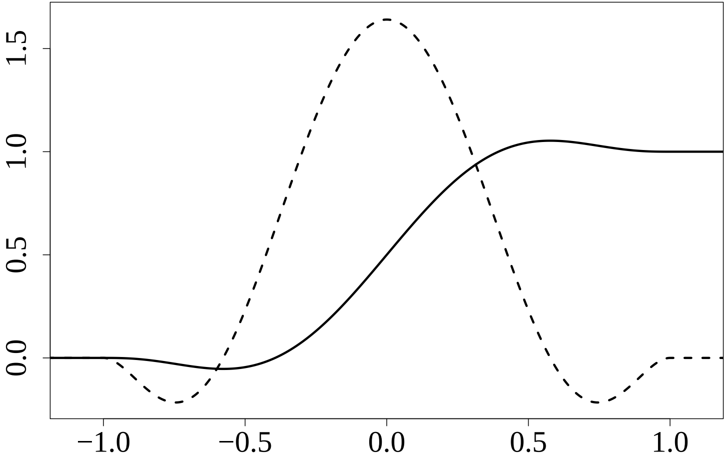
Smoothing the indicator function

$\tilde{I}(\cdot)$ = solid line



Smoothing the indicator function

$\tilde{I}(\cdot)$ = solid line $\tilde{I}'(\cdot)$ = broken line



Smoothed estimator

- Instead of solving sample moments $\mathbf{0} = \hat{\mathbb{E}}\big[\mathbf{Z}(\mathbb{1}\{\mathbf{X}'\hat{\boldsymbol{\beta}}_\tau - Y \geq 0\} - \tau)\big]$, replace $\mathbb{1}\{\cdot \geq 0\}$ with smoothed $\tilde{I}(\cdot/h_n)$, bandwidth h_n:

$$\mathbf{0} = n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i\left[\tilde{I}\left(\frac{\mathbf{X}_i'\hat{\boldsymbol{\beta}}_\tau - Y_i}{h_n}\right) - \tau\right]$$

Smoothed estimator

- Instead of solving sample moments $\mathbf{0} = \hat{\mathbb{E}}\big[\mathbf{Z}(\mathbb{1}\big\{\mathbf{X}'\hat{\boldsymbol{\beta}}_\tau - Y \geq 0\big\} - \tau)\big]$, replace $\mathbb{1}\{\cdot \geq 0\}$ with smoothed $\tilde{I}(\cdot/h_n)$, bandwidth h_n:

$$\mathbf{0} = n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i\left[\tilde{I}\left(\frac{\mathbf{X}_i'\hat{\boldsymbol{\beta}}_\tau - Y_i}{h_n}\right) - \tau\right]$$

- Can compute Jacobian (wrt $\boldsymbol{\beta}$)
- Easy/fast to compute (unless $h_n \approx 0$), standard solver
- $\tilde{I}'(\cdot)$: rth order kernel; $f_{U|\mathbf{Z},\mathbf{X}}(\cdot)$: $\geq r$ derivatives wrt u

Smoothed estimator

- Instead of solving sample moments $\mathbf{0} = \hat{\mathbb{E}}\big[\mathbf{Z}(\mathbb{1}\{\mathbf{X}'\hat{\boldsymbol{\beta}}_\tau - Y \geq 0\} - \tau)\big]$, replace $\mathbb{1}\{\cdot \geq 0\}$ with smoothed $\tilde{I}(\cdot/h_n)$, bandwidth h_n:

$$\mathbf{0} = n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i\left[\tilde{I}\left(\frac{\mathbf{X}_i'\hat{\boldsymbol{\beta}}_\tau - Y_i}{h_n}\right) - \tau\right]$$

- Can compute Jacobian (wrt $\boldsymbol{\beta}$)
- Easy/fast to compute (unless $h_n \approx 0$), standard solver
- $\tilde{I}'(\cdot)$: rth order kernel; $f_{U|\mathbf{Z},\mathbf{X}}(\cdot)$: $\geq r$ derivatives wrt u
- Why exact ID? Robust computation: know when numerical method returns correct $\hat{\boldsymbol{\beta}}$. (If overidentified: can use linear combination of moments, although not efficient.)
- Two-step GMM in progress

Connections: IV, Winsorized mean

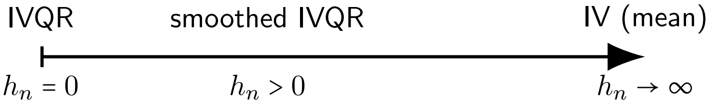
$$\mathbf{0} = n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i\left[\tilde{I}\left(\frac{\mathbf{X}_i'\hat{\boldsymbol{\beta}}_\tau - Y_i}{h_n}\right) - \tau\right]$$

- As $h_n \to \infty$, $\tilde{I}(\cdot/h_n)$ approx linear:

Connections: IV, Winsorized mean

$$\mathbf{0} = n^{-1} \sum_{i=1}^{n} \mathbf{Z}_i \left[\tilde{I}\left(\frac{\mathbf{X}_i' \hat{\boldsymbol{\beta}}_\tau - Y_i}{h_n} \right) - \tau \right]$$

- As $h_n \to \infty$, $\tilde{I}(\cdot / h_n)$ approx linear:



Connections: IV, Winsorized mean

$$\mathbf{0} = n^{-1}\sum_{i=1}^{n}\mathbf{Z}_i\left[\tilde{I}\left(\frac{\mathbf{X}_i'\hat{\boldsymbol{\beta}}_\tau - Y_i}{h_n}\right) - \tau\right]$$

- As $h_n \to \infty$, $\tilde{I}(\cdot/h_n)$ approx linear:

IVQR — smoothed IVQR — IV (mean)

$h_n = 0$ — $h_n > 0$ — $h_n \to \infty$

- Can try to pick h_n to improve efficiency (like median vs. mean); maybe even better to explicitly average IVQR and IV (and QR) like Hansen (2017)?

