Instrumental Variables Quantile Regression for Panel Data with Measurement Errors

Antonio Galvao
University of Illinois

Gabriel Montes-Rojas
City University London

Tokyo, August 2009
Panel Data

- Panel data allows the possibility of following the same individuals over time, which allows to control for individual specific effects.
- Recently, semiparametric panel data models have attracted considerable interest in both theory and application since the error distribution is not specified.
- Honore and Lewbell (ECMT 2002), Koenker (Journal of Multivariate Analysis 2004), Geraci and Bottai (Biostatistics 2007), Abrevaya and Dahl (JBES 2008)
Quantile Regression Panel Data

- Koenker (JMA 2004) - **Quantile Regression for Panel Data**

- Conditional quantile functions of the response of the $y_{it}$

$$Q_{y_{it}}(\tau|\eta_i, x_{it}) = \eta_i + \beta(\tau)x_{it} \quad i = 1, ..., n, \quad t = 1, ..., T$$

- In some applications, it is of interest to explore a broad class of covariate effects (location and scale shifts), while still accounting for individual specific effects

- Such models enable to explore various forms of **heterogeneity** associated with the covariates under less stringent distributional assumptions
Measurement Error

- Economic quantities are frequently measured with error, particularly if longitudinal information is collected through one-time retrospective surveys, which are notoriously susceptible to recall errors.

- If the regressors are indeed subject to measurement errors, it is well known that the slope coefficient of the least squares regression estimator is biased → attenuation bias.

- The bias may be exacerbated in panel data models.

- This topic has also attracted considerable attention in the quantile regression literature: Chesher (2001) and Schennach (ET 2008).
IV solution for ME in Panel Data

- There is an extensive literature that use IV strategies to solve the ME bias
- See for instance Griliches and Hausman (JoE 1986), Hsiao (1992, 2003), Wansbeek (JoE 2001), Biorn (ER 2000)
- The IV approach employs lagged (or lagged differences of the) regressors as instruments for the mismeasured variable
Our Contribution

- We show that measurement errors produce a similar attenuation bias as in OLS, by analytically deriving an approximation to the bias in QR.
- Propose an instrumental variables strategy to estimate fixed effects quantile regression panel data with measurement error and reduce the bias.
  - The estimator uses lagged dependent observations (or lagged differences) as instruments.
- We show consistency and asymptotic normality of the estimator, provided $N^a/T \to 0$, for some $a > 0$. 
Our Contribution

- Conduct Monte Carlo experiments to show that, even in short panels, the proposed estimators can substantially reduce the bias
- Propose Wald and Kolmogorov-Smirnov tests for general linear hypothesis, and derive the respective limiting distributions
- Illustrate the new approach to the Tobin’s $q$ theory of investment
Consider the following representation of a panel data model with individual fixed effects and measurement errors

\[ y_{it} = d_{it}' \eta + x_{it}^* \beta + u_{it} \quad i = 1, \ldots, N; \quad t = 1, \ldots, T. \]  \hspace{1cm} (1)

Suppose that we do not observe \( x_{it}^* \), but rather \( x_{it} \), which is a noisy measure of \( x_{it}^* \) subject to an additive measurement error \( \epsilon_{it} \),

\[ x_{it} = x_{it}^* + \epsilon_{it}. \]  \hspace{1cm} (2)

Using equation (2) we can express (1) in terms of the observed \( y \)'s and \( x \)'s as

\[ y_{it} = d_{it}' \eta + x_{it}' \beta + u_{it} - \epsilon_{it}' \beta \quad i = 1, \ldots, N; \quad t = 1, \ldots, T. \]  \hspace{1cm} (3)
Measurement Errors in OLS – Panel Data (Cont.)

- Define $v^* = [d', x^*]', \Lambda_\epsilon = [0', \epsilon']'$, $v = [d', x']' = v^* + \Lambda_\epsilon$ and $\varphi = [\eta', \beta']'$

- Two estimators:

  $$\varphi^* = \text{argmin}_\varphi E[y - v^* \varphi]^2,$$

  and

  $$\varphi^\circ = \text{argmin}_\varphi E[y - v' \varphi]^2.$$  

- Effect of ME (signal to noise-ratio)

  $$\varphi^\circ = \varphi^* - \left( E \left[ v^* v^*' + \Lambda_\epsilon \Lambda_\epsilon' \right] \right)^{-1} E[\Lambda_\epsilon \epsilon' \beta^*]$$
Consider now the $\tau$th conditional quantile function of the response $y$,

$$Q_y(\tau|d, x^*) = d'\eta^*(\tau) + x'^*\beta^*(\tau) \quad (6)$$

Define $v^* = [d', x'^*]'$, $\Lambda_\epsilon = [0', \epsilon']'$, $v = [d', x']' = v^* + \Lambda_\epsilon$ and $\varphi = [\eta', \beta']'$

