Smoothed IV quantile regression and quantile Euler equations

Luciano de Castro[1] Antonio F. Galvao[2] David M. Kaplan[3]

[1]University of Iowa [2]University of Arizona [3]University of Missouri

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

- Yesterday afternoon in Simon Hall 113:
  "Dynamic Quantile Models of Rational Behavior"

Quantile Euler equation: decision theory

- Yesterday afternoon in Simon Hall 113:
  "Dynamic Quantile Models of Rational Behavior"
- 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, $Q_\tau\big(\ln(W)\big) = \ln\big(Q_\tau(W)\big)$
- In contrast, $\mathbb{E}[\ln(W)] \le \ln\big(\mathbb{E}(W)\big)$ (Jensen's); approx error

Quantile Euler equation: estimation

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

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

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

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

- 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
- Before now: no feasible, theoretically justified IVQR estimator for time series data (or nonlinear structural model)

IV quantile regression (IVQR)

- So we can just run IVQR; but. . .
- Very difficult to compute IVQR estimator numerically
- Before now: no feasible, theoretically justified IVQR estimator for time series data (or nonlinear structural model)
- iid: Condition (i) on p. 310 in Chernozhukov and Hong (2003), Assumption 2.R1 in Chernozhukov and Hansen (2006), and Assumption 1 in Kaplan and Sun (2017)
- linear: (3.4) in Chernozhukov and Hansen (2006) and Assumption 1 in Kaplan and Sun (2017); Chernozhukov and Hong (2003) allows nonlinear, but computationally intensive MCMC
- We fill this gap, allowing non-iid data and nonlinear models, with fast, robust computation
- How? I'm glad you asked. Next slide, please—

Outline

Contribution: smoothed IVQR (SIVQR)

- 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

Contribution: smoothed IVQR (SIVQR)

- 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 quantile GMM (work in progress)

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 $\beta(\tau)$

IVQR moment conditions

(linear IV)

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

(linear IVQR)

IVQR moment conditions

(linear IV)

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

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

IVQR moment conditions

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

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

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

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

IVQR moment conditions

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

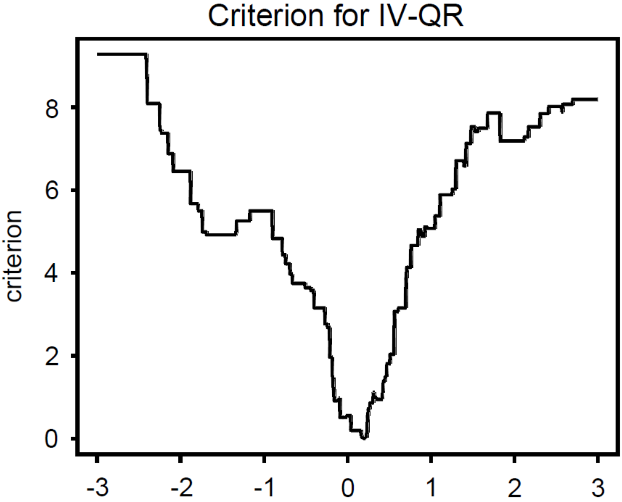
$$\mathbf{0} = \mathbb{E}[\mathbf{Z}(Y - \mathbf{X}'\boldsymbol{\beta}_0)],\ \boldsymbol{\beta}_0 = [\mathbb{E}(\mathbf{Z}\mathbf{X}')]^{-1}\,\mathbb{E}(\mathbf{Z}Y) \qquad \text{(linear IV)}$$

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

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

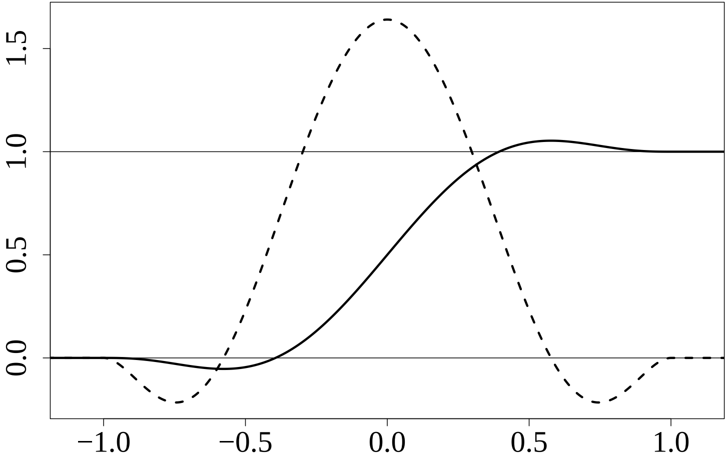
Just run GMM?