Connections: IV, Winsorized mean

$$\mathbf{0} = n^{-1} \sum_{i=1}^{n} \mathbf{Z}_i \left[\tilde{I}\left(\frac{\mathbf{X}_i' \hat{\boldsymbol{\beta}}_\tau - Y_i}{h_n} \right) - \tau \right]$$

- As $h_n \to \infty$, $\tilde{I}(\cdot / h_n)$ approx linear:

IVQR — smoothed IVQR — IV (mean)

$h_n = 0$ — $h_n > 0$ — $h_n \to \infty$

- Can try to pick h_n to improve efficiency (like median vs. mean); maybe even better to explicitly average IVQR and IV (and QR) like Hansen (2017)?
- Special case: $X_i = Z_i = 1$, $\tilde{I}'(u) = \mathbb{1}\{-1 \le u \le 1\}/2$, $\tau = 0.5 \implies$ "Winsorized" mean (Huber, 1964, Ex. iii, p. 79)

Outline

1. Consumption Euler equations
2. Smoothed IV quantile regression (SIVQR)
3. **Results**
4. Conclusion

MSE of SEE ("smoothed estimating equations"), iid/linear

$$Y_i = \mathbf{X}_i' \boldsymbol{\beta}_\tau + U_i, \quad \mathrm{P}(U_i \le 0 \mid \mathbf{Z}_i) = \tau$$

- Ultimately, care more about MSE of $\hat{\boldsymbol{\beta}}_\tau$ than MSE of SEE
- Large statistics literature on optimal EE leading to optimal point estimation for unbiased EE; here: biased
- Connection to MSE of $\hat{\boldsymbol{\beta}}_\tau$ (Kaplan and Sun, 2017)
- MSE of SEE: can compute finite-sample bias/variance; $\hat{\boldsymbol{\beta}}$: asy. approx.
- Also useful for inference; robust to weak IV

MSE of SEE

$$\mathbf{m}_n \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{Z}_i \left[\tilde{I}\left(\frac{\mathbf{X}_i' \boldsymbol{\beta}_{0\tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{W}_i$$

MSE of SEE

$$\mathbf{m}_n \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{Z}_i \left[\tilde{I}\left(\frac{\mathbf{X}_i' \boldsymbol{\beta}_{0\tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{W}_i$$

$$\mathbb{E}(\mathbf{W}_i) = \frac{h_n^r}{r!} \left(\int \tilde{I}'(v) v^r \, dv \right) \mathbb{E}\left[f_{U|\mathbf{Z}}^{(r-1)}(0 \mid \mathbf{Z}_j) \mathbf{Z}_j \right] + o(h_n^r)$$

MSE of SEE

$$\mathbf{m}_n \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{Z}_i \left[\tilde{I}\left(\frac{\mathbf{X}_i' \boldsymbol{\beta}_{0\tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{W}_i$$

$$\mathbb{E}(\mathbf{W}_i) = \frac{h_n^r}{r!} \left(\int \tilde{I}'(v) v^r \, dv \right) \mathbb{E}\left[f_{U|\mathbf{Z}}^{(r-1)}(0 \mid \mathbf{Z}_j) \mathbf{Z}_j \right] + o(h_n^r)$$

$$\underline{\mathbf{V}} \equiv \lim_{n \to \infty} \mathrm{Var}(\mathbf{W}_i) = \tau(1-\tau)\, \mathbb{E}(\mathbf{Z}_i \mathbf{Z}_i')$$

MSE of SEE

$$\mathbf{m}_n \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{Z}_i \left[\tilde{I}\left(\frac{\mathbf{X}_i' \boldsymbol{\beta}_{0\tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{W}_i$$

$$\mathbb{E}(\mathbf{W}_i) = \frac{h_n^r}{r!} \left(\int \tilde{I}'(v) v^r \, dv \right) \mathbb{E}\left[f_{U|\mathbf{Z}}^{(r-1)}(0 \mid \mathbf{Z}_j) \mathbf{Z}_j \right] + o(h_n^r)$$

$$\underline{\mathbf{V}} \equiv \lim_{n \to \infty} \operatorname{Var}(\mathbf{W}_i) = \tau(1-\tau)\, \mathbb{E}(\mathbf{Z}_i \mathbf{Z}_i')$$

$$\mathbb{E}(\mathbf{W}_i \mathbf{W}_i') = \underline{\mathbf{V}} - h_n \left[1 - \int_{-1}^{1} [\tilde{I}(u)]^2 \, du \right] \mathbb{E}\{ f_{U|\mathbf{Z}}(0 \mid \mathbf{Z}_i) \mathbf{Z}_i \mathbf{Z}_i' \} + o(h_n)$$

