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

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

\( \Omega_t \): information set at time \( t \)

\( R_{t+1} \): real rate of return of asset

\( C_t \): real consumption at time \( t \)

\( \beta \): discount factor (e.g., \( \beta = 0.99 \))

\( 1/\gamma \): elasticity of intertemporal substitution (EIS)

- Estimation: use variables in \( \Omega_t \) as instruments (inflation, etc.); run GMM, or IV/2SLS after log-linearization
Standard consumption Euler equation

- Expected utility maximization, \( U(C) = C^{1-\gamma}/(1 - \gamma) \):
  \[
  0 = E[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 | \Omega_t],
  \]

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

- Estimation: use variables in \( \Omega_t \) as instruments (inflation, etc.); run GMM, or IV/2SLS after log-linearization
- Drawback: no separation of EIS (\( 1/\gamma \)) and risk aversion (\( \gamma \))
- Drawback: approximation error from log-linearization (e.g., Carroll, 2001)
Quantile Euler equation?

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

- Standard: \( 0 = E\left[ \beta (1 + R_{t+1}) \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} - 1 \mid \Omega_t \right] \)
- Replace \( E[\cdot \mid \Omega_t] \) with conditional \( \tau \)-quantile \( Q_{\tau}[\cdot \mid \Omega_t] \):
  \[
  0 = Q_{\tau}\left[ \beta (1 + R_{t+1}) \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} - 1 \mid \Omega_t \right]
  \]
- Advantage: \( 1/\gamma \) is EIS, but both \( \tau \) and \( \gamma \) 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: \( 0 = \mathbb{E}[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 | \Omega_t] \)
- Replace \( \mathbb{E}[\cdot | \Omega_t] \) with conditional \( \tau \)-quantile \( Q_\tau[\cdot | \Omega_t] \):
  \[
  0 = Q_\tau[\beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma} - 1 | \Omega_t]
  \]

- Advantage: \( 1/\gamma \) is EIS, but both \( \tau \) and \( \gamma \) 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 ($\infty$-horizon): de Castro and Galvao (2017) show dynamic consistency and derive Euler equation
Quantile Euler equation: estimation

- Can write as $Q_T[\epsilon_{t+1} | \Omega_t] = 1$, $\epsilon_{t+1} \equiv \beta(1 + R_{t+1})(C_{t+1}/C_t)^{-\gamma}$
- Since $\ln(\cdot)$ is strictly increasing, $Q_T[\ln(W)] = \ln(Q_T[W])$
- 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) \text{ strictly } \downarrow \text{ in } \epsilon:

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

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

- Parameters for $\tau$-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

PDFs

\[ \ln(\varepsilon) \]
Log-linearization pictures

PDFs

$\tau$ $\tau$

$-\ln(\varepsilon)$ $\ln(\varepsilon)$
Log-linearization pictures

PDFs

$-\ln(\varepsilon)$

$1 - \tau$

$\tau$
Log-linearization pictures

Quantile functions

\[ -\frac{\ln(\varepsilon)}{\gamma} \]

\[ -\ln(\varepsilon) \]
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)
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) $\tau$-quantile treatment effects, or “structural quantile effects”
What does IVQR estimate?

- Chernozhukov and Hansen (2005): identification of (conditional) \( \tau \)-quantile treatment effects, or “structural quantile effects”

- Simplistic example: linear structural random coefficient model, \( Y = X'\beta(U) \), assume \( X'\beta(U) \) monotonic in unobserved \( U \sim \text{Unif}(0,1) \)

- If \( Y \) is wage, \( U \) is “ability”: \( X'\beta(0.5) \) traces out potential wage outcomes (given different \( X \)) for individual with median ability (\( P(U \leq 0.5) = 0.5 \))
What does IVQR estimate?

