

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

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

Standard consumption Euler equation

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

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

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

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

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 (e.g., Carroll, 2001)

Quantile Euler equation?

- Standard: $0 = E[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$
- Replace $E[\cdot \mid \Omega_t]$ 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: $0 = E[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$
- Replace $E[\cdot \mid \Omega_t]$ 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: $0 = E[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 \mid \Omega_t]$
- Replace $E[\cdot \mid \Omega_t]$ 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)
- Two-period: Giovannetti (2013)
- Dynamic setting (∞ -horizon): de Castro and Galvao (2017) show dynamic consistency and derive Euler equation

Quantile Euler equation: estimation

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

Quantile Euler equation: estimation

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

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

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

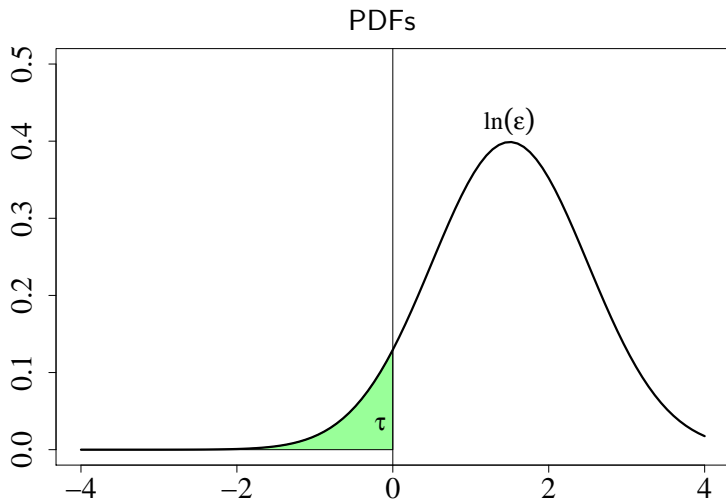
- $\gamma > 0 \implies -\gamma^{-1} \ln(\epsilon)$ strictly \downarrow in ϵ :

$$0 = \ln(1) = \ln(Q_\tau[\epsilon_{t+1} | \Omega_t]) = Q_\tau[\ln(\epsilon_{t+1}) | \Omega_t]$$

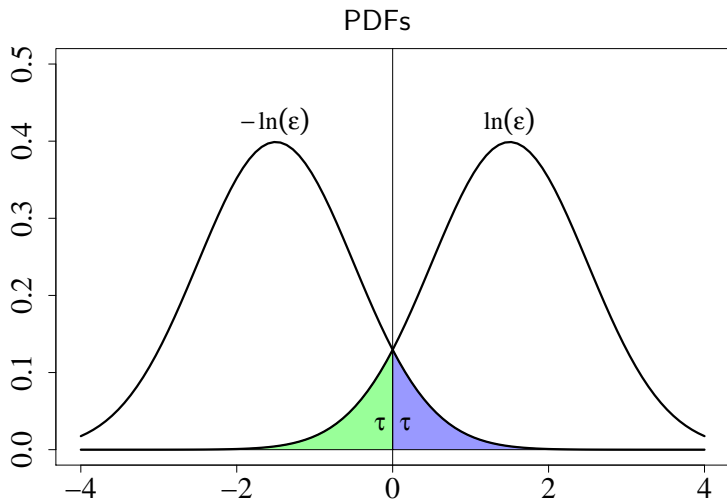
$$= Q_{1-\tau}[-\ln(\epsilon_{t+1}) | \Omega_t] = Q_{1-\tau}[-\gamma^{-1} \ln(\epsilon_{t+1}) | \Omega_t]$$