Two estimators:

$$\varphi^*(\tau) = \arg\min_{\varphi} E[\rho_{\tau}(y - v^*'\varphi)] \quad (7)$$

where $\rho_{\tau}(u) := u(\tau - I(u < 0))$ as in Koenker and Bassett (ECMT 1978), and

$$\varphi^\circ(\tau) = \arg\min_{\varphi} E[\rho_{\tau}(y - v'\varphi)] \quad (8)$$
The following Lemma shows that the measurement error bias in QR can be approximated to an expression similar to that in OLS, using the Angrist, Chernozhukov, and Fernandez-Val (ECMT 2006) approximation.

**Lemma 1.** Assume that: (i) the conditional density function \( f_y(y|v, \epsilon) \) exists and is bounded a.s.; (ii) \( E[y], E[Q_y(\tau|v, \epsilon)^2] \), and \( E\|v', \epsilon'\|^2 \) are finite; (iii) \( \varphi^*(\tau) \) and \( \varphi^\circ(\tau) \) uniquely solve equations (7) and (8) respectively; (iv) \( \epsilon \) is independent of \((d, x^*, u)\). Then,

\[
\varphi^\circ(\tau) = \varphi^*(\tau) - \left( \frac{1}{E[\omega_\tau(v, \epsilon) \cdot (vv')]} \right)^{-1} E[\omega_\tau(v, \epsilon) \cdot v \epsilon' \beta^*(\tau)]
\]

where

- \( \omega_\tau(v, \epsilon) = \int_0^1 f_u(\tau) (u \cdot \Delta_\tau(v, \epsilon; \varphi^\circ(\tau)))|v, \epsilon\) \( du/2 \) is a weighting function, and
- \( \Delta_\tau(v, \epsilon; \varphi^\circ(\tau)) = v' \cdot (\varphi^\circ(\tau) - \varphi^*(\tau))' + \epsilon' \beta^\circ(\tau) \) is the quantile regression specification error.
In particular, under some boundedness conditions:

\[ \varphi^*(\tau) \approx \varphi^*(\tau) - \left( E \left[ f_y(q) (v^* v^{*'} + \Lambda \epsilon \Lambda') \right] \right)^{-1} E \left[ f_y(q) \Lambda \epsilon' \beta^*(\tau) \right] \]

where \( f_y(q) = f_y(Q_\tau(y|v^*)|v^*) \)
Digression on IV for Quantile Regression

- Consider the following quantile regression model,

\[ Q_Y(\tau|X) = X\beta(\tau) \]

where \( Y \) is the outcome variable conditional on the exogenous variables of interest \( X \)

- Then note that

\[ Y = X\beta(U) \quad U|X \sim U(0,1) \]

- Koenker and Bassett (ECMT 1978) show that

\[ \hat{\beta} = \arg\min_{\beta} \sum \rho_\tau(y - x'\beta) \]

where \( \rho_\tau(u) = u(\tau - I(u < 0)) \)
Digression on IV for Quantile Regression (Cont.)

- $\hat{\beta}(\tau)$ based on the moment condition $\psi_{\tau}(R) \perp X$, where $R$ is the residual $Y - X\beta$ and $\psi_{\tau}(u) = \tau - I(u < 0)$
- Suppose that $\psi_{\tau}(R) \perp X$ does not hold
- But we can find $\psi_{\tau}(R) \perp W$
- We want to estimate $\beta$
- But now

$$Y = X\beta(U) \quad U|W \sim U(0, 1)$$
Digression on IV for Quantile Regression (Cont.)