- Chernozhukov and Hansen (2005): identification of (conditional) \( \tau \)-quantile treatment effects, or “structural quantile effects”

- Simplistic example: linear structural random coefficient model, \( Y = X' \beta(U) \), assume \( X' \beta(U) \) monotonic in unobserved \( U \sim \text{Unif}(0, 1) \)

- If \( Y \) is wage, \( U \) is “ability”: \( X' \beta(0.5) \) traces out potential wage outcomes (given different \( X \)) for individual with median ability (\( P(U \leq 0.5) = 0.5 \))

- If instrument vector \( Z \perp U \), then
  \[
P(Y \leq X' \beta(\tau) \mid Z) = P(U \leq \tau \mid Z) = P(U \leq \tau) = \tau: \text{ a conditional quantile restriction on the observables } Y, X, \text{ and } Z, \text{ and the parameter vector } \beta(\tau)
  \]
Alternative quantile models with endogeneity

- Triangular system: compared in Chernozhukov, Hansen, and Wüthrich (2017, §2.5); like Chesher (2003), Lee (2007), et al.
- Wüthrich (2016) studies IVQR estimator under LQTE framework: estimates QTE for compliers but at transformed quantile levels
IVQR moment conditions

\[ \tau = P(Y \leq h(X, \tau) \mid Z) \]

(linear IV)

(linear IVQR)
IVQR moment conditions

\[ \tau = P(Y \leq h(X, \tau) \mid Z), \quad h(X, \tau) = X' \beta_{0\tau}, \quad P(\cdot) = E(1\{\cdot\}) \implies 0 = E[Z(1\{Y - X' \beta_{0\tau} \leq 0\} - \tau)] \]

(linear IV)

(linear IVQR)
IVQR moment conditions

\[ Y = h(X) + U, \quad E(U \mid Z) = 0, \quad h(X) = X'\beta_0 \quad \implies \]
\[ 0 = E[Z(Y - X'\beta_0)] \quad \text{(linear IV)} \]

\[ \tau = P(Y \leq h(X, \tau) \mid Z), \quad h(X, \tau) = X'\beta_{0\tau}, \quad P(\cdot) = E(1\{\cdot\}) \quad \implies \]
\[ 0 = E[Z(1\{Y - X'\beta_{0\tau} \leq 0\} - \tau)] \quad \text{(linear IVQR)} \]
IVQR moment conditions

\[ Y = h(X) + U, \ E(U \mid Z) = 0, \ h(X) = X'\beta_0 \quad \implies \]
\[ 0 = E[Z(Y - X'\beta_0)], \ \beta_0 = [E(ZX')]^{-1} E(ZY) \quad \text{(linear IV)} \]
\[ \tau = P(Y \leq h(X, \tau) \mid Z), \ h(X, \tau) = X'\beta_{0\tau}, \ P(\cdot) = E(1\{\cdot\}) \quad \implies \]
\[ 0 = E[Z(1\{Y - X'\beta_{0\tau} \leq 0\} - \tau)] \quad \text{(linear IVQR)} \]
Just run GMM?

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

Criterion for IV-QR
Smoothing the indicator function
Smoothing the indicator function

\[ \tilde{I}(\cdot) = \text{solid line} \]
Smoothing the indicator function

\[ \tilde{I}(\cdot) = \text{solid line} \quad \tilde{I}'(\cdot) = \text{broken line} \]
Instead of solving sample moments
\[ 0 = \hat{E}\left[Z(1\{X'\hat{\beta}_\tau - Y \geq 0\} - \tau)\right], \]
replace \(1\{\cdot \geq 0\}\) with smoothed \(\tilde{I}(\cdot/h_n)\), bandwidth \(h_n\):
\[
0 = n^{-1} \sum_{i=1}^{n} Z_i \left[ \tilde{I}\left(\frac{X_i'\hat{\beta}_\tau - Y_i}{h_n}\right) - \tau \right]
\]
Smoothed estimator

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

  \[
  0 = n^{-1} \sum_{i=1}^{n} Z_i \left[ \hat{I}\left(\frac{X_i'\hat{\beta}_\tau - Y_i}{h_n}\right) - \tau \right]
  \]

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

- Instead of solving sample moments
  
  \[ 0 = \hat{E}[Z(1\{X'\hat{\beta}_{\tau} - Y \geq 0\} - \tau)] , \]

  replace \(1\{\cdot \geq 0\}\) with smoothed \(\tilde{I}(\cdot/h_n)\), bandwidth \(h_n\):

  \[
  0 = n^{-1} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i'\hat{\beta}_{\tau} - Y_i}{h_n} \right) - \tau \right]
  \]

- Can compute Jacobian (wrt \(\beta\))

