(Continuous) GMM Quantile Regression

Sergio Firpo[1] Antonio Galvao[2] Cristine Pinto[3]
Alexandre Poirier[4] and Graciela Sanroman[5]

[1]Insper, [2]Univ. Arizona, [3]EESP, FGV, [4]Georgetown Univ., [5]UDELAR

IESTA, FCEA

30 de Abril 2019

Outline

Motivation

- Koenker and Bassett (1978) introduced quantile regression (QR) models.
- By estimating conditional quantile functions, QR provides insight into heterogeneous effects of policy variables.
- This is especially valuable for program evaluation studies, where these methods help analyze how treatments or social programs affect the outcome's distribution.
- Since the work of Matkin (2003) QR has become a natural way to represent a structural relationship.

Motivation: Traditional Approach

- The Conditional Quantile Function (CQF) is defined as:

$$Q_\tau\left(Y|X\right) \equiv \inf\left\{y : F_y\left(y|X\right) \geq \tau\right\}.$$

- The CQF solves the following minimization problem:

$$Q_\tau\left(Y|X\right) \in \arg\min_{h(X)} \mathrm{E}\left[\rho_\tau\left(Y - h\left(X\right)\right)\right],$$

  where $\rho_\tau\left(u\right) = u \cdot \left(\tau - \mathbf{1}\left\{u \leq 0\right\}\right)$.
- Usually, we work with the linear quantile regression (QR):

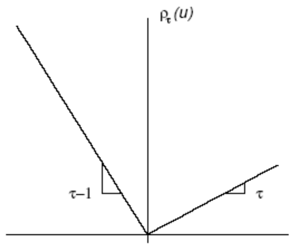
$$Q_\tau\left(Y|X\right) \in \arg\min_{\beta\in\mathbb{R}^d} E\left[\rho_\tau\left(Y - X^\top\beta(\tau)\right)\right].$$

Quantile Regression

- Conditional Quantile: $Q_\tau(Y|X) = X^\top \beta(\tau)$

$$\min_{\beta} \sum_{i=1}^{n} \rho_\tau(Y_i - X_i^\top \beta(\tau))$$

where $\rho_\tau(u) = u \cdot (\tau - \mathbf{1}\{u \leq 0\})$



- The solution for the QR problem is given by the following conditional moment restriction :

$$\mathrm{E}\left[\tau - 1\left\{Y - X^\top \beta(\tau) \leq 0\right\}\Big| X\right] = 0,$$

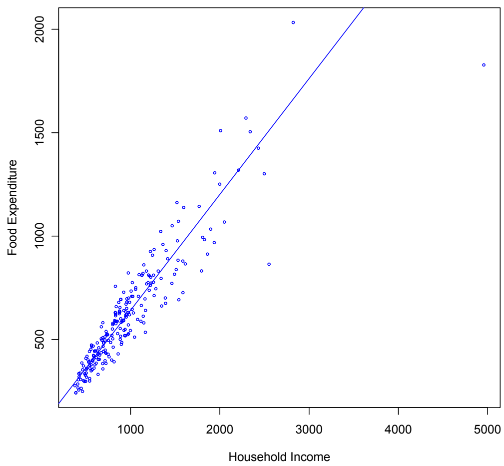
Quantile Regression – Simple Example

- Engle curve for food expenditure

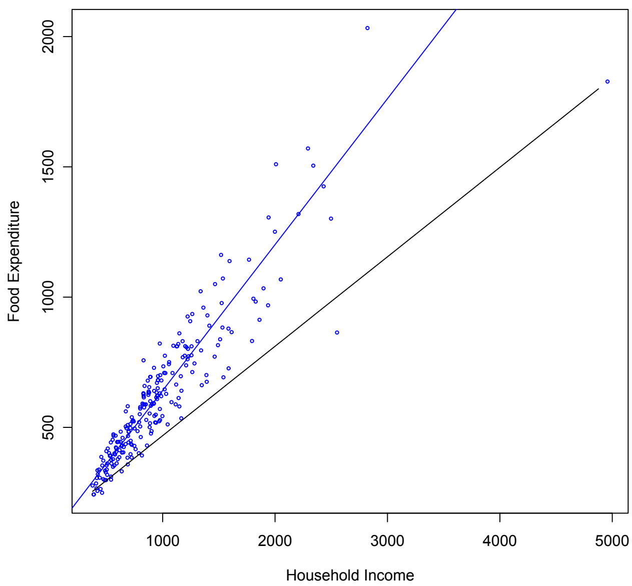
$$Q_\tau(Y|X) = X^\top \beta(\tau)$$

- Y is the household food expenditure,
- X is the household income.