- $W$ does not belong to the model
- Thus, for fixed $\beta$, in the quantile regression of $(Y - X\beta)$ on $W$, $W$ should have coefficient of zero
- Estimator is defined as:

$$\hat{\beta} = \arg\min_\beta \| \hat{\gamma}(\beta) \|_A$$

where

$$\hat{\gamma}(\beta) = \arg\min_\gamma \sum \rho_\tau(y - x'\beta - w'\gamma)$$

- $\hat{\beta}(\tau)$ that makes $\hat{\gamma}(\tau) \approx 0$ is the instrumental variables estimator
Now we consider a finite-sample analog of the above procedure for panel data

\[ Q_y(\tau|\eta, x^*, z) = d'\eta(\tau) + x^{'\prime}\beta(\tau) + z'\alpha(\tau), \]

Define

\[ Q_{NT} := \sum_{i=1}^{N} \sum_{t=1}^{T} \rho_\tau \left( y_{it} - d'_{it}\eta(\tau) - x'_it\beta(\tau) - z'_it\alpha(\tau) - w'_it\gamma(\tau) \right) \]
The IVQRFE is defined as follows:

For a given value of the structural parameter, say $\beta$, one estimates the panel QR to obtain

$$(\hat{\eta}(\beta, \tau), \hat{\beta}(\beta, \tau), \hat{\alpha}(\beta, \tau), \hat{\gamma}(\beta, \tau)) := \arg\min_{\eta, \alpha, \gamma} Q_{NT}(\tau, \eta, \beta, \alpha, \gamma)$$

To find an estimate for $\beta(\tau)$, we look for a value $\beta$ that makes the coefficient on the instrumental variable $\gamma(\beta, \tau)$ as close to 0 as possible

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \mathcal{B}} \| \hat{\gamma}(\beta, \tau) \|_A$$
Asymptotics for IVQRFE – Assumptions

**A1** The $y_{it}$ are independent with conditional distribution functions $F_{it}$, and differentiable conditional densities, $0 < f_{it} < \infty$, with bounded derivatives $f'_{it}$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T;
Asymptotics for IVQRFE – Assumptions (Cont.)

A2 Let \( D = I_N \otimes \nu_T \), and \( \nu_T \) a \( T \)-vector of ones, \( X = (x_{it}) \) be a \( NT \times \text{dim}(\beta) \) matrix, \( Z = (z_{it}) \) be a \( NT \times \text{dim}(\alpha) \) matrix, and \( W = (w_{it}) \) be a \( NT \times \text{dim}(\gamma) \) matrix. For

\[
\Pi(\eta, \beta, \alpha, \tau) := E[(\tau - 1(D'\eta + X'\beta + Z'\alpha))\tilde{X}]
\]

\[
\Pi(\eta, \beta, \alpha, \gamma, \tau) := E[(\tau - 1(D'\eta + X'\beta + Z'\alpha + W'\gamma))\tilde{X}]
\]

\( \tilde{X} := [D', Z', W']' \),

Jacobian matrices \( \frac{\partial}{\partial(\eta,\beta,\alpha)} \Pi(\eta, \beta, \alpha, \tau) \) and \( \frac{\partial}{\partial(\eta,\alpha,\gamma)} \Pi(\eta, \beta, \alpha, \gamma, \tau) \)

are continuous and have full rank, uniformly over \( \mathcal{E} \times \mathcal{B} \times \mathcal{A} \times \mathcal{G} \times \mathcal{I} \) and the image of \( \mathcal{E} \times \mathcal{B} \times \mathcal{A} \) under the map \( (\eta, \beta, \alpha) \rightarrow \Pi(\eta, \beta, \alpha, \tau) \) is simply connected;
Asymptotics for IVQRFE – Assumptions (Cont.)

\[ \textbf{A3} \] Denote \( \Phi(\tau) = \text{diag}(f_{it}(\xi_{it}(\tau))) \), where
\[
\xi_{it}(\tau) = d_{it}'\eta(\tau) + x_{it}'\beta(\tau) + z_{it}'\alpha(\tau) + w_{it}\gamma(\tau),
\]
\( M_D = I - P_D \) and \( P_D = D(D'\Phi(\tau)D)^{-1}D'\Phi(\tau) \). Let \( \tilde{X} = [W, Z]' \). Then, the following matrix is invertible:

\[
J_{\alpha\gamma} = E(\tilde{X}'M_D\Phi(\tau)M_D\tilde{X});
\]

Now define \([\bar{J}_\alpha', \bar{J}_\gamma']'\) as a partition of \( J_{\alpha\gamma}^{-1} \),
\[
J_\beta = E(\tilde{X}'M_D\Phi(\tau)M_DX) \quad \text{and} \quad H = \bar{J}_\gamma'A[\beta(\tau)]\bar{J}_\gamma.
\]
Then, \( J_\beta' H J_\beta \) is also invertible;
Asymptotics for IVQRFE – Assumptions (Cont.)