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

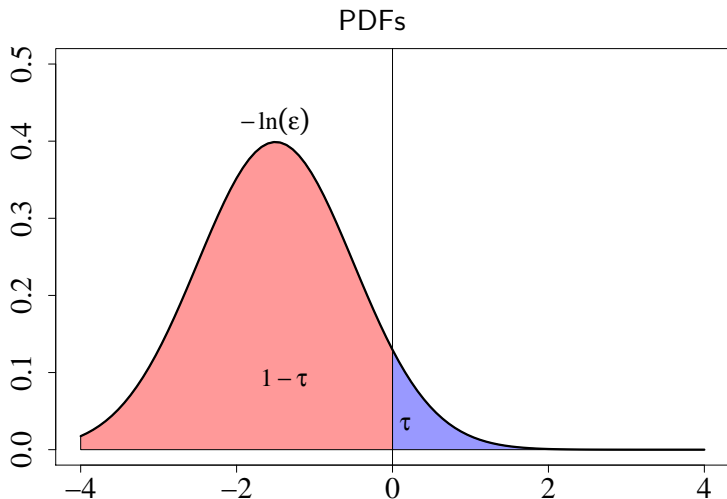
Log-linearization pictures



Log-linearization pictures

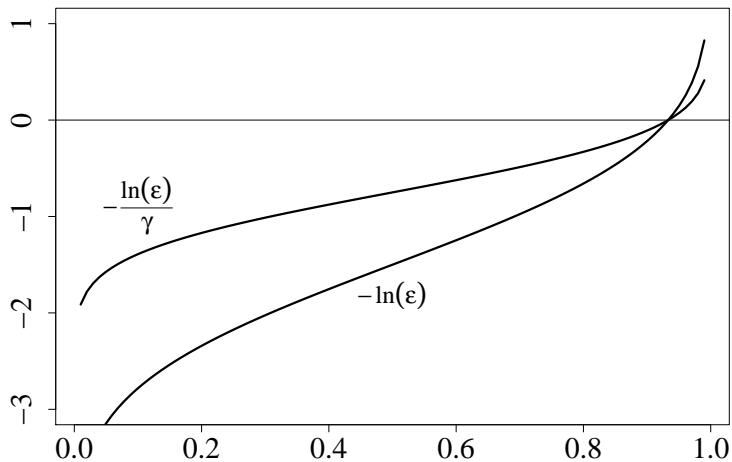


Log-linearization pictures



Log-linearization pictures

Quantile functions



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): MCMC (Chernozhukov and Hong, 2003), grid/MCMC (Lancaster and Jun, 2010), MIQP (Chen and Lee, 2017), binary treatment (Wüthrich, 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 and 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 and 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

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

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, 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 $U \sim \text{Unif}(0, 1)$
- 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 ($P(U \leq 0.5) = 0.5$)

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 $U \sim \text{Unif}(0, 1)$
- 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 ($P(U \leq 0.5) = 0.5$)
- If instrument vector $\mathbf{Z} \perp U$, then $P(Y \leq \mathbf{X}'\beta(\tau) \mid \mathbf{Z}) = P(U \leq \tau \mid \mathbf{Z}) = P(U \leq \tau) = \tau$: 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, Angrist, and Imbens (2002), Frölich and Melly (2013), et al.
- Wüthrich (2016) studies IVQR estimator under LQTE framework: estimates QTE for compliers but at transformed quantile levels

IVQR moment conditions

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

(linear IV)

(linear IVQR)

IVQR moment conditions

(linear IV)

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

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

IVQR moment conditions

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

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

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

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

IVQR moment conditions

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

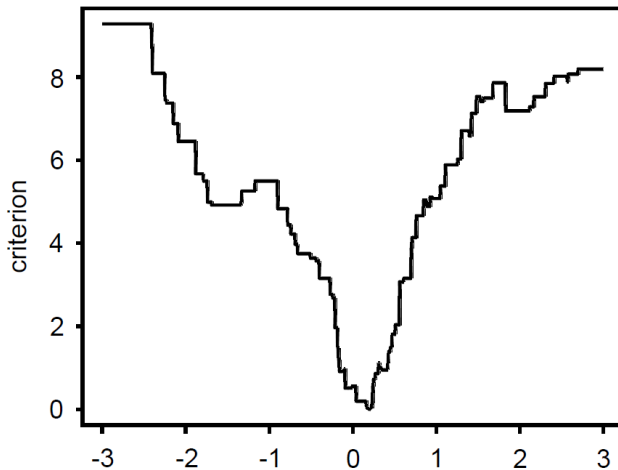
$$\mathbf{0} = \mathbb{E}[\mathbf{Z}(Y - \mathbf{X}'\boldsymbol{\beta}_0)], \boldsymbol{\beta}_0 = [\mathbb{E}(\mathbf{Z}\mathbf{X}')]^{-1} \mathbb{E}(\mathbf{Z}Y) \quad (\text{linear IV})$$

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

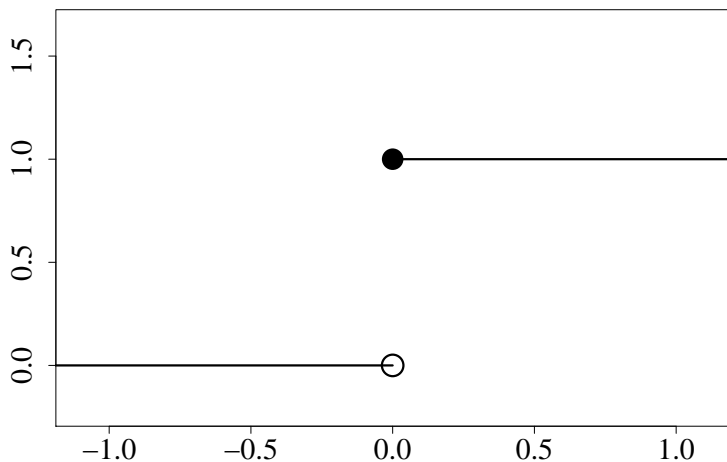
$$\mathbf{0} = \mathbb{E}[\mathbf{Z}(\mathbf{1}\{Y - \mathbf{X}'\boldsymbol{\beta}_{0\tau} \leq 0\} - \tau)] \quad (\text{linear IVQR})$$

Just run GMM?