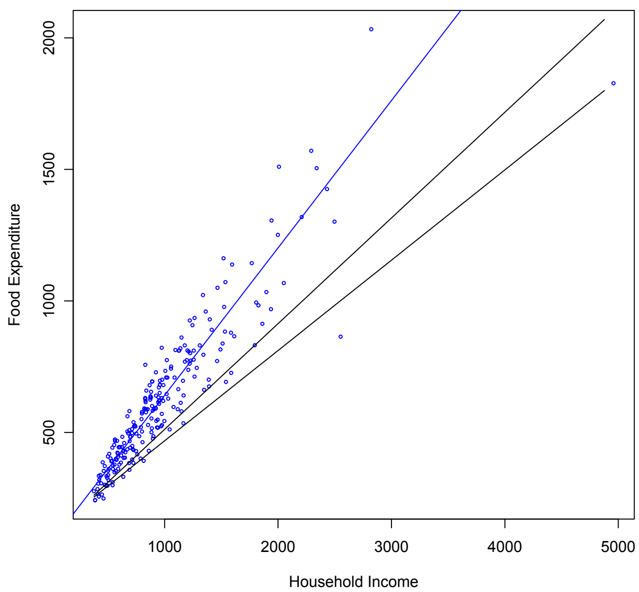
Estimated Conditional Quantiles of Engel Data



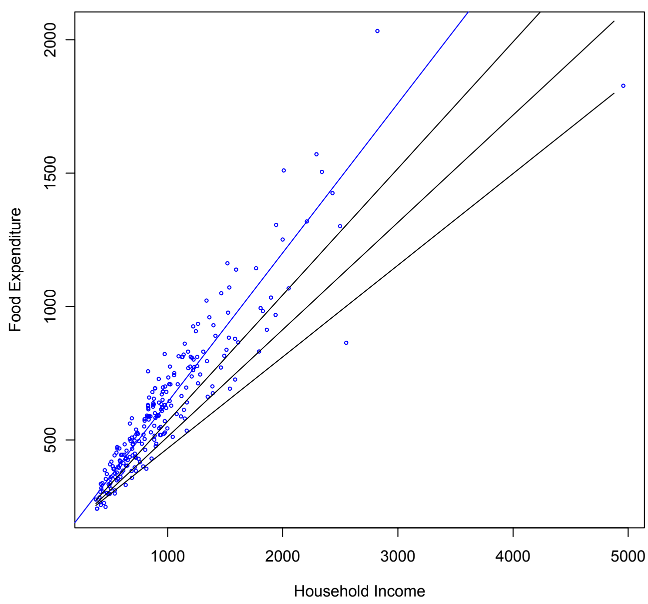
Estimated Conditional Quantiles of Engel Data



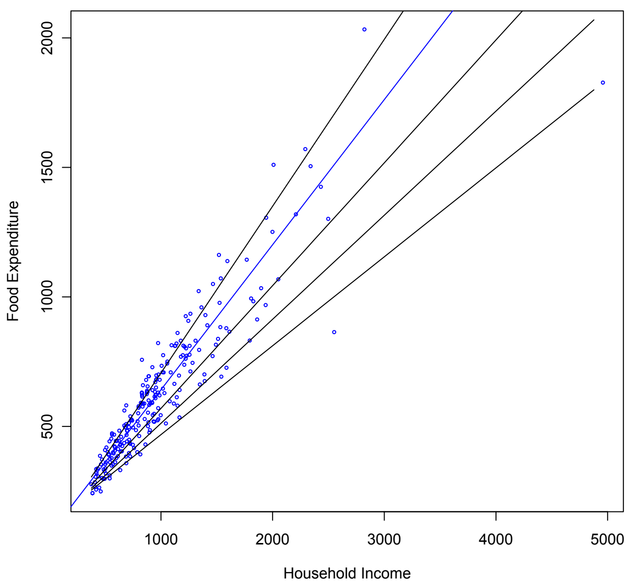
Estimated Conditional Quantiles of Engel Data



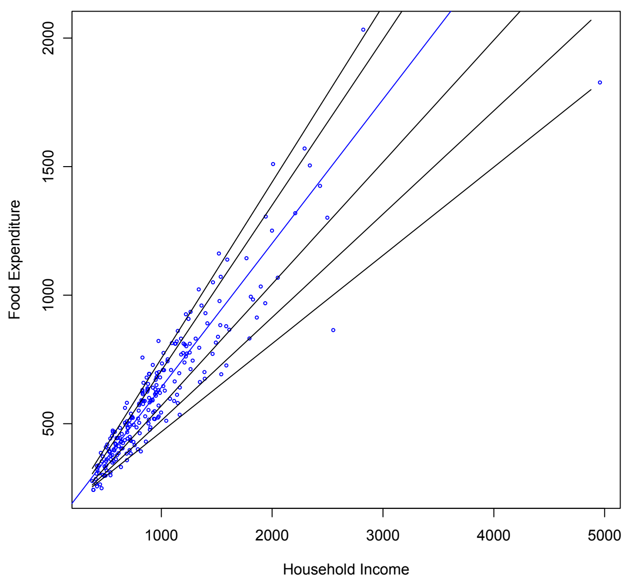
Estimated Conditional Quantiles of Engel Data