**A4** For all \( \tau \), \((\beta(\tau), \alpha(\tau))\) ∈ int \( \mathcal{B} \times \mathcal{A} \), and \( \mathcal{B} \times \mathcal{A} \) is compact and convex;

**A5** \( \max_{it} \| x_{it} \| = O(\sqrt{NT}) \); \( \max_{it} \| z_{it} \| = O(\sqrt{NT}) \);
\( \max_{it} \| w_{it} \| = O(\sqrt{NT}) \);

**A6** \( \frac{N^a}{T} \rightarrow 0 \), for some \( a > 0 \).
Asymptotics for IVQRFE – Results

Let $\theta(\tau) = (\beta(\tau), \alpha(\tau))$

**Theorem 1** Given assumptions A1-A6, $(\eta(\tau), \beta(\tau), \alpha(\tau))$ uniquely solves the equations $E[\psi_{\tau}(Y - D'\eta - X'\beta - Z'\alpha)\tilde{X}] = 0$ over $\mathcal{E} \times \mathcal{B} \times \mathcal{A}$, and $\varphi(\tau) = (\eta(\tau), \beta(\tau), \alpha(\tau))$ is consistently estimable.
Asymptotics for IVQRFE – Results (Cont.)

**Theorem 2 (Asymptotic Normality)**

Under conditions A1-A6, for a given \( \tau \in (0, 1) \), \( \hat{\theta}(\tau) \) converges to a Gaussian distribution as

\[
\sqrt{NT}(\hat{\theta}(\tau) - \theta(\tau)) \xrightarrow{d} N(0, \Omega(\tau)), \quad \Omega(\tau) = (K', L')' S(K', L')
\]

where \( S = \tau(1 - \tau)E[VV'] \), \( V = \tilde{X}'M_D \), \( K = (J'_\beta H J_\beta)^{-1} J_\beta H \),
\( H = \tilde{J}'_\gamma A[\beta(\tau)] \tilde{J}_\gamma \), \( L = \tilde{J}_\alpha M \), \( M = I - J_\beta K \),
\( J_\beta = E(\tilde{X}'M_D \Phi(\tau) M_D X) \), \( [\tilde{J}'_\alpha, \tilde{J}'_\gamma]' \) is a partition of
\( J^{-1}_{\alpha\gamma} = (E(\tilde{X}'M_D \Phi M_D \tilde{X}))^{-1} \), \( \Phi(\tau) = diag(f_{it}(\xi_{it}(\tau))) \), and
\( \tilde{X} = [Z, W]' \).
Inference for IVQRFE

- Test the hypothesis $R\theta(\tau) = r$, when $r$ is known
- Under the linear hypothesis $H_0 : R\theta(\tau) = r$, assumptions A1-A6, we have

$$\mathcal{V}_n = \sqrt{NT}[R\Sigma R']^{-1/2}(R\hat{\theta} - r) \Rightarrow B_q(\tau),$$

where $B_q(\tau)$ represents a $q$-dimensional standard Brownian Bridge
- Thus, for given $\tau$, the regression Wald process can be constructed as

$$\mathcal{W}_n = NT(R\hat{\theta} - r)'[R\hat{\Omega} R']^{-1}(R\hat{\theta} - r)$$

where $\hat{\Omega}$ is a consistent estimators of $\Omega$
Inference for IVQRFE (Cont.)

- **Theorem 3** (Wald Test Inference). Under $H_0 : R\theta(\tau) = r$ and conditions A1-A6, for fixed $\tau$,

  $\mathcal{W}_n(\tau) \xrightarrow{a} \chi^2_q$.

- **Theorem 4** (Kolmogorov-Smirnov test). Under $H_0$ and conditions A1-A6,

  $$ KS\mathcal{W}_n = \sup_{\tau \in T} \mathcal{W}_n(\tau) \Rightarrow \sup_{\tau \in T} Q^2_q(\tau). $$

- Critical values for $\sup Q^2_q(\tau)$ have been tabled by DeLong (1981) and, more extensively, by Andrews (1993) using simulation methods.
Monte Carlo - Description