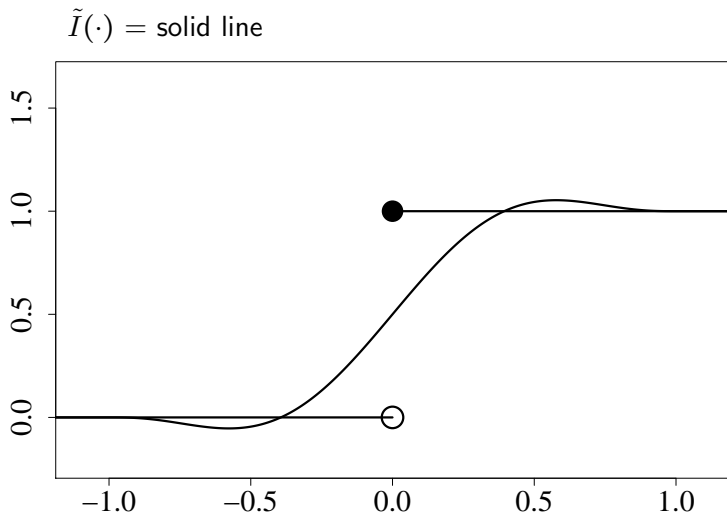
Chernozhukov and Hong (2003), Figure 1(a)
Criterion for IV-QR



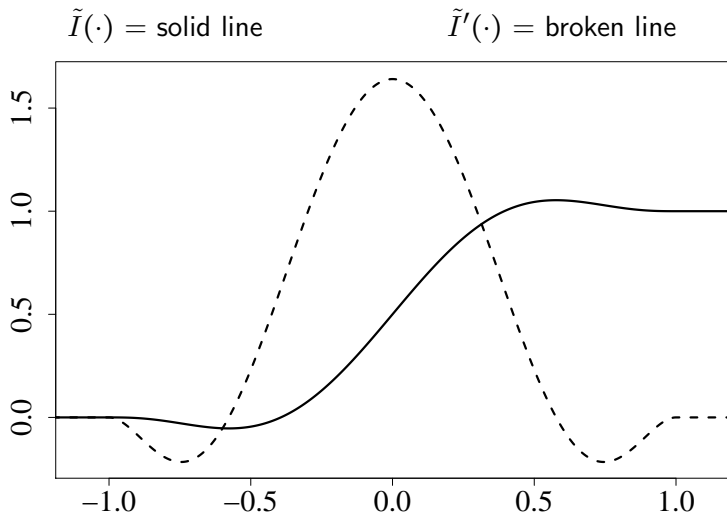
Smoothing the indicator function



Smoothing the indicator function



Smoothing the indicator function



Smoothed estimator

- Instead of solving sample moments

$\mathbf{0} = \hat{\mathbb{E}}[\mathbf{Z}(\mathbb{1}\{\mathbf{X}'\hat{\boldsymbol{\beta}}_{\tau} - Y \geq 0\} - \tau)]$, 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

$\mathbf{0} = \hat{\mathbb{E}}[\mathbf{Z}(\mathbb{1}\{\mathbf{X}'\hat{\boldsymbol{\beta}}_{\tau} - Y \geq 0\} - \tau)]$, 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
- $\tilde{I}'(\cdot)$: r th order kernel; $f_{U|\mathbf{Z},\mathbf{X}}(\cdot)$: $\geq r$ derivatives wrt u

Smoothed estimator

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

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

- Can compute Jacobian (wrt β)
- Easy/fast to compute (unless $h_n \approx 0$), standard solver
- $\tilde{I}'(\cdot)$: r th 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 $\hat{\beta}$. (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 $h_n \rightarrow \infty$, $\tilde{I}(\cdot/h_n)$ approx linear:

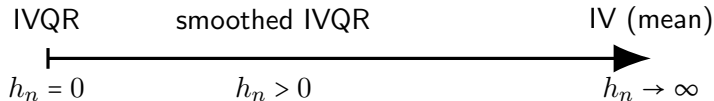
$$\mathbf{0} \approx \sum_{i=1}^n \mathbf{Z}_i \left[\left(0.5 + \frac{105}{64} \frac{\mathbf{X}'_i \hat{\boldsymbol{\beta}} - Y_i}{h} \right) - \tau \right] = \underline{\mathbf{Z}}' \underline{\mathbf{X}} \hat{\boldsymbol{\beta}} - \underline{\mathbf{Z}}' \mathbf{Y} - \underline{\mathbf{Z}}' \overbrace{\mathbf{X} \mathbf{e}_1}^{=(1, \dots, 1)'} \frac{64h}{105} (\tau - 0.5)$$

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

$$\mathbf{0} \approx \sum_{i=1}^n \mathbf{Z}_i \left[\left(0.5 + \frac{105}{64} \frac{\mathbf{X}'_i \hat{\boldsymbol{\beta}} - Y_i}{h} \right) - \tau \right] = \underline{\mathbf{Z}}' \underline{\mathbf{X}} \hat{\boldsymbol{\beta}} - \underline{\mathbf{Z}}' \mathbf{Y} - \underline{\mathbf{Z}}' \overbrace{\mathbf{X} \mathbf{e}_1}^{=(1, \dots, 1)'} \frac{64h}{105} (\tau - 0.5)$$

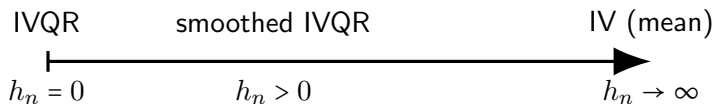


Connections: IV, Winsorized mean

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

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

$$\mathbf{0} \approx \sum_{i=1}^n \mathbf{z}_i \left[\left(0.5 + \frac{105}{64} \frac{\mathbf{X}'_i \hat{\boldsymbol{\beta}} - Y_i}{h} \right) - \tau \right] = \underline{\mathbf{z}}' \underline{\mathbf{X}} \hat{\boldsymbol{\beta}} - \underline{\mathbf{z}}' \mathbf{Y} - \underline{\mathbf{z}}' \overbrace{\mathbf{X} \mathbf{e}_1}^{=(1, \dots, 1)'} \frac{64h}{105} (\tau - 0.5)$$



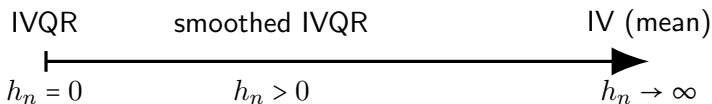
- Can pick h_n to improve efficiency (like median vs. mean); maybe better to average IVQR and IV like Hansen (2017)?

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

$$\mathbf{0} \approx \sum_{i=1}^n \mathbf{Z}_i \left[\left(0.5 + \frac{105}{64} \frac{\mathbf{X}_i' \hat{\boldsymbol{\beta}} - Y_i}{h} \right) - \tau \right] = \underline{\mathbf{Z}}' \underline{\mathbf{X}} \hat{\boldsymbol{\beta}} - \underline{\mathbf{Z}}' \mathbf{Y} - \underline{\mathbf{Z}}' \overbrace{\mathbf{X} \mathbf{e}_1}^{=(1, \dots, 1)'} \frac{64h}{105} (\tau - 0.5)$$



- Can pick h_n to improve efficiency (like median vs. mean); maybe better to average IVQR and IV like Hansen (2017)?
- Special case: $X_i = Z_i = 1$, $\tilde{I}'(u) = \mathbf{1}\{-1 \leq u \leq 1\}/2$, $\tau = 0.5 \implies$ "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

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

- 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

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

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

MSE of SEE

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

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

$$\underline{\mathbf{V}} \equiv \lim_{n \rightarrow \infty} \text{Var}(\mathbf{W}_i) = \tau(1 - \tau) \mathbf{E}(\mathbf{Z}_i \mathbf{Z}'_i)$$

MSE of SEE

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

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

$$\underline{\mathbf{V}} \equiv \lim_{n \rightarrow \infty} \text{Var}(\mathbf{W}_i) = \tau(1 - \tau) \mathbf{E}(\mathbf{Z}_i \mathbf{Z}'_i)$$

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

MSE of SEE

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

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

$$\underline{\mathbf{V}} \equiv \lim_{n \rightarrow \infty} \text{Var}(\mathbf{W}_i) = \tau(1 - \tau) \mathbf{E}(\mathbf{Z}_i \mathbf{Z}'_i)$$

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

■ \uparrow bias \implies \Downarrow var

MSE of SEE

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

MSE of SEE

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

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

MSE of SEE

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

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

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

MSE of SEE

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

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

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

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

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 $\Lambda(\mathbf{Y}, \mathbf{X}, \boldsymbol{\beta})$, like $Y_1 - (Y_2, \mathbf{X}')\boldsymbol{\beta}$

Setup

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

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

Setup

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

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

- Smoothed estimator:

$$\begin{aligned}\mathbf{0} &= \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^n \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau), \\ \mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) &\equiv \mathbf{g}_n(\mathbf{Y}_i, \mathbf{X}_i, \mathbf{Z}_i, \boldsymbol{\beta}, \tau) \\ &\equiv \mathbf{Z}_i[\tilde{I}(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n) - \tau].\end{aligned}$$

Assumptions

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^n \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$
$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i \left[\tilde{I}(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n) - \tau \right].$$

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

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

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

Assumption A4

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

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 $h_n \rightarrow 0$; WLLN from Andrews (1988).