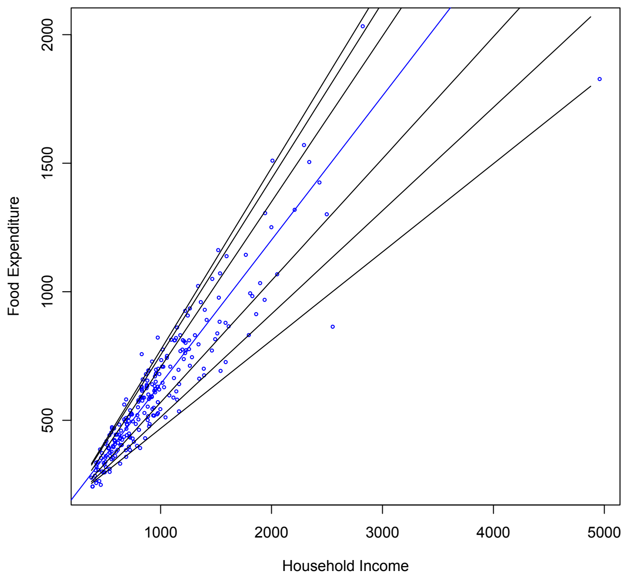
Estimated Conditional Quantiles of Engel Data



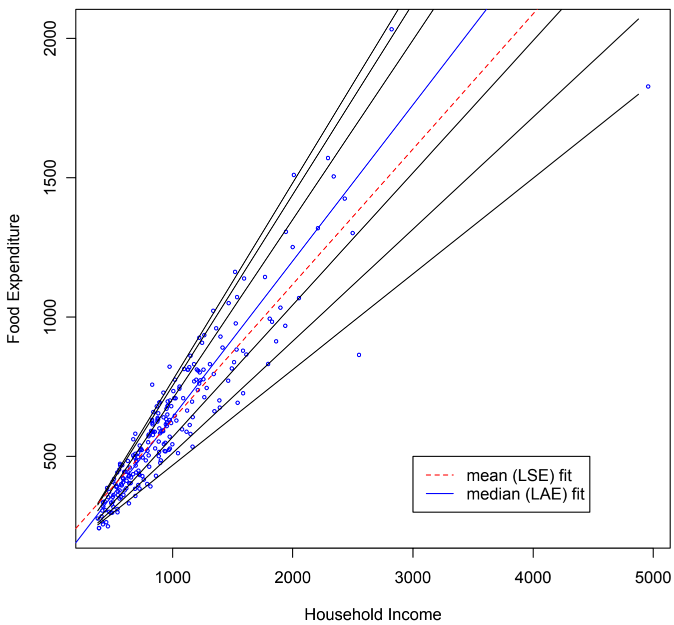
Estimated Conditional Quantiles of Engel Data



Estimated Conditional Quantiles of Engel Data



Estimated Conditional Quantiles of Engel Data



Motivation

- Sometimes, the researcher is interested in estimating a specific relationship between quantiles:
  - Kelly and Jiang (2014): tail risk over time.
  - Donald and Paarsch (1993): test predictions of game theory models for actions.
  - Koenker and Geling (2001): QR survival analysis models.
- Or, they need the effects over the entire conditional distribution:
  - Constant treatment effects.
  - Effects of changes over the distribution (Katz and Murphy (1992), Juhn, Murphy and Pierce (1993)).

Motivation: Traditional Approach

- Estimation: Solve the sample analog of the minimization problem for each τ (except, e.g., Koenker (1984), Koenker (2004)).
  - Inefficient
- Difficult to impose, and test, important restrictions across quantiles.
- Hard to implement when we look at the extreme quantiles (non-gaussian asymptotic distributions).

Our paper

- This paper develops (continuous) generalized method of moments estimation and inference procedures for QR models when allowing for general parametric restrictions on the parameters of interest over a set of quantiles.

Our paper

- This paper develops (continuous) generalized method of moments estimation and inference procedures for QR models when allowing for general parametric restrictions on the parameters of interest over a set of quantiles.
- We develop GMM-QR for fixed and large number of quantile partitions.
- We estimate imposing restrictions across quantiles.
- Characterize GMM-QR estimators in relation to the MLE (Efficiency Bound).

Literature

- Estimation QR with multiple quantiles: Koenker (1984).
- There is also large literature using moment conditions (and GMM) to estimate QR models, among many others, Xu, Sit, Wang and Huang (2017), Chen and Lee (2017), Kaplan and Sun (2017), Chen, Wan and Zhou (2015), Chen and Liao (2015), Chen and Pouzo (2012, 2009), Chernozhukov and Hong (2003), and Buchinsky (1998).
- Continuum GMM: Carrasco et al (2007), Chacko and Viceira (2005), Carrasco and Florens (2002), Jiang and Knight (2002), Singleton (2001) and Carrasco and Florens (2000).