- Evaluate the finite sample performance of the quantile regression instrumental variables estimator
- Model:
  \[ y_{it} = \eta_i + x_{it}'\beta + z_{it}'\alpha + u_{it}, \]
- Two schemes to generate the disturbances \( u_{it} \)
  - \( u_{it} \sim N(0, \sigma^2_u) \)
  - \( u_{it} \sim t_3 \)
In both cases we have an additive measurement error of the form

\[ x_{it} = x^*_t + \epsilon_{it}, \]

where \( x^*_t \) follows an ARMA(1,1) process

\[ (1 - \phi L) x^*_t = \mu_i + \varepsilon_{it} + \theta \varepsilon_{it-1} \]

and \( \epsilon_{it} \) follows the same distribution as \( u_{it} \), that is, normal distribution and \( t_3 \).
The fixed effects, $\mu_i$ and $\alpha_i$, are generated as

$$\mu_i = e_{1i} + T^{-1} \sum_{t=1}^{T} \epsilon_{it}, \quad e_{1i} \sim N(0, \sigma_{e_1}^2),$$

$$\eta_i = e_{2i} + T^{-1} \sum_{t=1}^{T} x_{it}, \quad e_{2i} \sim N(0, \sigma_{e_2}^2).$$

In the simulations, we experiment with $T = 10$ and $N = 100$. 1000 replications

Consider the following values for the remaining parameters:

$$(\beta, \alpha) = (1, 1);$$

$$\phi = 0.6, \quad \theta = 0.7, \quad \gamma = 1, \quad \sigma_u^2 = \sigma_{e_1}^2 = \sigma_{e_2}^2 = 1.$$
### Location Shift Model: Bias and RMSE of estimators for normal distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$</td>
<td>Bias</td>
<td>-0.195</td>
<td>0.006</td>
<td>-0.197</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.1957</td>
<td>0.0572</td>
<td>0.1998</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>Bias</td>
<td>0.002</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0362</td>
<td>0.0462</td>
<td>0.0447</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for normal distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$</td>
<td>Bias</td>
<td>$-0.195$</td>
<td>0.006</td>
<td>$-0.197$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.1957</td>
<td>0.0572</td>
<td>0.1998</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>Bias</td>
<td>0.002</td>
<td>0.001</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>0.0362</td>
<td>0.0462</td>
<td>0.0447</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for normal distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$ Bias</td>
<td>$-0.195$</td>
<td>$0.006$</td>
<td>$-0.197$</td>
<td>$0.008$</td>
</tr>
<tr>
<td>RMSE</td>
<td>$0.1957$</td>
<td>$0.0572$</td>
<td>$0.1998$</td>
<td>$0.0736$</td>
</tr>
<tr>
<td>$\alpha = 1$ Bias</td>
<td>$0.002$</td>
<td>$0.001$</td>
<td>$0.003$</td>
<td>$0.001$</td>
</tr>
<tr>
<td>RMSE</td>
<td>$0.0362$</td>
<td>$0.0462$</td>
<td>$0.0447$</td>
<td>$0.0475$</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for normal distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$ Bias</td>
<td>$-0.195$</td>
<td>$0.006$</td>
<td>$-0.197$</td>
<td>$0.008$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.1957$</td>
<td>$0.0572$</td>
<td>$0.1998$</td>
</tr>
<tr>
<td>$\alpha = 1$ Bias</td>
<td>$0.002$</td>
<td>$0.001$</td>
<td>$0.003$</td>
<td>$0.001$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.0362$</td>
<td>$0.0462$</td>
<td>$0.0447$</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for normal distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$ Bias</td>
<td>-0.195</td>
<td>0.006</td>
<td>-0.197</td>
<td>0.008</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.1957</td>
<td>0.0572</td>
<td>0.1998</td>
<td>0.0736</td>
</tr>
<tr>
<td>$\alpha = 1$ Bias</td>
<td>0.002</td>
<td>0.001</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0362</td>
<td>0.0462</td>
<td>0.0447</td>
<td>0.0475</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for normal distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$ Bias</td>
<td>−0.195</td>
<td>0.006</td>
<td>−0.197</td>
<td>0.008</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.1957</td>
<td>0.0572</td>
<td>0.1998</td>
<td>0.0736</td>
</tr>
<tr>
<td>$\alpha = 1$ Bias</td>
<td>0.002</td>
<td>0.001</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0362</td>
<td>0.0462</td>
<td>0.0447</td>
<td>0.0475</td>
</tr>
</tbody>
</table>
## Location Shift Model: Bias and RMSE of estimators for $t_3$ distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$</td>
<td>Bias</td>
<td>$-0.201$</td>
<td>$-0.004$</td>
<td>$-0.151$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.2075$</td>
<td>$0.0619$</td>
<td>$0.1531$</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>Bias</td>
<td>$0.001$</td>
<td>$0.001$</td>
<td>$-0.001$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.0396$</td>
<td>$0.0471$</td>
<td>$0.0342$</td>
</tr>
</tbody>
</table>