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



Smoothing the indicator function

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



Smoothed estimator

- Instead of solving sample moments $\mathbf{0} = \hat{\mathbb{E}}\big[\mathbf{Z}(\mathbb{1}\big\{\mathbf{X}'\hat{\boldsymbol{\beta}}_\tau - Y \geq 0\big\} - \tau)\big]$, replace $\mathbb{1}\{\cdot \geq 0\}$ with smoothed $\tilde{I}(\cdot)$: can compute Jacobian (wrt $\boldsymbol{\beta}$), fast and easy with standard solver

Outline

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

Brief comments on proofs

- 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

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

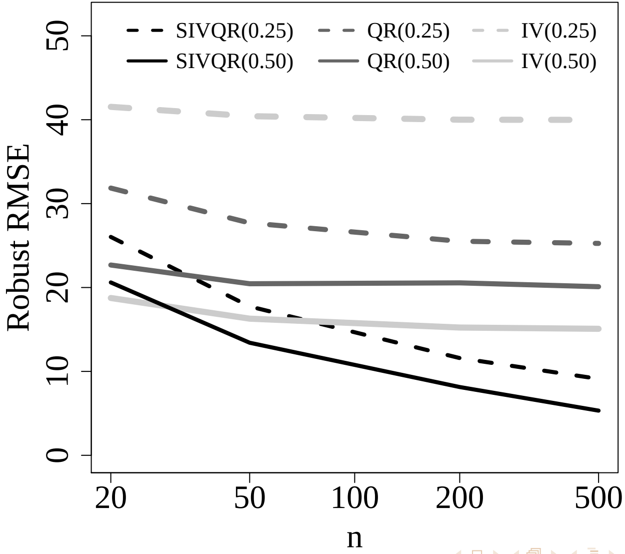
Simulation setup

- 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...)

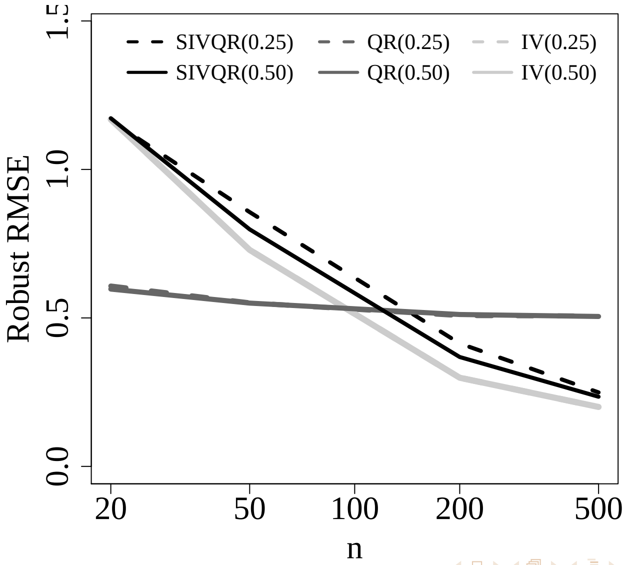
Simulation setup

- 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).