Our Paper in the Literature

- Compared to the existing procedures for estimation and inference in QR models, our approach has several distinctive advantages:
  - First, an important application of the proposed methods is to allow researchers to estimate models under restrictions on the coefficient functions.
  - Second, the estimator is flexible and does not necessarily require modeling the entire conditional distribution of the variable of interest.
  - Third, we derive the optimal CGMM-QR, which is efficient.
  - Fourth, our algorithm is computationally simple and easy to implement in practice because we explore the partitioning argument.
  - Finally, a direct implication of imposing restrictions on coefficients is to test hypotheses on the shape of the quantile curve.

Outline

1. Introduction
2. **Framework**
3. Partition Result
4. GMM-QR Estimator
5. A Continuum of Moments
6. Monte Carlo
7. Conclusion

Framework

Consider the structural model:

$$Y = X^\top \beta(U),$$

with U being:

- The heterogeneity in responses
- Scalar and unspecified function.

We consider models such that:

- $X^\top \beta(U)$ is strictly increasing in U.
- $U \sim \text{Unif}[0, 1]$.
- Exogeneity: U is independent from X

Framework

- Under these assumptions,

$$Q_\tau(Y|X) = X^\top \beta(Q_\tau(U|X)),$$

  where $\tau \in (0,1)$ is a specific quantile.
- Under the exogeneity assumption, we obtain the following representation:

$$Q_\tau(Y|X) = X^\top \beta(\tau),$$

  with $X^\top \beta(\tau)$ is strictly increasing in τ.
  - For a fixed τ, $\beta(\tau)$ represents the marginal effect of X on $Q_\tau(Y|X)$.
  - The relationship between τ's is left unspecified.

Framework

- In this paper, we add to the linear QR model a flexible assumption about how τ affects the conditional quantile function.
$$\beta(\tau) = g(\theta, \tau), \tag{1}$$
where $g(\cdot)$ is a K-vector of known functions and θ is a vector of unknown parameters.
- Under the restriction (1), the structural model relates the dependent variable Y as a function of observables X and the unobservable random variable U as,
$$Y = X^{\top} g(\theta, U),$$
and the QR model is then parametrically specified in both X and τ as,
$$Q_{\tau}(Y|X) = X^{\top}\beta(\tau) = X^{\top} g(\theta, \tau).$$

Example

- Consider the simple case of a linear location-scale-shift QR model with one regressor. The model can the written as

$$Y = b_0 + b_1 X + \sigma(X)e, \text{ where } e = F^{-1}(U)$$

if $\sigma(X) = (\gamma_0 + \gamma_1 X)$. Thus, the conditional quantile function can be written as following

$$\begin{aligned} Q_\tau(Y|X) &= b_0 + b_1 X + \sigma(X)F_\tau^{-1}(U) \\ &= (b_0 + \gamma_0 F_\tau^{-1}(U)) + (b_1 + \gamma_1 F_\tau^{-1}(U))X \\ &= \beta_0(\tau) + \beta_1(\tau)X, \end{aligned}$$

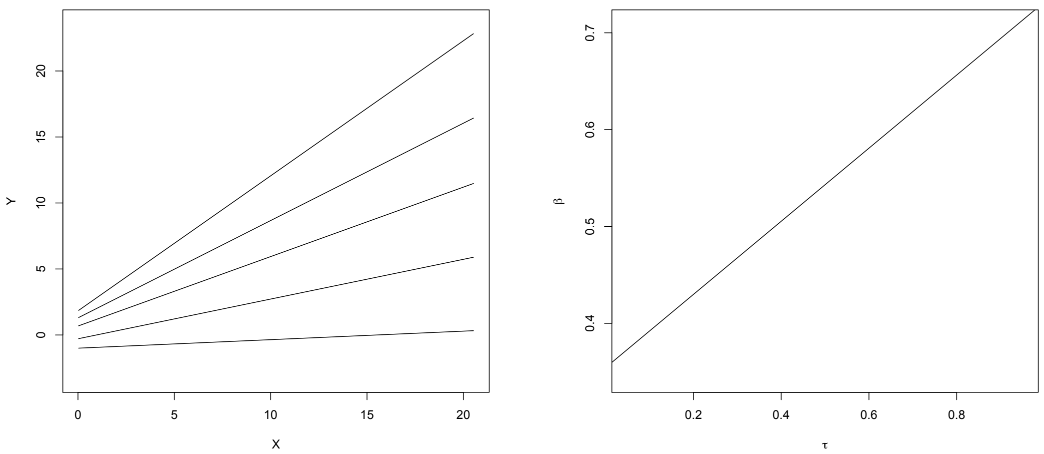
where

$$\begin{aligned} \beta_0(\tau) &= (b_0 + \gamma_0 F_\tau^{-1}(U)) \\ \beta_1(\tau) &= (b_1 + \gamma_1 F_\tau^{-1}(U)). \end{aligned}$$

Example

- Let $F(U)$ be the Logistic distribution with location and scale parameters zero and one, respectively.
- Then, $F_{\tau}^{-1}(U) = \ln \tau/(1-\tau)$, and we have both the constant and slope coefficients varying with the quantile τ.
- The function $g(\theta, \tau)$ is a simple linear function as

$$\beta_0(\tau) = g_0(\theta, \tau) = \theta_1 + \theta_2 \ln \frac{\tau}{1-\tau}, \quad \beta_1(\tau) = g_1(\theta, \tau) = \theta_3 + \theta_4 \ln \frac{\tau}{1-\tau}.$$



Framework : Assumptions

A1. : Define $W_i \equiv (Y_i, X_i)$, $\{W_i\}_{i=1}^n$ i.i.d.