MSE of SEE

$$\mathbf{m}_n \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{Z}_i \left[\tilde{I}\left(\frac{\mathbf{X}_i' \boldsymbol{\beta}_{0\tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} \mathbf{W}_i$$

$$\mathbb{E}(\mathbf{W}_i) = \frac{h_n^r}{r!} \left(\int \tilde{I}'(v) v^r \, dv \right) \mathbb{E}\left[f_{U|\mathbf{Z}}^{(r-1)}(0 \mid \mathbf{Z}_j) \mathbf{Z}_j \right] + o(h_n^r)$$

$$\underline{\mathbf{V}} \equiv \lim_{n \to \infty} \operatorname{Var}(\mathbf{W}_i) = \tau(1-\tau) \, \mathbb{E}(\mathbf{Z}_i \mathbf{Z}_i')$$

$$\mathbb{E}(\mathbf{W}_i \mathbf{W}_i') = \underline{\mathbf{V}} - h_n \left[1 - \int_{-1}^{1} [\tilde{I}(u)]^2 \, du \right] \mathbb{E}\{ f_{U|\mathbf{Z}}(0 \mid \mathbf{Z}_i) \mathbf{Z}_i \mathbf{Z}_i' \} + o(h_n)$$

- $\uparrow$ bias $\Longrightarrow$ $\downdownarrows$ var

MSE of SEE

$$h^* \equiv \arg\min_h \mathbb{E}\{\mathbf{m}_n' \underline{\mathbf{V}}^{-1} \mathbf{m}_n\}$$

MSE of SEE

$$h^* \equiv \arg\min_h \mathbb{E}\{\mathbf{m}_n' \underline{\mathbf{V}}^{-1} \mathbf{m}_n\}$$

$$\mathbb{E}\{\mathbf{m}_n' \underline{\mathbf{V}}^{-1} \mathbf{m}_n\} = n^{-1} \sum_{j=1}^{n} \left(\mathbb{E}\{\mathbf{W}_j' \underline{\mathbf{V}}^{-1} \mathbf{W}_j\} + \sum_{i \neq j} \mathbb{E}\{\mathbf{W}_i \underline{\mathbf{V}}^{-1} \mathbf{W}_j\} \right)$$

MSE of SEE

$$h^* \equiv \arg\min_h \mathbb{E}\{\mathbf{m}_n' \underline{\mathbf{V}}^{-1} \mathbf{m}_n\}$$

$$\mathbb{E}\{\mathbf{m}_n' \underline{\mathbf{V}}^{-1} \mathbf{m}_n\} = n^{-1} \sum_{j=1}^{n} \left(\mathbb{E}\{\mathbf{W}_j' \underline{\mathbf{V}}^{-1} \mathbf{W}_j\} + \sum_{i \neq j} \mathbb{E}\{\mathbf{W}_i \underline{\mathbf{V}}^{-1} \mathbf{W}_j\} \right)$$

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

MSE of SEE

$$h^* \equiv \arg\min_h \mathbb{E}\{\mathbf{m}_n' \underline{\mathbf{V}}^{-1} \mathbf{m}_n\}$$

$$\mathbb{E}\{\mathbf{m}_n' \underline{\mathbf{V}}^{-1} \mathbf{m}_n\} = n^{-1} \sum_{j=1}^{n} \left(\mathbb{E}\{\mathbf{W}_j' \underline{\mathbf{V}}^{-1} \mathbf{W}_j\} + \sum_{i \neq j} \mathbb{E}\{\mathbf{W}_i \underline{\mathbf{V}}^{-1} \mathbf{W}_j\} \right)$$

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

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

Brief comments on proofs in de Castro et al. (2017)

- Direct treatment of smoothed estimator (vs. show within $o_p(n^{-1/2})$ of unsmoothed); triangular array (U)LLN/CLT
- Three high-level assumptions, but primitive conditions with dependent data given in each case (using Andrews, 1987, 1988; Kato, 2012; Wooldridge, 1986)
- Smoothing allows the usual mean-value expansion (of the sample moment conditions) to derive asymptotic normality
- For consistency, can just smooth as little as possible

Theoretical results

- Consistency
- Asymptotic normality
- Inference: Wald test based on normality; but bootstrap works better (not proved theoretically); and neither is robust to weak identification like Andrews and Mikusheva (2016), Chernozhukov, Hansen, and Jansson (2009), and others

Setup

- Now: endogenous vector $\mathbf{Y}$, instruments $\mathbf{Z}$ with subset $\mathbf{X}$
- Residual fn $\Lambda(\mathbf{Y}, \mathbf{X}, \beta)$, like $Y_1 - (Y_2, \mathbf{X}')\beta$

Setup

- Now: endogenous vector $\mathbf{Y}$, instruments $\mathbf{Z}$ with subset $\mathbf{X}$
- Residual fn $\Lambda(\mathbf{Y}, \mathbf{X}, \boldsymbol{\beta})$, like $Y_1 - (Y_2, \mathbf{X}')\boldsymbol{\beta}$

$$\mathbf{0} = \mathbb{E}\Big\{\mathbf{Z}_i\Big[\mathbb{1}\{\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta}_{0\tau}) \le 0\} - \tau\Big]\Big\} = \mathbf{M}(\boldsymbol{\beta}_{0\tau}, \tau)$$

Setup

- Now: endogenous vector $\mathbf{Y}$, instruments $\mathbf{Z}$ with subset $\mathbf{X}$
- Residual fn $\Lambda(\mathbf{Y}, \mathbf{X}, \boldsymbol{\beta})$, like $Y_1 - (Y_2, \mathbf{X}')\boldsymbol{\beta}$

$$\mathbf{0} = \mathbb{E}\big\{\mathbf{Z}_i\big[\mathbb{1}\{\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta}_{0\tau}) \le 0\} - \tau\big]\big\} = \mathbf{M}(\boldsymbol{\beta}_{0\tau}, \tau)$$

- Smoothed estimator:

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\begin{aligned} \mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) &\equiv \mathbf{g}_n(\mathbf{Y}_i, \mathbf{X}_i, \mathbf{Z}_i, \boldsymbol{\beta}, \tau) \\ &\equiv \mathbf{Z}_i\big[\tilde{I}\big(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\big) - \tau\big]. \end{aligned}$$

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i\Big[\tilde{I}\big(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\big) - \tau\Big].$$

Assumption A1

Strictly stationary, weakly dependent data.