Assumptions

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

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

Assumption A5

$E(\mathbf{Z}_i \mathbf{Z}_i')$ 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

$$\mathbf{0} = \hat{\mathbf{M}}_n(\hat{\boldsymbol{\beta}}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^n \mathbf{g}_{ni}(\hat{\boldsymbol{\beta}}_\tau, \tau),$$
$$\mathbf{g}_{ni}(\boldsymbol{\beta}, \tau) \equiv \mathbf{Z}_i \left[\tilde{I}'(-\Lambda(\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta})/h_n) - \tau \right].$$

Assumption A7

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

Assumption A8

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

Assumptions

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

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

Assumption A9

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

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

(ii) \implies local identification

Assumptions

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

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

Assumption A10

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

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

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

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

Assumption A11

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

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

$$A1-A3 \text{ and } A5-A8 \implies \sup_{\beta \in \mathcal{B}} |E[\hat{\mathbf{M}}_n(\beta, \tau)] - \mathbf{M}(\beta, \tau)| = o(1).$$

Proof.

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

Consistency

Theorem 2

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

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 $\beta_{0\tau}$ (A3), “well-separated” b/c compact \mathcal{B} (A3), cts $\mathbf{M}(\cdot)$ (can show). □

Asymptotic normality

Lemma 3

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

Proof.

- 1) $E[\hat{\mathbf{M}}_n(\boldsymbol{\beta}_{0\tau}, \tau)] = O(h_n^2)$, 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

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

Proof.

Mean value expansion: $\mathbf{0} = \hat{\mathbf{M}}_n(\beta_{0\tau}) + \underline{\dot{\mathbf{M}}}_n(\hat{\beta}_\tau - \beta_{0\tau})$, so
 $\sqrt{n}(\hat{\beta}_\tau - \beta_{0\tau}) = -[\underline{\dot{\mathbf{M}}}_n]^{-1} \sqrt{n} \hat{\mathbf{M}}_n(\beta_{0\tau})$.

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

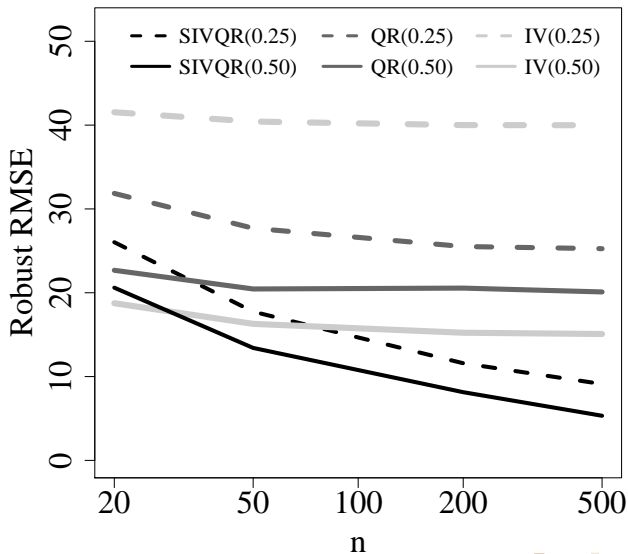
Simulation setup

- 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 $\text{IQR}/1.35$, so equals RMSE for normal distribution. (IV has no mean. . .)

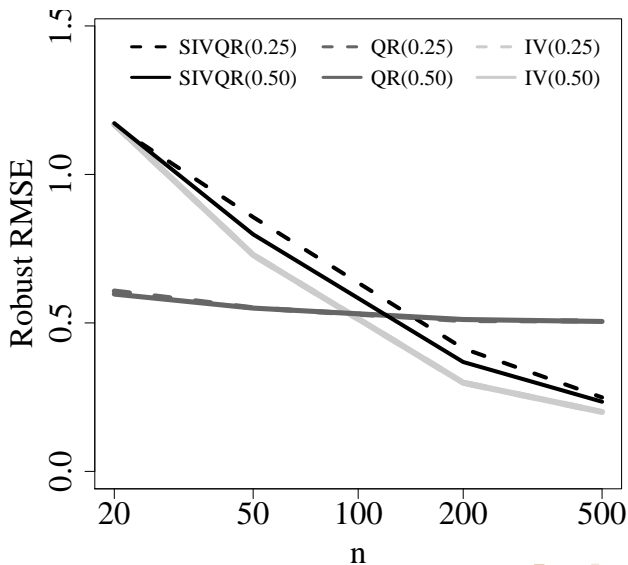
Simulation setup

- 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).