A2. : $f_{Y_i}(y|X_i)$ exists almost surely for every $i = 1, \ldots, n$. It is bounded above and continuous in y, uniformly over the support of X_i.

A3. : $\beta(\tau) = g(\theta_0, \tau)$ for $\theta_0 \in \mathbb{R}^K$ for all $\tau \in \mathcal{T} \subset (0, 1)$.

- The function g is continuously differentiable in both arguments and known up to the vector θ.

A4. : For fixed M, $q_M(X, \tau) = q_M(X)$, and the matrix $\Sigma_M := \mathrm{E}[q_M(X) q_M(X)^\top]$ is positive definite.

Framework

- Under assumption **A3**, the quantile regression model is:

$$Q_{\tau}\left(Y|X\right) = F_{\tau}^{-1}\left(Y|X\right) = X^{\top}\beta\left(\tau\right) = X^{\top}g\left(\theta,\tau\right)$$

- The QR is parametrically specified in X and τ

- The conditional moment restriction for the QR problem is:

$$\mathrm{E}\left[\tau - 1\left\{Y - X^{\top}\beta_0(\tau) \leq 0\right\}\middle| X\right] = 0,$$

for all $\tau \in \mathcal{T}$.

Framework

- For all $\tau \in (0,1)$, the unconditional GMM-QR problem considers the moments restrictions:

$$m\left(W;\theta,\tau\right) = q_M(X,\tau)\left(\tau - 1\left\{Y - X^\top \beta\left(\tau\right) \leq 0\right\}\right)$$

  with $q_M(X,\tau)$ is a M_τ-vector of subset of the conditioning (or instrumental) variables of X, and $W = [Y,X]$.
- For a fixed quantile τ, $M_\tau \geq K$.
- Notice that, in the standard linear QR case, $q_M(X,\tau) = X$ for all τ.

Framework

- We can then define the following unconditional moment condition for the QR problem,

$$\mathrm{E}\left[m\left(W;\beta_0,\tau\right)\right]=0. \tag{2}$$

- When we impose the structure $g(\cdot)$ across quantiles, we obtain the set of moments conditions in (2) with the following $m(\cdot)$ function where, for a given τ, $M_\tau \geq d_\theta$

$$m\left(W;g\left(\theta,\tau\right),\tau\right):=q_M(X,\tau)\psi(W,\theta,\tau), \tag{3}$$

where $\psi(W,\theta,\tau):=\left(\tau-1\left\{Y-X^\top g\left(\theta,\tau\right)\leq 0\right\}\right)$.
- We work with the set of unconditional moment conditions in (2) for this problem with $m(\cdot)$ given by (3).

Framework

- To make estimation practical in applications, we use a partition of the space of quantiles.
- From equation (3), let

$$\begin{aligned} \mathrm{m}(W,\theta,\tau_1,...,\tau_{L-1}) &:= \left[m\left(W;g\left(\theta,\tau_1\right),\tau_1\right)^\top,...,m\left(W;g\left(\theta,\tau_{L-1}\right),\tau_{L-1}\right)^\top\right]^\top, \\ &= \left[q_M(X,\tau_1)^\top\psi(W,\theta,\tau_1),...,q_M(X,\tau_{L-1})^\top\psi(W,\theta,\tau_{L-1})\right]^\top, \end{aligned}$$

be the $M\cdot(L-1)$-vector of moments restrictions for $L-1$ values of $\{\tau_1,\ldots,\tau_{L-1}\}\subset(0,1)$.

- The index L defines the number of partitions of the quantile space, hence $L-1$ is the number of quantiles.
- For instance, when $L=4$ we have the three quartiles and four partitions.

Framework

- We note that the weights $q_M(X, \tau)$ may be a function of the quantile τ in the general statement of the problem.
- We follow the QR literature, and with abuse of notation, set $q_M(X, \tau) = q_M(X)$ for all τ.
- Nevertheless, we will see below that for the optimal GMM, the weights are a function of the quantiles.
- When the function $q_M(\cdot)$ does not depend on τ, the vector $\mathrm{m}(\cdot)$ will be written as

$$\mathrm{m}(W, \theta, \tau_1, ..., \tau_{L-1}) = \left[\psi(W, \theta, \tau_1), ..., \psi(W, \theta, \tau_{L-1})\right]^\top \otimes q_M(X).$$

Framework

- We find the vector of parameters θ that solves the following minimization problem considering a set of values of τ, $\tau_1, ..., \tau_L$,

$$\theta_0 = \underset{\theta}{\operatorname{argmin}} \, \mathsf{E}\left[\mathsf{m}(W, \theta, \tau_1, ..., \tau_{L-1})\right]^\top \Omega(\theta, \tau_1, ..., \tau_{L-1})^{-1} \mathsf{E}\left[\mathsf{m}(W, \theta, \tau_1, ..., \tau_{L-1}) \tag{4}$$

where $\Omega(\theta, \tau_1, ..., \tau_{L-1})$ is the weight matrix that is equal to the variance-covariance matrix of the moments conditions.