- Easy/fast to compute (unless \(h_n \approx 0\)), standard solver

- \(\tilde{I}'(\cdot)\): \(r\)th order kernel; \(f_{U|Z,X}(\cdot)\): \(\geq r\) 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

\[ 0 = n^{-1} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i \hat{\beta} - Y_i}{h_n} \right) - \tau \right] \]

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

\[ 0 \approx \sum_{i=1}^{n} Z_i \left[ \left( 0.5 + \frac{105}{64} \frac{X_i \hat{\beta} - Y_i}{h} \right) - \tau \right] = \mathbf{Z}' \mathbf{X} \hat{\beta} - \mathbf{Z}' \mathbf{Y} - \mathbf{Z}' \mathbf{X e}_1 \frac{64h}{105} (\tau - 0.5) \]
Connections: IV, Winsorized mean

\[ 0 = n^{-1} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i \hat{\beta}_\tau - Y_i}{h_n} \right) - \tau \right] \]

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

\[ 0 \approx \sum_{i=1}^{n} Z_i \left[ \left( 0.5 + \frac{105}{64} \frac{X_i \hat{\beta} - Y_i}{h} \right) - \tau \right] = Z'X\hat{\beta} - Z'Y - Z' \overrightarrow{Xe_1} \frac{64h}{105}(\tau-0.5) \]

- IVQR
- Smoothed IVQR
- IV (mean)

\( h_n = 0 \)
\( h_n > 0 \)
\( h_n \to \infty \)
Connections: IV, Winsorized mean

\[ 0 = n^{-1} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i \hat{\beta} - Y_i}{h_n} \right) - \tau \right] \]

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

\[ 0 \approx \sum_{i=1}^{n} Z_i \left[ \left( 0.5 + \frac{105}{64} \frac{X_i \hat{\beta} - Y_i}{h} \right) - \tau \right] = Z'X\hat{\beta} - Z'Y - Z' \widehat{Xe_1} \frac{64h}{105} (\tau - 0.5) \]

- 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

\[ 0 = n^{-1} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i \hat{\beta} - Y_i}{h_n} \right) - \tau \right] \]

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

\[ 0 \approx \sum_{i=1}^{n} Z_i \left[ \left( 0.5 + \frac{105}{64} \frac{X_i \hat{\beta} - Y_i}{h} \right) - \tau \right] = Z'X \hat{\beta} - Z'Y - Z' \frac{64h}{105} (\tau - 0.5) \]

IVQR \quad \text{smoothed IVQR} \quad \text{IV (mean)}

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

- Special case: \( X_i = Z_i = 1 \), \( \tilde{I}'(u) = 1 \{ -1 \leq u \leq 1 \} / 2 \), \( \tau = 0.5 \) \( \implies \) “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

\[ Y_i = X_i' \beta_\tau + U_i, \quad P(U_i \leq 0 \mid Z_i) = \tau \]

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

\[ m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i' \beta_0 - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i \]
MSE of SEE

\[ m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i' \beta_{0 \tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i \]

\[ E(W_i) = \frac{h_n^r}{r!} \left( \int \tilde{I}'(v)v^r \, dv \right) E\left[ f_U^{(r-1)}(0 \mid Z_j)Z_j \right] + o(h_n^r) \]
Theoretical Simulations Empirical

MSE of SEE

\[
m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i' \beta_{0\tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i
\]

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

\[
V \equiv \lim_{n \to \infty} \text{Var}(W_i) = \tau(1-\tau) E(Z_iZ_i')
\]
MSE of SEE

\[ m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[ \hat{I} \left( \frac{X'_i \beta_{0\tau} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i \]

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

\[ V \equiv \lim_{n \to \infty} \text{Var}(W_i) = \tau(1 - \tau) E(Z_i Z'_i) \]

\[ E(W_i W'_i) = V - h_n \left[ 1 - \int_{-1}^{1} [\hat{I}(u)]^2 \, du \right] E \{ f_{U|Z}(0 \mid Z_i)Z_i Z'_i \} + o(h_n) \]
MSE of SEE

\[ m_n \equiv n^{-1/2} \sum_{i=1}^{n} Z_i \left[ \tilde{I} \left( \frac{X_i' \beta_{0T} - Y_i}{h_n} \right) - \tau \right] \equiv n^{-1/2} \sum_{i=1}^{n} W_i \]

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

\[ \overline{V} \equiv \lim_{n \to \infty} \text{Var}(W_i) = \tau (1 - \tau) E(Z_i Z'_i) \]