Assumption A2

$\Lambda(\cdot)$ *known, differentiable in* $\boldsymbol{\beta}$*.*

Assumption A3

Global point identification of $\boldsymbol{\beta}_{0\tau}$*; interior of compact* $\mathcal{B}$*.*

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i\Big[\tilde{I}\Big(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\Big) - \tau\Big].$$

Assumption A4

ULLN: $\sup_{\boldsymbol{\beta}\in\mathcal{B}}\Big|\hat{\mathbf{M}}_n(\boldsymbol{\beta}, \tau) - \mathbb{E}[\hat{\mathbf{M}}_n(\boldsymbol{\beta}, \tau)]\Big| = o_p(1)$.

Note $\mathbb{E}[\hat{\mathbf{M}}_n(\boldsymbol{\beta}, \tau)] \neq \mathbf{M}(\boldsymbol{\beta}, \tau)$. Paper: example primitive conditions, using Andrews (1987). Use: $\mathbf{g}_{ni} \leq 2|\mathbf{Z}_i|$ and $h_n \to 0$; WLLN from Andrews (1988).

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i\big[\tilde{I}\big(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\big) - \tau\big].$$

Assumption A5

$\mathbb{E}(\mathbf{Z}_i\mathbf{Z}_i')$ *is positive definite (and finite).*

(No moment restrictions on $\mathbf{Y}_i$.)

Assumption A6

Distribution of $\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})$ *given* $(\boldsymbol{\beta}, \mathbf{Z}_i = \mathbf{z})$ *is continuous at zero.*

E.g., Y_1 cts given $(Y_2, \mathbf{Z})$ for linear IVQR.

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i\big[\tilde{I}\big(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\big) - \tau\big].$$

Assumption A7

$\tilde{I}'(\cdot)$ is kernel fn (bdd support), like picture.

Assumption A8

$h_n = o(n^{-1/4})$.

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i\Big[\tilde{I}\Big(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\Big) - \tau\Big].$$

Assumption A9

Let $\Lambda_i \equiv \Lambda\big(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta}_{0\tau}\big)$ *and* $\mathbf{D}_i \equiv \nabla_{\boldsymbol{\beta}}\Lambda\big(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta}_{0\tau}\big)$. *(i)* $f_{\Lambda|\mathbf{Z}}(\cdot \mid \mathbf{z})$ *twice differentiable (also* $f_{\Lambda|\mathbf{Z},\mathbf{D}}$*). (ii) Nonsingular* $\underline{\mathbf{G}} = \nabla_{\beta}\mathbf{M}(\boldsymbol{\beta}_{0\tau}, \tau) = -\mathbb{E}\{\mathbf{Z}_i\mathbf{D}_i' f_{\Lambda|\mathbf{Z},\mathbf{D}}(0 \mid \mathbf{Z}_i, \mathbf{D}_i)\}$.

E.g., $\mathbf{D}_i$ is regressor vector for linear IVQR.

(ii) $\Longrightarrow$ local identification

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i\left[\tilde{I}\left(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\right) - \tau\right].$$

Assumption A10

$$-\frac{1}{nh_n}\sum_{i=1}^{n} \tilde{I}'\left(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \hat{\boldsymbol{\beta}}_\tau)/h_n\right)\mathbf{Z}_i \nabla_{\boldsymbol{\beta}}\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \hat{\boldsymbol{\beta}}_\tau)' \xrightarrow{p} \underline{\mathbf{G}}.$$

Closely related to Powell (1984, 1991) kernel estimator for QR covariance. Kato (2012): primitive conditions (w/ weakly dependent data) for linear QR ($\mathbf{Y} = Y$, $\mathbf{Z} = \mathbf{X} = \mathbf{D}$). Readily extended to linear IVQR, but harder if non-constant $\nabla_{\boldsymbol{\beta}}\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \hat{\boldsymbol{\beta}}_\tau)$.

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$

$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i\Big[\tilde{I}\Big(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n\Big) - \tau\Big].$$

Assumption A11

Pointwise CLT: $\sqrt{n}\big\{\hat{\mathbf{M}}_n(\boldsymbol{\beta}_{0\tau}, \tau) - \mathbb{E}[\hat{\mathbf{M}}_n(\boldsymbol{\beta}_{0\tau}, \tau)]\big\} \xrightarrow{d} \mathrm{N}(\mathbf{0}, \underline{\boldsymbol{\Sigma}}_\tau)$.

Primitive conditions: moment and dependence restrictions.
Ex: iid, $\mathbb{E}(\|\mathbf{Z}_i\|^2) < \infty$.
Ex: Wooldridge (1986), NED (. . .), $\mathbb{E}(\|\mathbf{Z}_i\|^{2+\epsilon}) < \infty$.

Consistency

Lemma 1

A1–A3 and A5–A8 $\Longrightarrow \sup_{\boldsymbol{\beta}\in\mathcal{B}} \left| \mathbb{E}[\hat{\mathbf{M}}_n(\boldsymbol{\beta},\tau)] - \mathbf{M}(\boldsymbol{\beta},\tau) \right| = o(1)$.