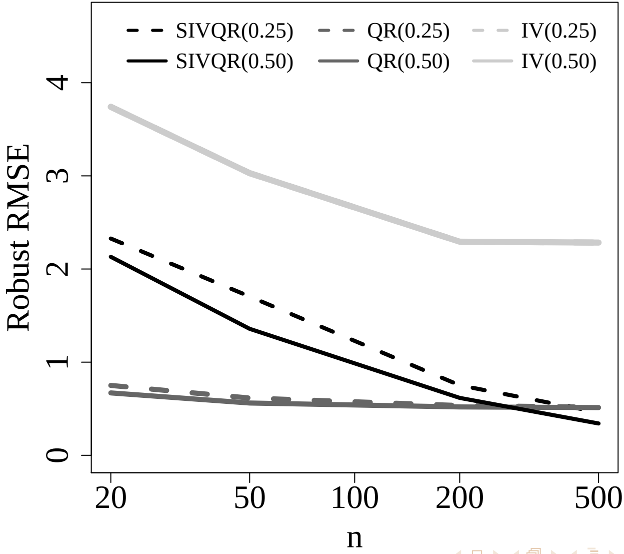
Robust RMSE: JTPA



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

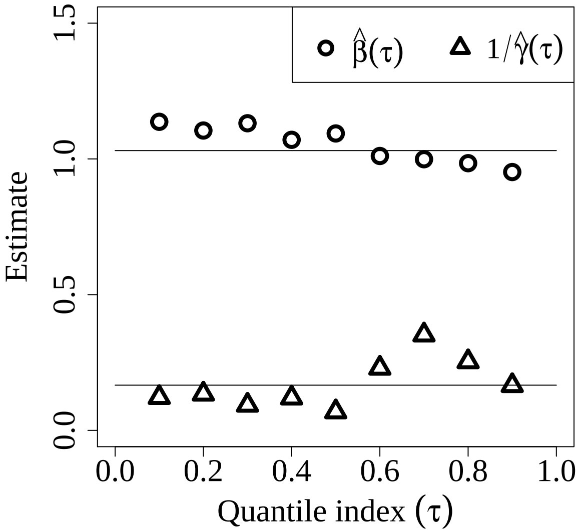
Size, 2-sided test of $H_0 : \gamma_\tau = \gamma_0$

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

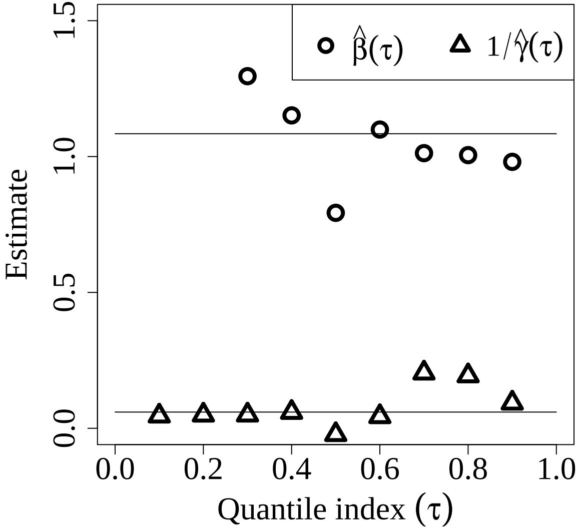
Quantile Euler equation estimates

- 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

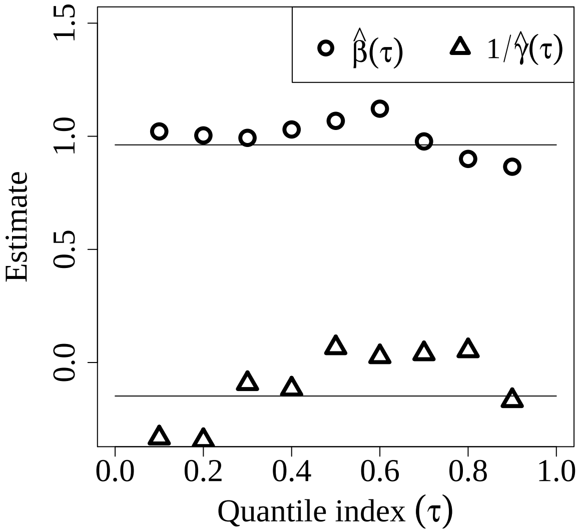
Quantile Euler equation estimates: UK



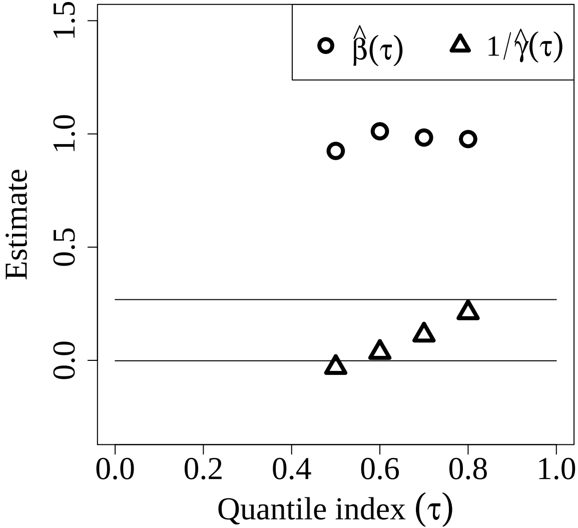
Quantile Euler equation estimates: USA



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)

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 et al. (2016) (in progress)? semi/nonparametric? uniformity in τ? determination of τ?
- Thank you!
- (And further questions or comments are welcome)

References I

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.

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

References II

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.

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.

References III

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.

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

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

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

References IV

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.

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

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

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

References V

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.