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

\[ \uparrow \text{bias} \implies \downarrow \text{var} \]
MSE of SEE

\[ h^* \equiv \arg \min_h E\{m'_n V^{-1} m_n\} \]
MSE of SEE

\[ h^* \equiv \arg \min_h E\{m_n' V^{-1} m_n\} \]

\[ E\{m_n' V^{-1} m_n\} = n^{-1} \sum_{j=1}^{n} \left( E\{W_j' V^{-1} W_j\} + \sum_{i \neq j} E\{W_i' V^{-1} W_j\} \right) \]
MSE of SEE

\[ h^* \equiv \arg \min_h E\{m_n' V^{-1} m_n\} \]

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

\[ h^* = n^{-1/(2r-1)} \cdot \frac{1}{2^{r-1}} \]
MSE of SEE

\[ h^* \equiv \arg \min_h \mathbb{E}\{m'_n V^{-1} m_n\} \]

\[ \mathbb{E}\{m'_n V^{-1} m_n\} = n^{-1} \sum_{j=1}^{n} \left( \mathbb{E}\{W'_j V^{-1} W_j\} + \sum_{i \neq j} \mathbb{E}\{W_i V^{-1} W_j\} \right) \]

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

\[ h^* = \left( \frac{(r!)^2 \left[ 1 - \int_{-1}^{1} \tilde{I}^2(u) \, du \right] f_U(0) \, d}{2r \left( \int \tilde{I}'(v) v^r \, dv \right)^2 \left[ f_U^{(r-1)}(0) \right]^2 n} \right)^{\frac{1}{2r-1}} \quad \text{if } U \perp Z \]
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
- 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 $Y$, instruments $Z$ with subset $X$
- Residual fn $\Lambda(Y, X, \beta)$, like $Y_1 - (Y_2, X')\beta$
Setup

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

$$0 = E\{Z_i[1\{\Lambda(Y_i, X_i, \beta_{0\tau}) \leq 0\} - \tau]\} = M(\beta_{0\tau}, \tau)$$
Setup

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

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

- Smoothed estimator:

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

$$g_{ni}(\beta, \tau) \equiv g_n(\mathbf{Y}_i, \mathbf{X}_i, \mathbf{Z}_i, \beta, \tau)$$

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

\[ 0 = \hat{M}_n(\hat{\beta}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^{n} g_{ni}(\hat{\beta}_\tau, \tau), \]

\[ g_{ni}(\beta, \tau) \equiv Z_i \left[ \tilde{I}(\Lambda(Y_i, X_i, \beta)/h_n) - \tau \right]. \]

Assumption A1

*Strictly stationary, weakly dependent data.*

Assumption A2

*\( \Lambda(\cdot) \) known, differentiable in \( \beta \).*

Assumption A3

*Global point identification of \( \beta_{0\tau} \); interior of compact \( \mathcal{B} \).*
Assumptions

\[ 0 = \hat{M}_n(\hat{\beta}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^{n} g_{ni}(\hat{\beta}_\tau, \tau), \]

\[ g_{ni}(\beta, \tau) \equiv Z_i \left[ \tilde{I} \left( -\Lambda(Y_i, X_i, \beta) / h_n \right) - \tau \right]. \]

Assumption A4

**ULLN:** \( \sup_{\beta \in B} |\hat{M}_n(\beta, \tau) - E[\hat{M}_n(\beta, \tau)]| = o_p(1). \)

Note \( E[\hat{M}_n(\beta, \tau)] \neq M(\beta, \tau). \) Paper: example primitive conditions, using Andrews (1987). Use: \( g_{ni} \leq 2|Z_i| \) and \( h_n \to 0; \) WLLN from Andrews (1988).
Assumptions

\[ 0 = \hat{M}_n(\hat{\beta}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^{n} g_{ni}(\hat{\beta}_\tau, \tau), \]

\[ g_{ni}(\beta, \tau) \equiv Z_i \left[ \tilde{I}(\Lambda(Y_i, X_i, \beta)/h_n) - \tau \right]. \]

**Assumption A5**

\( E(Z_iZ'_i) \) is positive definite (and finite).

(No moment restrictions on \( Y_i \).)