Galvao, Montes-Rojas

Instrumental Variables Quantile Regression for Panel Data with Measurement Errors
Location Shift Model: Bias and RMSE of estimators for $t_3$ distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$</td>
<td>Bias</td>
<td>$-0.201$</td>
<td>$-0.004$</td>
<td>$-0.151$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.2075$</td>
<td>$0.0619$</td>
<td>$0.1531$</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>Bias</td>
<td>$0.001$</td>
<td>$0.001$</td>
<td>$-0.001$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.0396$</td>
<td>$0.0471$</td>
<td>$0.0342$</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for $t_3$ distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$</td>
<td>Bias</td>
<td>$-0.201$</td>
<td>$-0.004$</td>
<td>$-0.151$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.2075$</td>
<td>$0.0619$</td>
<td>$0.1531$</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>Bias</td>
<td>$0.001$</td>
<td>$0.001$</td>
<td>$-0.001$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.0396$</td>
<td>$0.0471$</td>
<td>$0.0342$</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for $t_3$ distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$ Bias</td>
<td>-0.201</td>
<td>-0.004</td>
<td>-0.151</td>
<td>0.009</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.2075</td>
<td>0.0619</td>
<td>0.1531</td>
<td>0.0542</td>
</tr>
<tr>
<td>$\alpha = 1$ Bias</td>
<td>0.001</td>
<td>0.001</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.0396</td>
<td>0.0471</td>
<td>0.0342</td>
<td>0.0363</td>
</tr>
</tbody>
</table>
Location Shift Model: Bias and RMSE of estimators for $t_3$ distribution ($T = 10$ and $N = 100$)

<table>
<thead>
<tr>
<th></th>
<th>FE</th>
<th>IVFE</th>
<th>QRFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta = 1$</td>
<td>Bias</td>
<td>$-0.201$</td>
<td>$-0.004$</td>
<td>$-0.151$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.2075$</td>
<td>$0.0619$</td>
<td>$0.1531$</td>
</tr>
<tr>
<td>$\alpha = 1$</td>
<td>Bias</td>
<td>$0.001$</td>
<td>$0.001$</td>
<td>$-0.001$</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>$0.0396$</td>
<td>$0.0471$</td>
<td>$0.0342$</td>
</tr>
</tbody>
</table>
Empirical application: Investment, Tobin’s $q$ and cash flow

- Tobin’s $q$ is the ratio of the market valuation of a firm and the replacement value of its assets.
- Firms with a high value of $q$ are considered attractive as to the investment opportunities, whereas a low value of $q$ indicates the opposite.
- Since the operationalization of $q$ is not clear-cut and unambiguous, estimation poses a measurement error problem.
- Many empirical investment studies found a very disappointing performance of the $q$ theory of investment, although this theory has a good performance when measurement error is purged as in Erickson and Whited (JPE 2000).
Empirical application: Investment, Tobin’s $q$ and cash flow

- In parallel, investment theory is also interested in the effect of cash flow, as the theory predicts that financially constrained firms are more likely to rely on internal funds to finance investment.
- Alti (JF 2003), Almeida, Campello and Weisbach (JF 2004)
- However, Erickson and Whited (2000) argue that cash flow has no effect on investment once measurement error in Tobin’s $q$ is taken into account.
Empirical application: Investment, Tobin’s $q$ and cash flow

- QR is useful to describe differences in investment ratios across firms in the presence of unobserved heterogeneity
- Unobserved characteristics (for the econometrician) across firms may be due to several reasons:
  1. The ability of managers (insiders) or the characteristics of investors (outsiders) (Bertrand and Schoar, QJE 2003; Walentin and Lorenzoni, NBER 2007)
  2. Firms’ capital structure (Chirinko, JEDC 1993)
Main hypothesis