- Except for Koenker (1984, 2004), the studies in the QR literature estimate separate regression models for each value of τ. These simple approaches correspond to solving a GMM set-up for a single given value of τ separately.

Framework

- Note that in this system of unconditional moments restrictions, the dimension of unknown parameters is equal to the dimension of θ, d_θ.
- In this case, the dimension of the vector of instruments M needs to be at least as large as d_θ.
- In our case, θ is the same for all values of τ, and it is more efficient to estimate a model that considers the moments conditions for different values of τ all together.
- When we consider the moments conditions for all τ together, even in the case that $M = d_\theta$, we have a overidentified model since the number of equations is larger than the number of parameters.

Outline

Partition

- Split the interval $(0, 1)$ into L equally spaced intervals.
- $L - 1$ quantiles, $\beta(j/L) = g(\theta, j/L)$, $j = 1, ..., L - 1$

$$
\begin{aligned}
\mathrm{H}_L(W;\theta) &:= \left[\frac{1}{L} - 1\left\{Y - X^\top g\left(\theta, \frac{1}{L}\right) \le 0\right\}, ..., \frac{L-1}{L} - 1\left\{Y - X^\top g\left(\theta, \frac{L-1}{L}\right) \le 0\right\}\right]^\top \\
&= \left[\psi\left(W, \theta, \frac{1}{L}\right), ..., \psi\left(W, \theta, \frac{L-1}{L}\right)\right]^\top .
\end{aligned}
$$

Partition

- Based on this partition of the quantile space, we can construct a $M \cdot (L-1) \times 1$ vector of moments denoted by $\mathrm{m}_L(W; \mathrm{g}_L(\theta))$ and defined as follows:

$$\mathrm{m}_L(W; \mathrm{g}_L(\theta)) := \left[m\left(W; g\left(\theta, \frac{1}{L}\right), \frac{1}{L}\right)^\top, m\left(W; g\left(\theta, \frac{2}{L}\right), \frac{2}{L}\right)^\top, ..., m\left(W; g\left(\theta, \frac{L-1}{L}\right) \tag{5}$$

- We also define the following:

$$\mathrm{g}_L(\theta) := \left[g(\theta, 1/L)^\top, g(\theta, 2/L)^\top, ..., g(\theta, (L-1)/L)^\top \right]^\top$$

$$\Omega_L(\theta) := \mathrm{E}\left[\mathrm{m}_L(W; \mathrm{g}_L(\theta)) \mathrm{m}_L(W; \mathrm{g}_L(\theta))^\top \right],$$

where $\mathrm{g}_L(\theta)$ is a $K \cdot (L-1)$ vector, and $\Omega_L(\theta)$ is a $M \cdot (L-1) \times M \cdot (L-1)$ covariance matrix.

Partition

- Our parameter of interest is the solution of the following problem

$$\theta_0 = \underset{\theta}{\operatorname{argmin}}\, Q_L(\theta), \tag{6}$$

$$Q_L(\theta) := \mathrm{E}[\mathrm{m}_L(W; \mathrm{g}_L(\theta))]^\top \Omega_L^{-1}(\theta_0)\, \mathrm{E}[\mathrm{m}_L(W; \mathrm{g}_L(\theta))].$$

Partition

- Next we present a result where we compute the population function $Q_L(\theta)$ as a function of the number of partitions of the quantile space.
- Define

$$\Sigma_L^{-1} := L \cdot \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}. \tag{7}$$

Partition: Lemma 1

Lemma

*Under conditions **A1**-**A4**, the objective function of the partition GMM problem $Q_L(\theta)$ in (6) can be expressed as:*

$$Q_L(\theta) = \left(\mathrm{E}\left[\mathrm{H}_L(W;\theta) \otimes q_M(X)\right]\right)^\top \cdot \left(\Sigma_L^{-1} \otimes \Sigma_M^{-1}\right) \cdot \left(\mathrm{E}\left[\mathrm{H}_L(W;\theta) \otimes q_M(X)\right]\right)$$

Outline

1. Introduction
2. Framework
3. Partition Result
4. **GMM-QR Estimator**
5. A Continuum of Moments
6. Monte Carlo
7. Conclusion

GMM-QR Estimator

- Now we suggest a GMM-QR estimator based on a set of the moments that are analog of the ones in equation (6).
- The GMM-QR estimator is defined as,

$$\widehat{\theta}^{M,L}_{GMM} = \operatorname*{argmin}_{\theta} \left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{m}_L(W_i; \mathrm{g}_L(\theta)) \right)^{\top} (\Sigma_L^{-1} \otimes \widehat{\Sigma}_M^{-1}) \left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{m}_L(W_i; \mathrm{g}_L(\theta)) \right), \tag{8}$$

with

$$\widehat{\Sigma}_M = \frac{1}{n} \sum_{i=1}^{n} q_M(X_i) q_M(X_i)^{\top},$$

$\mathrm{m}_L(\cdot)$ given in (5), and Σ_L^{-1} given in (7).