**Assumption A6**

*Distribution of \( \Lambda(Y_i, X_i, \beta) \) given \( (\beta, Z_i = z) \) is continuous at zero.*

E.g., \( Y_1 \) cts given \( (Y_2, Z) \) for linear IVQR.
Assumptions

\[ 0 = \hat{M}_n(\hat{\beta}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^{n} g_{ni}(\hat{\beta}_\tau, \tau), \]

\[ g_{ni}(\beta, \tau) \equiv Z_i \left[ \tilde{I}(\Lambda(Y_i, X_i, \beta)/h_n) - \tau \right]. \]

**Assumption A7**

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

**Assumption A8**

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

\[ 0 = \hat{M}_n(\hat{\beta}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^{n} g_{ni}(\hat{\beta}_\tau, \tau), \]

\[ g_{ni}(\beta, \tau) \equiv Z_i [\tilde{I}(\Lambda(Y_i, X_i, \beta)/h_n) - \tau]. \]

Assumption A9

Let \( \Lambda_i \equiv \Lambda(Y_i, X_i, \beta_{0\tau}) \) and \( D_i \equiv \nabla_\beta \Lambda(Y_i, X_i, \beta_{0\tau}) \). (i) \( f_{\Lambda|Z}(\cdot | z) \) twice differentiable (also \( f_{\Lambda|Z,D} \)). (ii) Nonsingular \( G = \nabla_\beta M(\beta_{0\tau}, \tau) = -E\{Z_i D_i' f_{\Lambda|Z,D}(0 | Z_i, D_i)\} \).

E.g., \( D_i \) is regressor vector for linear IVQR. (ii) \( \implies \) local identification
Assumptions

\[ 0 = \hat{M}_n(\hat{\beta}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^{n} g_{ni}(\hat{\beta}_\tau, \tau), \]

\[ g_{ni}(\beta, \tau) \equiv Z_i \left[ \tilde{I}(\Lambda(Y_i, X_i, \beta)/h_n) - \tau \right]. \]

Assumption A10

\[-\frac{1}{nh_n} \sum_{i=1}^{n} \tilde{I}'(\Lambda(Y_i, X_i, \hat{\beta}_\tau)/h_n) Z_i \nabla_\beta \Lambda(Y_i, \hat{X}_i, \hat{\beta}_\tau)' \xrightarrow{p} G. \]

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

\[ 0 = \hat{M}_n(\hat{\beta}_\tau, \tau) = \frac{1}{n} \sum_{i=1}^{n} g_{ni}(\beta, \tau), \]

\[ g_{ni}(\beta, \tau) \equiv Z_i \left[ \tilde{I} \left( -\Lambda(Y_i, X_i, \beta)/h_n \right) - \tau \right]. \]

Assumption A11

*Pointwise CLT:* \( \sqrt{n} \left\{ \hat{M}_n(\beta_{0\tau}, \tau) - E[\hat{M}_n(\beta_{0\tau}, \tau)] \right\} \overset{d}{\to} N(0, \Sigma_\tau). \)

Primitive conditions: moment and dependence restrictions.

Ex: iid, \( E(\|Z_i\|^2) < \infty. \)

Ex: Wooldridge (1986), NED (\ldots), \( E(\|Z_i\|^{2+\epsilon}) < \infty. \)
Lemma 1

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

Proof.

Use dominated convergence theorem. Need cts distribution of \( \Lambda(Y, X, \beta) \) since \( \tilde{I}(0) = 0.5 \neq 1 = 1 \{0 \geq 0\} \).
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): $\hat{M}_n(\cdot) \xrightarrow{p} M(\cdot)$ uniformly. Maximizer of $-\|M(\cdot)\|$ is uniquely $\beta_{0\tau}$ (A3), “well-separated” b/c compact $B$ (A3), cts $M(\cdot)$ (can show).
Asymptotic normality

Lemma 3

$A1$–$A3$, $A5$, $A7$–$A9$, and $A11 \iff \sqrt{n}\hat{M}_n(\beta_{0\tau}, \tau) \xrightarrow{d} N(0, \Sigma_\tau)$.

Proof.

1) $E[\hat{M}_n(\beta_{0\tau}, \tau)] = O(h_n^2)$, 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

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

Proof.