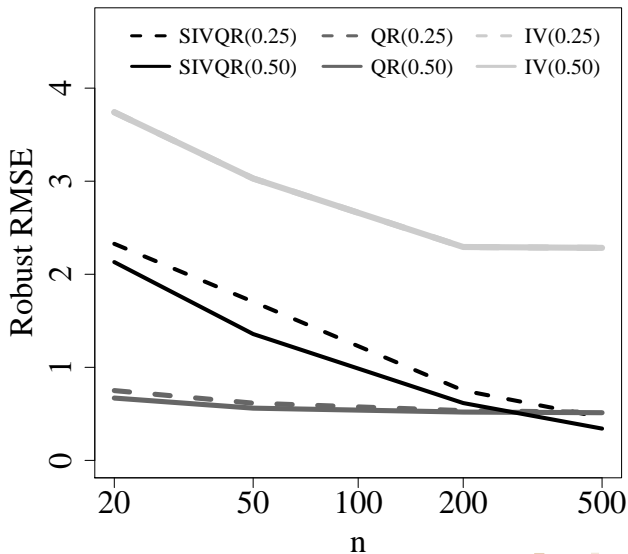
Robust RMSE: JTPA



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: $P(D_i = 1 | Z_i = 1) = 0.62$. 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

Regressor	Method	Quantile index τ				
		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

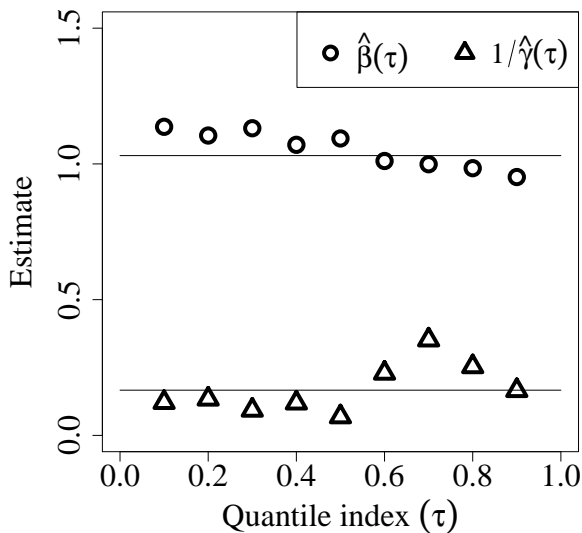
JTPA: results

Regressor	Method	Quantile index τ				
		0.15	0.25	0.50	0.75	0.85
<i>Original controls</i>						
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		
<i>Modified controls</i>						
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		