In the upper conditional quantiles of investment, investment may be driven by insiders knowledge of business opportunities or a particular capital structure that requires more investment, and therefore, we expect that investment would be more responsive to changes in $q$ and cash flow, as the firms would use all available resources to finance its projects.
The baseline model in the literature is

\[ \frac{I_{it}}{K_{it}} = \eta_i + \beta q^*_{it} + \alpha \frac{CF_{it}}{K_{it}} + u_{it}, \]

where \( I \) denotes investment, \( K \) capital stock, \( CF \) cash flow, \( q^* \) well-measured Tobin’s \( q \), \( \eta \) is the firm-specific effects and \( u \) is the innovation term. Our objective is estimating the following conditional quantile function:

\[ Q_{\frac{I}{K}}(\tau | \eta, CF/K, q) = \eta_i(\tau) + \beta(\tau)q^*_it + \alpha(\tau)CF_{it}/K_{it}. \]
Data

- We follow Almeida, Campello and Weisbach (JF 2004) approach by considering a sample of manufacturing firms (SICs 2000 to 3999) over the 1980 to 2005 period with data available from COMPUSTAT’s P/S/T, Full Coverage.

- Only firms with observations in every year are used, in order to construct a balanced panel of firms for the 26 year period (the choice of a balanced panel is made to reduce the computational burden).

- Moreover, following those authors we eliminate firms for which cash-holdings exceeded the value of total assets and those displaying asset or sales growth exceeding 100%.

- Our final sample consists of 4550 firm-years and 175 firms.

Galvao, Montes-Rojas
Instrumental Variables Quantile Regression for Panel Data with Measurement Errors
<table>
<thead>
<tr>
<th>Variable</th>
<th>FE</th>
<th>QRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q$</td>
<td>0.087</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>$CF/K$</td>
<td>0.095</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>$\Delta q_t^{-1}$</td>
<td>0.282</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>$CF/K$</td>
<td>0.074</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Inst.</td>
<td>$\Delta q_t^{-1}$</td>
<td>$\Delta q_t^{-1}$</td>
</tr>
<tr>
<td>$q$</td>
<td>0.203</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>$CF/K$</td>
<td>0.083</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Inst.</td>
<td>$\Delta q_t^{-1}, \Delta q_t^{-2}$</td>
<td>$\Delta q_t^{-1}, \Delta q_t^{-2}$</td>
</tr>
</tbody>
</table>
Figure: FE and QRFE

QRFE Tobin's Q

QRFE Cash Flow

Galvao, Montes-Rojas
Instrumental Variables Quantile Regression for Panel Data with Measurement Errors
### Motivation

- Model
- IVQRFE Estimator
- Asymptotics for IVQRFE
- Monte Carlo
- Application
- Conclusions

### Table

<table>
<thead>
<tr>
<th>Variable</th>
<th>FE</th>
<th>QRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>τ</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( q )</td>
<td>0.087</td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>( CF/K )</td>
<td>0.095</td>
<td>0.060</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>IVFE</th>
<th>IVQRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q )</td>
<td>0.282</td>
<td>0.144</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>( CF/K )</td>
<td>0.074</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Inst.</td>
<td>( \Delta q_{t-1} )</td>
<td>( \Delta q_{t-1} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>FE</th>
<th>QRFE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>τ</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( q )</td>
<td>0.203</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>( CF/K )</td>
<td>0.083</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Inst.</td>
<td>( \Delta q_{t-1}, \Delta q_{t-2} )</td>
<td>( \Delta q_{t-1}, \Delta q_{t-2} )</td>
</tr>
</tbody>
</table>
Motivation
Model
IVQRFE Estimator
Asymptotics for IVQRFE
Monte Carlo
Application
Conclusions

Figure: IVFE and IVQRFE (Instrument $\Delta q_{t-1}$)

IVQRFE Tobin's Q

IVQRFE Cash Flow

Galvao, Montes-Rojas
Instrumental Variables Quantile Regression for Panel Data with Measurement Errors
Figure: IVFE and IVQRFE (Instruments $\Delta q_{t-1}, \Delta q_{t-2}$)

IVQRFE (DL1x DL2x) Tobin’s Q

IVQRFE (DL1x DL2x) Cash Flow

Galvao, Montes-Rojas
Instrumental Variables Quantile Regression for Panel Data with Measurement Errors
Conclusions

- Propose an instrumental variables quantile regression estimator for panel data that solves measurement error problems
- Show consistency and asymptotic normality of the estimators
- Propose a Wald and Kolmogorov-Smirnov tests for linear hypotheses
- Apply the estimator and test to Tobin’s $q$ theory of investment