Mean value expansion: $0 = \hat{M}_n(\beta_{0\tau}) + \dot{M}_n(\hat{\beta}_\tau - \beta_{0\tau})$, so

$\sqrt{n}(\hat{\beta}_\tau - \beta_{0\tau}) = -[\dot{M}_n]^{-1}\sqrt{n}\dot{M}_n(\beta_{0\tau}).$

Apply CMT to Lemma 3 and $\frac{\dot{M}_n}{\sqrt{n}} \xrightarrow{p} G$ (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
- Stationary bootstrap from Politis and Romano (1994).
Robust RMSE: JTPA

![Graph showing Robust RMSE for different methods: SIVQR(0.25), SIVQR(0.50), QR(0.25), QR(0.50), IV(0.25), IV(0.50). The x-axis represents the sample size (n) ranging from 20 to 500, and the y-axis represents Robust RMSE ranging from 0 to 50. The methods are plotted as lines with different styles and colors.](image-url)
Robust RMSE: TS–IV, normal

![Graph showing Robust RMSE for different methods: SIVQR(0.25), SIVQR(0.50), QR(0.25), QR(0.50), IV(0.25), IV(0.50). The x-axis represents the sample size (n) ranging from 20 to 500, and the y-axis represents the Robust RMSE, ranging from 0.0 to 1.5. The graph illustrates the performance of these methods under normal conditions.](image-url)
Robust RMSE: TS–IV, Cauchy

![Graph showing robust RMSE for various methods.

- SIVQR(0.25)
- QR(0.25)
- IV(0.25)
- SIVQR(0.50)
- QR(0.50)
- IV(0.50)
Size, 2-sided test of $H_0 : \gamma_T = \gamma_0$

<table>
<thead>
<tr>
<th>DGP</th>
<th>$\tau$</th>
<th>$n$</th>
<th>$\alpha$</th>
<th>Wald</th>
<th>BS-t</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>JTPA</td>
<td>0.25</td>
<td>100</td>
<td>0.10</td>
<td>0.411</td>
<td>0.094</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>0.10</td>
<td>0.415</td>
<td>0.067</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>0.10</td>
<td>0.226</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>0.50</td>
<td>100</td>
<td>0.10</td>
<td>0.550</td>
<td>0.074</td>
<td>0.120</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.10</td>
<td>0.342</td>
<td>0.041</td>
<td>0.101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10,000</td>
<td>0.10</td>
<td>0.134</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>
## Size, 2-sided test of $H_0 : \gamma_T = \gamma_0$

<table>
<thead>
<tr>
<th>DGP</th>
<th>$\tau$</th>
<th>$n$</th>
<th>$\alpha$</th>
<th>Wald</th>
<th>BS–t</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS–IV.N</td>
<td>0.25</td>
<td>100</td>
<td>0.10</td>
<td>0.269</td>
<td>0.098</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>0.10</td>
<td>0.154</td>
<td>0.107</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>0.10</td>
<td>0.116</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>100</td>
<td>0.10</td>
<td>0.246</td>
<td>0.103</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1000</td>
<td>0.10</td>
<td>0.147</td>
<td>0.119</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>0.10</td>
<td>0.101</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
Size, 2-sided test of \( H_0 : \gamma_\tau = \gamma_0 \)

<table>
<thead>
<tr>
<th>DGP</th>
<th>( \tau )</th>
<th>( n )</th>
<th>( \alpha )</th>
<th>Wald</th>
<th>BS-( t )</th>
<th>BS</th>
</tr>
</thead>
<tbody>
<tr>
<td>TS–IV.C</td>
<td>0.25</td>
<td>100</td>
<td>0.10</td>
<td>0.389</td>
<td>0.059</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.10</td>
<td>0.179</td>
<td>0.075</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10,000</td>
<td>0.10</td>
<td>0.116</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>100</td>
<td>0.10</td>
<td>0.296</td>
<td>0.067</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>0.10</td>
<td>0.159</td>
<td>0.110</td>
<td>0.079</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10,000</td>
<td>0.10</td>
<td>0.092</td>
<td>n/a</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>
JTPA: context

- Abadie, Angrist, and Imbens (2002), 5102 adult men
- Randomized offer of services to individuals ($Z_i$), 62% uptake: $P(D_i = 1 \mid Z_i = 1) = 0.62$. 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