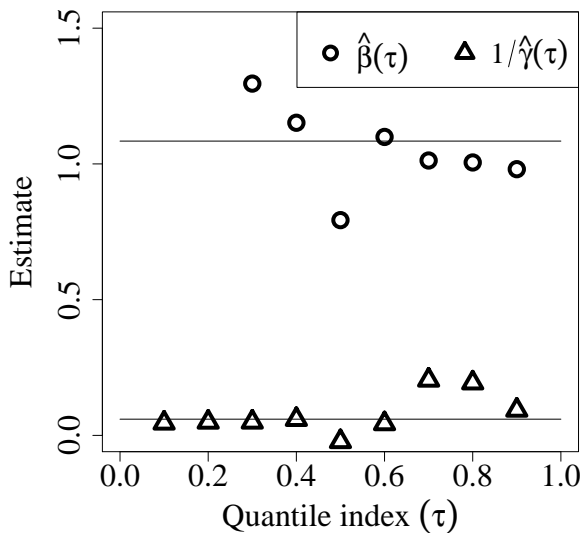
Quantile Euler equation estimates

- 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

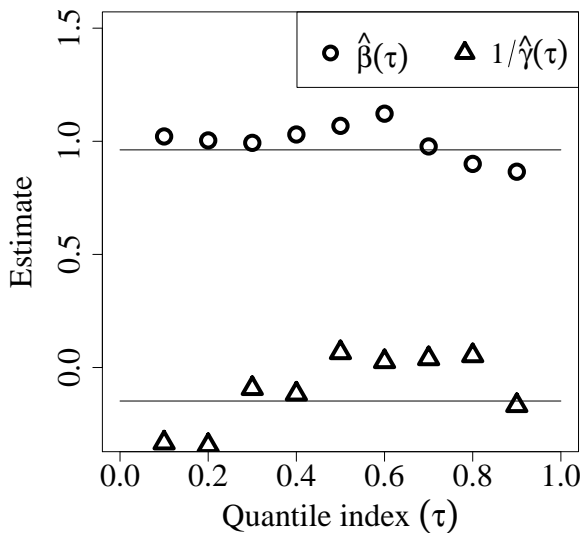
Quantile Euler equation estimates: UK



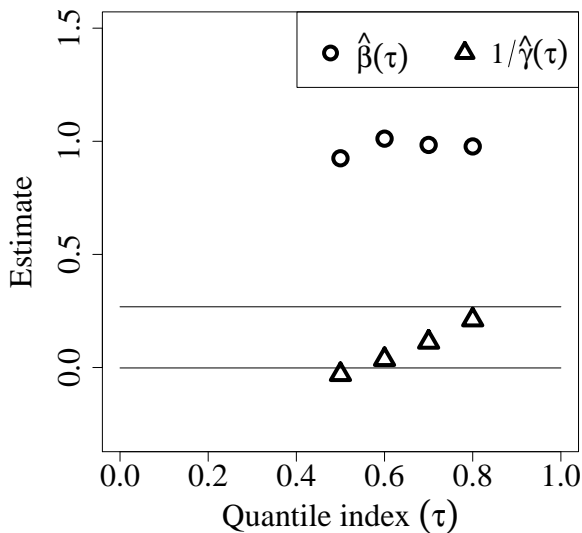
Quantile Euler equation estimates: USA



Quantile Euler equation estimates: Netherlands



Quantile Euler equation estimates: Sweden



Quantile Euler equation estimates

τ	USA		UK	
	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$
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

τ	US		UK		NTH		SWE	
	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$	$\hat{\beta}_\tau$	$\hat{\gamma}_\tau$
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

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

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., Angrist, J., Imbens, G., 2002. Instrumental variables estimates of the effect of subsidized training on the quantiles of trainee earnings. *Econometrica* 70 (1), 91–117.
URL <https://doi.org/10.1111/1468-0262.00270>
- Andrews, D. W. K., 1987. Consistency in nonlinear econometric models: A generic uniform law of large numbers. *Econometrica* 55 (6), 1465–1471.
URL <http://www.jstor.org/stable/1913568>
- Andrews, D. W. K., 1988. Laws of large numbers for dependent non-identically distributed random variables. *Econometric Theory* 4 (3), 458–467.
URL <http://www.jstor.org/stable/3532335>

References II

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

URL <http://www.jstor.org/stable/2938229>

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

URL <https://doi.org/10.3982/ECTA12868>

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

URL [https://doi.org/10.1016/S1574-0102\(03\)01022-7](https://doi.org/10.1016/S1574-0102(03)01022-7)

References III

- Carroll, C. D., 2001. Death to the log-linearized consumption Euler equation! (and very poor health to the second-order approximation). *Advances in Macroeconomics* 1 (1).
URL <https://doi.org/10.2202/1534-6013.1003>
- Chambers, C. P., 2009. An axiomatization of quantiles on the domain of distribution functions. *Mathematical Finance* 19, 335–342.
- Chen, L.-Y., Lee, S., 2017. Exact computation of GMM estimators for instrumental variable quantile regression models, working paper, available at <https://arxiv.org/abs/1703.09382>.
- Cheng, X., Liao, Z., Shi, R., 2016. An averaging GMM estimator robust to misspecification, working paper, available at http://www.econ.ucla.edu/liao/papers_pdf/ChengLiaoShi_AvgGMM.pdf.

References IV

- Chernozhukov, V., Hansen, C., 2005. An IV model of quantile treatment effects. *Econometrica* 73 (1), 245–261.
URL <http://www.jstor.org/stable/3598944>
- Chernozhukov, V., Hansen, C., 2006. Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics* 132 (2), 491–525.
URL <https://doi.org/10.1016/j.jeconom.2005.02.009>
- Chernozhukov, V., Hansen, C., Jansson, M., 2009. Finite sample inference for quantile regression models. *Journal of Econometrics* 152 (2), 93–103.
URL <https://doi.org/10.1016/j.jeconom.2009.01.004>

References V

- Chernozhukov, V., Hansen, C., Wüthrich, K., 2017. Instrumental variable quantile regression. In: Koenker, R., Chernozhukov, V., He, X., Peng, L. (Eds.), Handbook of Quantile Regression. CRC/Chapman-Hall, p. forthcoming.
- Chernozhukov, V., Hong, H., 2003. An MCMC approach to classical estimation. *Journal of Econometrics* 115 (2), 293–346. URL [https://doi.org/10.1016/S0304-4076\(03\)00100-3](https://doi.org/10.1016/S0304-4076(03)00100-3)
- Chesher, A., 2003. Identification in nonseparable models. *Econometrica* 71 (5), 1405–1441.
- de Castro, L., Galvao, A. F., 2017. Dynamic quantile models of rational behavior, university of Iowa, mimeo.