Proof.

Use dominated convergence theorem. Need cts distribution of $\Lambda(\mathbf{Y},\mathbf{X},\boldsymbol{\beta})$ since $\tilde{I}(0) = 0.5 \neq 1 = \mathbb{1}\{0 \geq 0\}$. □

Consistency

Theorem 2

A1–A8 $\implies \hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_{0\tau} = o_p(1)$.

Proof.

Use Thm 5.9 in van der Vaart (1998), or Thm 2.1 in Newey and McFadden (1994). Combine ULLN (A4) with Lemma 1 (and triangle inequality): $\hat{\mathbf{M}}_n(\cdot) \xrightarrow{p} \mathbf{M}(\cdot)$ uniformly. Maximizer of $-\|\mathbf{M}(\cdot)\|$ is uniquely $\boldsymbol{\beta}_{0\tau}$ (A3), "well-separated" b/c compact $\mathcal{B}$ (A3), cts $\mathbf{M}(\cdot)$ (can show). □

Asymptotic normality

Lemma 3

A1–A3, A5, A7–A9, and A11 $\implies \sqrt{n}\hat{\mathbf{M}}_n(\boldsymbol{\beta}_{0\tau}, \tau) \xrightarrow{d} \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_\tau)$.

Proof.

1) $\mathbb{E}[\hat{\mathbf{M}}_n(\boldsymbol{\beta}_{0\tau}, \tau)] = O(h_n^2)$, like kernel bias.
2) $O(\sqrt{n}h_n^2) = o(1)$ if $h_n = o(n^{-1/4})$ (A8).
3) Apply CLT (A11). □

Asymptotic normality

Theorem 4

A1–A11 $\implies \sqrt{n}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_{0\tau}) \xrightarrow{d} \mathrm{N}\big(\mathbf{0}, \underline{\mathbf{G}}^{-1}\underline{\boldsymbol{\Sigma}}_\tau[\underline{\mathbf{G}}']^{-1}\big)$.

Proof.

Mean value expansion: $\mathbf{0} = \hat{\mathbf{M}}_n(\boldsymbol{\beta}_{0\tau}) + \underline{\dot{\mathbf{M}}}_n(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_{0\tau})$, so
$\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_{0\tau}) = -[\underline{\dot{\mathbf{M}}}_n]^{-1}\sqrt{n}\hat{\mathbf{M}}_n(\boldsymbol{\beta}_{0\tau})$.
Apply CMT to Lemma 3 and $\underline{\dot{\mathbf{M}}}_n \xrightarrow{p} \underline{\mathbf{G}}$ (A10). □

Simulation setup

- Compare SIVQR, QR (ignore endogeneity), IV (ignore heterogeneity)
- "JTPA" DGP: iid, binary treatment, randomized offer but self-selection endogeneity
- "TS–IV" DGP: time series regression of y_t on mismeasured x_t, where x_{t-1} is valid IV; normal or Cauchy errors

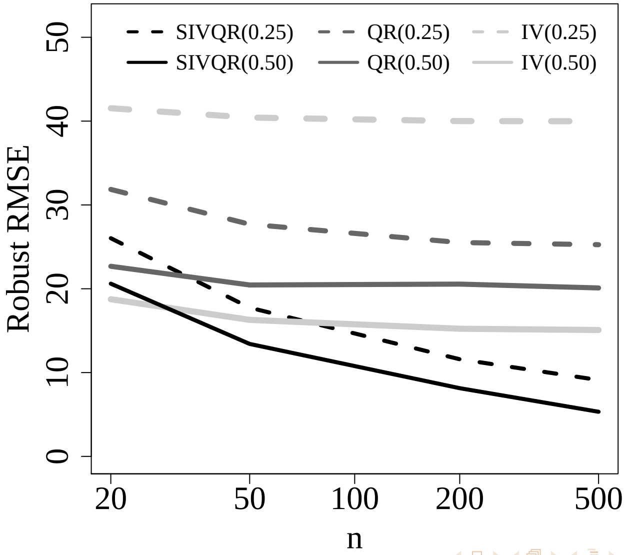
Simulation setup

- Compare SIVQR, QR (ignore endogeneity), IV (ignore heterogeneity)
- "JTPA" DGP: iid, binary treatment, randomized offer but self-selection endogeneity
- "TS–IV" DGP: time series regression of y_t on mismeasured x_t, where x_{t-1} is valid IV; normal or Cauchy errors
- "Robust RMSE": use median bias, and IQR/1.35, so equals RMSE for normal distribution. (IV has no mean. . .)