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

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Method</th>
<th>0.15</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.85</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>Original controls</td>
<td>SEE ((\hat{h}))</td>
<td>57</td>
<td>381</td>
<td>1080</td>
<td>2630</td>
</tr>
<tr>
<td>Training</td>
<td>CH</td>
<td>-125</td>
<td>341</td>
<td>385</td>
<td>2557</td>
<td>3137</td>
</tr>
<tr>
<td>Training</td>
<td>tiny (h)</td>
<td>-129</td>
<td>500</td>
<td>381</td>
<td>2760</td>
<td>3114</td>
</tr>
<tr>
<td>Training</td>
<td>2SLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1593</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Modified controls</th>
<th>SEE ((\hat{h}))</th>
<th>74</th>
<th>398</th>
<th>1045</th>
<th>2748</th>
<th>2974</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>CH</td>
<td>-20</td>
<td>451</td>
<td>911</td>
<td>2577</td>
<td>3415</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>tiny (h)</td>
<td>-50</td>
<td>416</td>
<td>721</td>
<td>2706</td>
<td>3555</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>huge (h)</td>
<td>1568</td>
<td>1573</td>
<td>1582</td>
<td>1590</td>
<td>1595</td>
<td></td>
</tr>
<tr>
<td>Training</td>
<td>2SLS</td>
<td></td>
<td>1582</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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)
- $\beta$: discount factor ($1 = \text{no discount}$)
- $1/\gamma$: EIS, elasticity of intertemporal substitution
Quantile Euler equation estimates: UK

![Graph showing quantile Euler equation estimates for UK](image)
Quantile Euler equation estimates: USA

![Graph showing quantile Euler equation estimates for USA]
Quantile Euler equation estimates: Netherlands

![Graph showing quantile Euler equation estimates for Netherlands. The graph plots the estimates of \( \hat{\beta}(\tau) \) and \( 1/\hat{\gamma}(\tau) \) against quantile index \( \tau \). The estimates are represented by circles and triangles, respectively.]
Quantile Euler equation estimates: Sweden

\[ \hat{\beta}(\tau) \quad \Delta \quad \frac{1}{\hat{\gamma}(\tau)} \]

Quantile index \((\tau)\)

Estimate

0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0 1.5

Kaplan and: Sun, de Castro, Galvao, Liu

Smoothed IVQR & quantile Euler equations
## Quantile Euler equation estimates

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

*: significantly different from 1 at 10% level (2-sided)
Quantile Euler equation estimates

<table>
<thead>
<tr>
<th></th>
<th>US</th>
<th></th>
<th>UK</th>
<th></th>
<th>NTH</th>
<th></th>
<th>SWE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\beta}_T$</td>
<td>$\hat{\gamma}_T$</td>
<td>$\hat{\beta}_T$</td>
<td>$\hat{\gamma}_T$</td>
<td>$\hat{\beta}_T$</td>
<td>$\hat{\gamma}_T$</td>
<td>$\hat{\beta}_T$</td>
<td>$\hat{\gamma}_T$</td>
</tr>
<tr>
<td>$\tau$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.70</td>
<td>1.01</td>
<td>4.9</td>
<td>1.00</td>
<td>2.8</td>
<td>0.98</td>
<td>25.5</td>
<td>0.98</td>
<td>8.9</td>
</tr>
<tr>
<td>0.80</td>
<td>1.01</td>
<td>5.2</td>
<td>0.98</td>
<td>3.9</td>
<td>0.90</td>
<td>19.1</td>
<td>0.98</td>
<td>4.7</td>
</tr>
<tr>
<td>2SLS</td>
<td>1.08</td>
<td>16.7</td>
<td>1.03</td>
<td>6.0</td>
<td>0.96</td>
<td>−6.8</td>
<td>0.27</td>
<td>−544.4</td>
</tr>
</tbody>
</table>
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)

Thank you! (And further questions or comments are welcome)
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)
- Thank you!
- (And further questions or comments are welcome)


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

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

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


References VI


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

URL http://www.jstor.org/stable/2999619

URL http://www.jstor.org/stable/2238020

URL https://doi.org/10.1017/S0266466615000407
URL https://doi.org/10.1007/s10463-010-0310-9

URL https://doi.org/10.1002/jae.1069

URL https://doi.org/10.1016/j.jeconom.2007.01.014


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

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

URL https://doi.org/10.1016/0304-4076(84)90004-6
References XI


URL https://doi.org/10.1111/j.1467-937X.2009.00564.x

References XII

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


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