References VI

- de Castro, L., Galvao, A. F., Kaplan, D. M., 2017. Smoothed instrumental variables quantile regression, with estimation of quantile Euler equations, working paper, available at <https://faculty.missouri.edu/~kaplandm>.
- Fernandes, M., Guerre, E., Horta, E., 2017. Smoothing quantile regressions, mimeo, available at <http://bibliotecadigital.fgv.br/dspace/handle/10438/18390>.
- Frölich, M., Melly, B., 2013. Unconditional quantile treatment effects under endogeneity. *Journal of Business & Economic Statistics* 31 (3), 346–357.
- Giovannetti, B. C., 2013. Asset pricing under quantile utility maximization. *Review of Financial Economics* 22 (4), 169–179. URL <https://doi.org/10.1016/j.rfe.2013.05.008>

References VII

- 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.
URL <http://www.jstor.org/stable/2999619>
- Huber, P. J., 1964. Robust estimation of a location parameter. *Annals of Mathematical Statistics* 35 (1), 73–101.
URL <http://www.jstor.org/stable/2238020>
- Kaplan, D. M., Sun, Y., 2017. Smoothed estimating equations for instrumental variables quantile regression. *Econometric Theory* 33 (1), 105–157.
URL <https://doi.org/10.1017/S0266466615000407>

References VIII

- Kato, K., 2012. Asymptotic normality of Powell's kernel estimator. *Annals of the Institute of Statistical Mathematics* 64 (2), 255–273.
URL <https://doi.org/10.1007/s10463-010-0310-9>
- Lancaster, T., Jun, S. J., 2010. Bayesian quantile regression methods. *Journal of Applied Econometrics* 25 (2), 287–307.
URL <https://doi.org/10.1002/jae.1069>
- Lee, S., 2007. Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics* 141 (2), 1131–1158.
URL <https://doi.org/10.1016/j.jeconom.2007.01.014>
- MaCurdy, T., Hong, H., 1999. Smoothed quantile regression in generalized method of moments, mimeo.

References IX

- Manski, C. F., 1988. Ordinal utility models of decision making under uncertainty. *Theory and Decision* 25 (1), 79–104.
URL <https://doi.org/10.1007/BF00129169>
- Melly, B., Wüthrich, K., 2017. Local quantile treatment effects. In: Koenker, R., Chernozhukov, V., He, X., Peng, L. (Eds.), *Handbook of Quantile Regression*. CRC/Chapman-Hall, p. forthcoming.
- Newey, W. K., McFadden, D., 1994. Large sample estimation and hypothesis testing. In: Engle, R. F., McFadden, D. L. (Eds.), *Handbook of Econometrics*. Vol. 4. Elsevier, Ch. 36, pp. 2111–2245.
URL [https://doi.org/10.1016/S1573-4412\(05\)80005-4](https://doi.org/10.1016/S1573-4412(05)80005-4)

References X

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

URL <https://doi.org/10.1016/j.jeconom.2007.08.016>

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

URL <https://doi.org/10.1080/01621459.1994.10476870>

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

URL [https://doi.org/10.1016/0304-4076\(84\)90004-6](https://doi.org/10.1016/0304-4076(84)90004-6)

References XI

- Powell, J. L., 1991. Estimation of monotonic regression models under quantile restrictions. In: Barnett, W. A., Powell, J., Tauchen, G. (Eds.), Nonparametric and semiparametric methods in econometrics and statistics: proceedings of the Fifth International Symposium in Economic Theory and Econometrics. Cambridge University Press, pp. 357–384.
- Rostek, M., 2010. Quantile maximization in decision theory. *Review of Economic Studies* 77 (1), 339–371.
URL
<https://doi.org/10.1111/j.1467-937X.2009.00564.x>
- van der Vaart, A. W., 1998. *Asymptotic Statistics*. Cambridge University Press, Cambridge.

References XII

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

URL <https://doi.org/10.1017/S0266466606060087>

Wooldridge, J. M., 1986. Asymptotic properties of econometric estimators. Ph.D. thesis, Department of Economics, University of California, San Diego.

Wüthrich, K., 2016. A comparison of two quantile models with endogeneity, working paper.

Wüthrich, K., 2017. A closed-form estimator for quantile treatment effects with endogeneity, working paper.

References XIII

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

URL <https://doi.org/10.1162/0034653041811770>