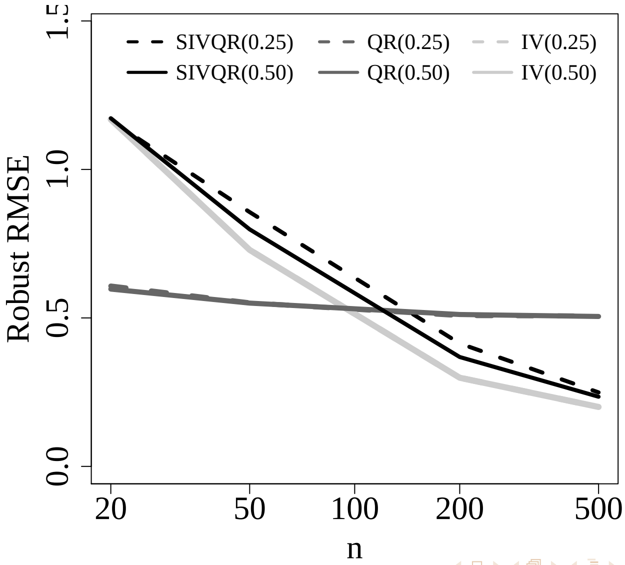
Simulation setup

- Compare SIVQR, QR (ignore endogeneity), IV (ignore heterogeneity)
- "JTPA" DGP: iid, binary treatment, randomized offer but self-selection endogeneity
- "TS–IV" DGP: time series regression of y_t on mismeasured x_t, where x_{t-1} is valid IV; normal or Cauchy errors
- "Robust RMSE": use median bias, and IQR/1.35, so equals RMSE for normal distribution. (IV has no mean. . .)
- Bandwidth h_n: smallest possible for estimation (only second-order effects over wide range); rule of thumb from Kato (2012) for inference
- LRV est: Bartlett kernel, data-dependent bandwidth from Andrews (1991).
- Stationary bootstrap from Politis and Romano (1994).

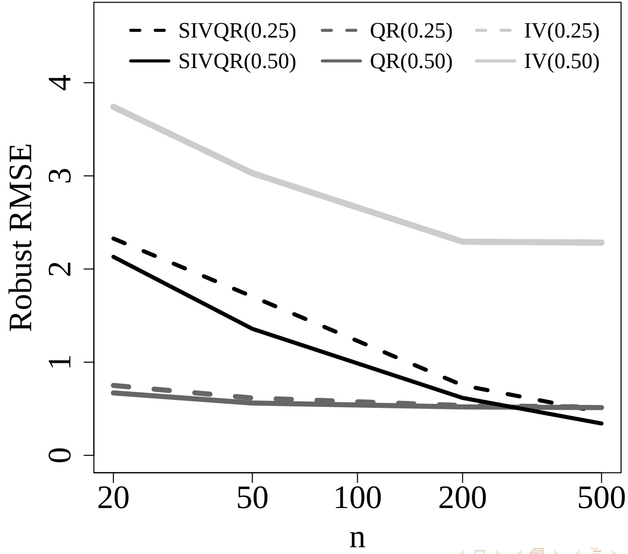
Robust RMSE: JTPA



Robust RMSE: TS–IV, normal



Robust RMSE: TS–IV, Cauchy



Size, 2-sided test of $H_0 : \gamma_\tau = \gamma_0$

DGP	τ	n	α	Wald	BS-t	BS
JTPA	0.25	100	0.10	0.411	0.094	0.196
		1000	0.10	0.415	0.067	0.115
		10,000	0.10	0.226	n/a	n/a
	0.50	100	0.10	0.550	0.074	0.120
		1000	0.10	0.342	0.041	0.101
		10,000	0.10	0.134	n/a	n/a

Size, 2-sided test of $H_0 : \gamma_\tau = \gamma_0$

DGP	τ	n	α	Wald	BS-t	BS
TS–IV.N	0.25	100	0.10	0.269	0.098	0.040
		1000	0.10	0.154	0.107	0.088
		10,000	0.10	0.116	n/a	n/a
	0.50	100	0.10	0.246	0.103	0.040
		1000	0.10	0.147	0.119	0.090
		10,000	0.10	0.101	n/a	n/a

Size, 2-sided test of $H_0 : \gamma_\tau = \gamma_0$

DGP	τ	n	α	Wald	BS-t	BS
TS–IV.C	0.25	100	0.10	0.389	0.059	0.022
		1000	0.10	0.179	0.075	0.059
		10,000	0.10	0.116	n/a	n/a
	0.50	100	0.10	0.296	0.067	0.019
		1000	0.10	0.159	0.110	0.079
		10,000	0.10	0.092	n/a	n/a

JTPA: context

- Abadie, Angrist, and Imbens (2002), 5102 adult men
- Randomized offer of services to individuals (Z_i), 62% uptake: $\mathrm{P}(D_i = 1 \mid Z_i = 1) = 0.62$. Other regressors: age, race, etc.
- Endogeneity from self-selection into treatment; OLS estimate twice as big as IV est
- Y_i: 30-month earnings (US dollars) in "after" period

JTPA: results

		Quantile index τ				
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

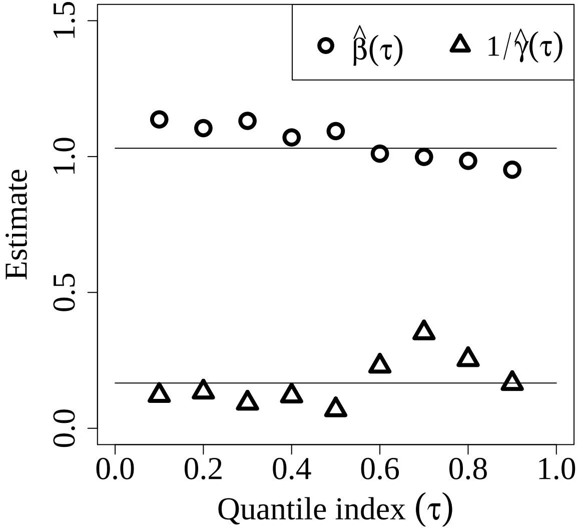
JTPA: results

		Quantile index τ				
Regressor	Method	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		