GMM-QR Estimator

- Imposing additional assumptions in our model, we can establish consistency and asymptotic normality of GMM-QR estimator.
- Consider the following conditions

A5. $\theta_0 \in \text{int}(\Theta)$ where Θ is a compact subset of $\mathbb{R}^{d_\theta}$.

A6. Let $h(X, \theta, U) = X^\top g(\theta, U)$. For all $\theta, \theta' \in \Theta$, $\left|h^{-1}(X, \theta, Y) - h^{-1}(X, \theta', Y)\right| \leq \kappa(X, Y)\|\theta - \theta'\|$, where $\mathrm{E}[\kappa(X, Y)^2] < \infty$.

A7. There is a unique θ_0 such that $Q_\tau(Y|X) = X^\top g(\theta_0, \tau)$.

GMM-QR Estimator

- Before we present the results on the limiting properties of the estimator, we define the several population quantities.
- Let

$$G_{lL}(X,\theta_0) := f_{Y|X}\left(X^\top g\left(\theta_0, \frac{l}{L}\right)\right) X^\top g_\theta\left(\theta_0, \frac{l}{L}\right)$$
$$G_L(X,\theta_0) := (G_{1L}(X,\theta_0)^\top, \ldots, G_{L-1,L}(X,\theta_0)^\top),$$

where $G_{lL}(X,\theta_0)$ is a $1 \times d_\theta$ vector.

GMM-QR Estimator

- Now we state the result for asymptotic normality of the GMM-QR estimator.

Theorem

*Let Assumptions **A1–A7** hold and let* $M \cdot (L-1) \geq d_\theta$. *Then, as* $n \to \infty$

$$\widehat{\theta}^{M,L}_{GMM} \xrightarrow{p} \theta_0,$$

and

$$\sqrt{n}(\widehat{\theta}^{M,L}_{GMM} - \theta_0) \xrightarrow{d} N\left(0, V_L^{-1}\right).$$

where
$V_L := (\mathrm{E}[G_L(X,\theta_0) \otimes q_M(X)])^\top (\Sigma_L \otimes \Sigma_M)^{-1} \mathrm{E}[G_L(X,\theta_0) \otimes q_M(X)]$.

Optimal GMM Estimator

- We can define the vector of conditional moment restrictions at the true value θ_0 as,

$$\mathrm{E}\left[\mathrm{H}_L(W;\theta_0)|X\right] = 0.$$

- This vector of conditional moments conditions implies a vector of unconditional moments conditions:

$$\mathrm{E}\left[q_M^*(X,L)\mathrm{H}_L(W,\theta_0)\right] = 0,$$

  where $q_M^*(X,L)$ is a $d_\theta \times (L-1)$ matrix of functions of X that minimizes the asymptotic variance of the GMM estimator.
- We write $q_M^*(X,L)$ instead of $q_M^*(X,\tau)$ to emphasize the use of the partition of the quantile space.

Optimal GMM Estimator

- Using the optimal GMM theory, $q_M^*(X, L)$ is given by:

$$q_M^*(X, L) = \frac{\partial}{\partial \theta^\top} \mathrm{E}\left[\mathrm{H}_L(W; \theta)|X\right]\big|_{\theta=\theta_0} \cdot \mathrm{E}\left[\mathrm{H}_L(X, Y, \theta_0) \cdot \mathrm{H}_L(X, Y, \theta_0)^\top \,\middle|\, X\right]^{-1},$$
$$= G_L(X, \theta_0) \cdot \Sigma_L^{-1},$$

since $\mathrm{E}\left[\mathrm{H}_L(X, Y, \theta_0) \cdot \mathrm{H}_L(X, Y, \theta_0)^\top \,\middle|\, X\right] = \Sigma_L$, and $\frac{\partial}{\partial \theta^\top} \mathrm{E}\left[\mathrm{H}_L(W; \theta)|X\right]\big|_{\theta=\theta_0} = G_L(X, \theta_0)$.

- From the optimal GMM theory (see, e.g., Hall (2005)), an optimal estimator reaches the efficiency bound for the optimal GMM, which is given by the inverse of the following variance-covariance matrix

$$V_L^* = \mathrm{E}[G_L(X, \theta_0)^\top \Sigma_L^{-1} G_L(X, \theta_0)]. \tag{9}$$

Optimal GMM Estimator

- We can now formally compare the variance of the GMM-QR, V_L, given in Theorem 1 with the variance of the optimal GMM, V_L^*, given in equation (9).
- The next result shows that, for a fixed number of partitions L, the variance of the optimal GMM is smaller than that of the GMM-QR.