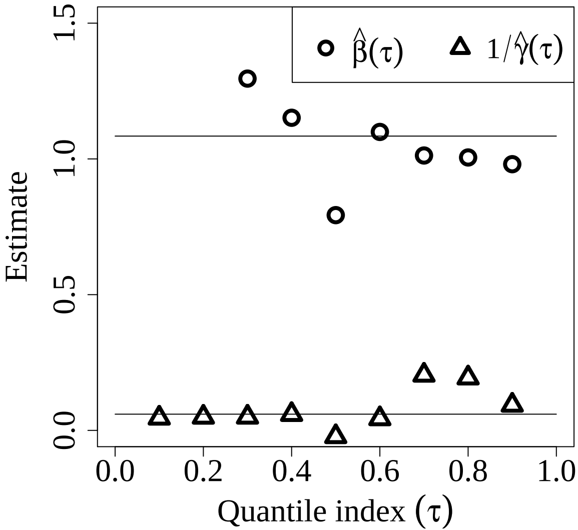
Quantile Euler equation estimates

- Data: from Yogo (2004) (from Campbell, 2003), country-level aggregate time series
- Specification: same as Yogo (2004) Table 2 (but with quantiles): IVQR of $\ln(C_{t+1}/C_t)$ on a constant and $\ln(1 + R_{t+1})$, R =real interest rate
- Excluded instruments are $t-1$ values of: nominal interest rate, inflation, log dividend-price ratio, and $\ln(C_{t-1}/C_{t-2})$; "weak instruments are not a problem" (Yogo, 2004, p. 805)
- Very little smoothing (for estimation)
- β: discount factor ($1 =$ no discount)
- $1/\gamma$: EIS, elasticity of intertemporal substitution

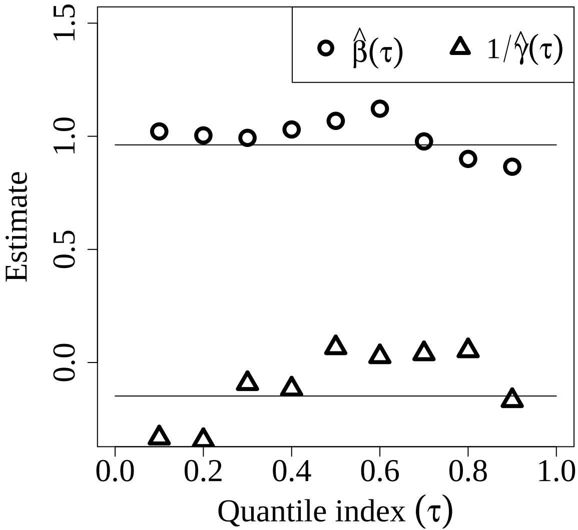
Quantile Euler equation estimates: UK



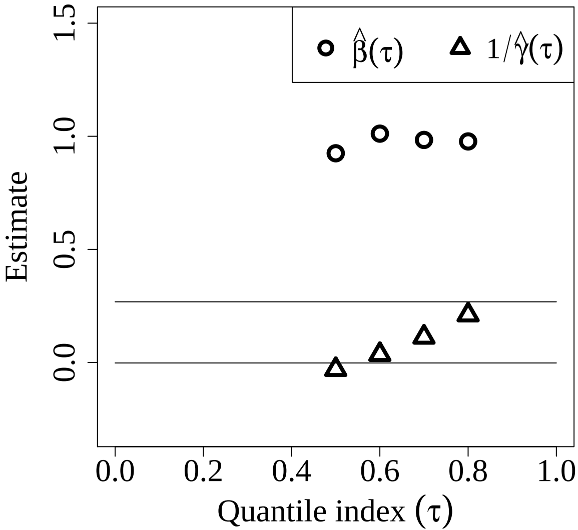
Quantile Euler equation estimates: USA



Quantile Euler equation estimates: Netherlands



Quantile Euler equation estimates: Sweden



Quantile Euler equation estimates

	USA		UK	
τ	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$
0.10	3.18*	22.0*	1.14	8.2*
0.20	1.64	20.5*	1.11	7.5*
0.30	1.30	20.4*	1.13	10.8*
0.40	1.15	17.0*	1.07	8.4*
0.50	0.79	−43.9*	1.09	14.6*
0.60	1.10	23.2*	1.01	4.4*
0.70	1.01	4.9*	1.00	2.8
0.80	1.01	5.2*	0.98	3.9*
0.90	0.98	10.8*	0.95	6.1*
2SLS	1.08	16.7*	1.03	6.0*

*: significantly different from 1 at 10% level (2-sided)

Quantile Euler equation estimates

	US		UK		NTH		SWE	
τ	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$
0.70	1.01	4.9	1.00	2.8	0.98	25.5	0.98	8.9
0.80	1.01	5.2	0.98	3.9	0.90	19.1	0.98	4.7
2SLS	1.08	16.7	1.03	6.0	0.96	−6.8	0.27	−544.4

Outline

Conclusion

- Smoothed IVQR: fast, scalable, robust computation (and better MSE)
- First theoretical results for feasible IVQR with dependent data (and nonlinear model)
- Quantile Euler equations: decouple EIS and risk attitude, robust to fat tails, no error in log-linearization, more reasonable estimates than 2SLS (in our example)
- Many open questions: quantile GMM (in progress)? averaging estimator like in Hansen (2017) or Cheng, Liao, and Shi (2016) (in progress)? semi/nonparametric? uniformity in τ? determination of τ?

Conclusion

- Smoothed IVQR: fast, scalable, robust computation (and better MSE)
- First theoretical results for feasible IVQR with dependent data (and nonlinear model)
- Quantile Euler equations: decouple EIS and risk attitude, robust to fat tails, no error in log-linearization, more reasonable estimates than 2SLS (in our example)
- Many open questions: quantile GMM (in progress)? averaging estimator like in Hansen (2017) or Cheng, Liao, and Shi (2016) (in progress)? semi/nonparametric? uniformity in τ? determination of τ?
- Thank you!
- (And further questions or comments are welcome)

References I

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

Andrews, D. W. K. (1987). Consistency in nonlinear econometric models: A generic uniform law of large numbers. *Econometrica 55*(6), 1465–1471.

Andrews, D. W. K. (1988). Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory 4*(3), 458–467.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*(3), 817–858.

References II

Andrews, I. and A. Mikusheva (2016). Conditional inference with a functional nuisance parameter. *Econometrica 84*(4), 1571–1612.

Campbell, J. Y. (2003). Consumption-based asset pricing. In G. M. Constantinides, M. Harris, and R. M. Stulz (Eds.), *Handbook of the Economics of Finance: Financial Markets and Asset Pricing*, Volume 1, Part B, Chapter 13, pp. 803–887. North Holland.

Chambers, C. P. (2009). An axiomatization of quantiles on the domain of distribution functions. *Mathematical Finance 19*, 335–342.

Chen, L.-Y. and S. Lee (2017). Exact computation of GMM estimators for instrumental variable quantile regression models. Working paper, available at https://arxiv.org/abs/1703.09382.

References III

Cheng, X., Z. Liao, and R. Shi (2016). An averaging GMM estimator robust to misspecification. Working paper, available at http://www.econ.ucla.edu/liao/papers_pdf/ChengLiaoShi_AvgGMM.pdf.

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

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

Chernozhukov, V., C. Hansen, and M. Jansson (2009). Finite sample inference for quantile regression models. *Journal of Econometrics 152*(2), 93–103.

References IV

Chernozhukov, V., C. Hansen, and K. Wüthrich (2017). Instrumental variable quantile regression. In R. Koenker, V. Chernozhukov, X. He, and L. Peng (Eds.), *Handbook of Quantile Regression*, pp. forthcoming. CRC/Chapman-Hall.

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

Chesher, A. (2003). Identification in nonseparable models. *Econometrica 71*(5), 1405–1441.

de Castro, L. and A. F. Galvao (2017). Dynamic quantile models of rational behavior. University of Iowa, mimeo.

de Castro, L., A. F. Galvao, and D. M. Kaplan (2017). Smoothed instrumental variables quantile regression, with estimation of quantile Euler equations. Working paper, available at https://faculty.missouri.edu/~kaplandm.

References V

Fernandes, M., E. Guerre, and E. Horta (2017). Smoothing quantile regressions. Mimeo, available at http://bibliotecadigital.fgv.br/dspace/handle/10438/18390.

Frölich, M. and B. Melly (2013). Unconditional quantile treatment effects under endogeneity. *Journal of Business & Economic Statistics 31*(3), 346–357.

Hansen, B. E. (2017). A Stein-like 2SLS estimator. *Econometric Reviews XX*(X), XXX.

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

Huber, P. J. (1964). Robust estimation of a location parameter. *Annals of Mathematical Statistics 35*(1), 73–101.

References VI

Kaplan, D. M. and Y. Sun (2017). Smoothed estimating equations for instrumental variables quantile regression. *Econometric Theory 33*(1), 105–157.

Kato, K. (2012). Asymptotic normality of Powell's kernel estimator. *Annals of the Institute of Statistical Mathematics 64*(2), 255–273.

Lancaster, T. and S. J. Jun (2010). Bayesian quantile regression methods. *Journal of Applied Econometrics 25*(2), 287–307.

Lee, S. (2007). Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics 141*(2), 1131–1158.

MaCurdy, T. and H. Hong (1999). Smoothed quantile regression in generalized method of moments. mimeo.

References VII

Manski, C. F. (1988). Ordinal utility models of decision making under uncertainty. *Theory and Decision 25*(1), 79–104.

Melly, B. and K. Wüthrich (2017). Local quantile treatment effects. In R. Koenker, V. Chernozhukov, X. He, and L. Peng (Eds.), *Handbook of Quantile Regression*, pp. forthcoming. CRC/Chapman-Hall.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. In R. F. Engle and D. L. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, Chapter 36, pp. 2111–2245. Elsevier.

Otsu, T. (2008). Conditional empirical likelihood estimation and inference for quantile regression models. *Journal of Econometrics 142*(1), 508–538.

References VIII

Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association 89*(428), 1303–1313.

Powell, J. L. (1984). Least absolute deviations estimation for the censored regression model. *Journal of Econometrics 25*(3), 303–325.

Powell, J. L. (1991). Estimation of monotonic regression models under quantile restrictions. In W. A. Barnett, J. Powell, and G. Tauchen (Eds.), *Nonparametric and semiparametric methods in econometrics and statistics: proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, pp. 357–384. Cambridge University Press.

Rostek, M. (2010). Quantile maximization in decision theory. *Review of Economic Studies 77*(1), 339–371.

References IX

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.

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

Wooldridge, J. M. (1986). *Asymptotic Properties of Econometric Estimators*. Ph. D. thesis, Department of Economics, University of California, San Diego.

Yogo, M. (2004). Estimating the elasticity of intertemporal substitution when instruments are weak. *Review of Economics and Statistics 86*(3), 797–810.