Lemma

*Under Assumptions **A1–A7**, we have that for a fixed L,*

$$V_L^* \geq V_L.$$

Outline

1. Introduction
2. Framework
3. Partition Result
4. GMM-QR Estimator
5. **A Continuum of Moments**
6. Monte Carlo
7. Conclusion

GMM Objective Function with Large Number of Moments

- We start by investigating the properties of the GMM objective function $Q_L(\theta)$ in (6) when L diverges to infinity and the partitioning set becomes dense.
- The dimension of the vector of instruments $(q_M(X, \cdot))$ is kept constant. In other words, L increases, and M is fixed.

GMM Objective Function with Large Number of Moments

Lemma

*Assume **A1**-**A7** and that*

$$\lim_{\tau\downarrow 0} \mathrm{E}\left[m(W;g(\theta,\tau),\tau)^{\top}\right]\Sigma_M^{-1}\,\mathrm{E}\left[m(W;g(\theta,\tau),\tau)\right]$$
$$=\lim_{\tau\uparrow 1} \mathrm{E}\left[m(W;g(\theta,\tau),\tau)^{\top}\right]\Sigma_M^{-1}\,\mathrm{E}\left[m(W;g(\theta,\tau),\tau)\right]=0.$$

Then, for a fixed θ,

$$\lim_{L\to\infty} Q_L(\theta)=Q(\theta)\,,$$

where

$$Q(\theta)=\int_0^1\left(\mathrm{E}\left[\left(\frac{f_{Y|X}\left(X^{\top}g(\tau;\theta);\theta_0\right)}{f_{Y|X}\left(X^{\top}g(\tau;\theta);\theta\right)}\right)q_M(X)^{\top}\right]\Sigma_M^{-1}\,\mathrm{E}\left[q_M(X)\left(\frac{f_{Y|X}\left(X^{\top}g(\tau;\theta);\theta_0\right)}{f_{Y|X}\left(X^{\top}g(\tau;\theta);\theta\right)}\right)\right]\right)d\tau$$
$$-\mathrm{E}\left[q_M(X)^{\top}\right]\Sigma_M^{-1}\,\mathrm{E}\left[q_M(X)\right],$$

which is uniquely minimized at θ_0 and $Q(\theta_0)=0$.

Efficiency Bound

- For comparison purpose, we define the MLE estimator for this problem.
- By definition, the MLE maximizes a recentered version of the average log-likelihood function,

$$\begin{aligned} \widehat{\theta}_{MLE} &= \arg\max_{\theta} \frac{1}{n} \sum_{i=1}^{n} \ln \left(f_{Y|X}\left(Y_i, X_i; \theta\right) \right) \\ &= \arg\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \ln \left(\frac{f_{Y|X}\left(Y_i, X_i; \theta_0\right)}{f_{Y|X}\left(Y_i, X_i; \theta\right)} \right). \end{aligned} \tag{10}$$

Efficiency Bound

- In order to derive the properties of the MLE, we use again the representation $Y = X^\top g(\theta, U)$ where $U \sim Unif[0,1]$ independently of X.
- We start by deriving the log-likelihood function.

$$\begin{aligned} P(Y \leq y | X = x; \theta) &= P(x^\top g(\theta, U) \leq y | X = x; \theta) \\ &= P(h(x, \theta, U) \leq y | X = x; \theta). \end{aligned}$$

Assuming that h has a well-defined inverse function,

$$P(Y \leq y | X = x; \theta) = P(U \leq h^{-1}(x, \theta, y)),$$

and the conditional density of Y given X at θ is

$$\begin{aligned} f_{Y|X}(y|x;\theta) &= \frac{\partial}{\partial y} h^{-1}(x, \theta, y) \\ &= \frac{1}{x^\top g_u(\theta, h^{-1}(x, \theta, y))}, \end{aligned} \tag{11}$$

Efficiency Bound

- Using this derivation, the log-likelihood at a given θ is

$$l_{y|X}\left(y|\, x;\theta\right)=\sum_{i=1}^{n}-\ln x^{\top}g_{u}(\theta,h^{-1}\left(x,\theta,y\right)).$$

- We can also derive the score function $S_{Y|X}(y|x;\theta)\in\mathbb{R}^{d_{\theta}}$.
- Using the fact that $h\left(X,\theta,U\right)=X^{\top}g(\theta,U)$ in Assumption **A6**, we obtain the following expression for $S(X,U,\theta_0)$,

$$S(X,U,\theta_0)=\left(\frac{X^{\top}g_{uu}(\theta_0,U)}{X^{\top}g_{u}(\theta_0,U)}\frac{X^{\top}g_{\theta}(\theta_0,U)}{X^{\top}g_{u}(\theta_0,U)}-\frac{X^{\top}g_{u\theta}(\theta_0,U)}{X^{\top}g_{u}(\theta_0,U)}\right)^{\top}.$$

Efficiency Bound

- Based on the score evaluated at the true parameter, we are able to find the efficiency bound for this problem based on the following information matrix,

$$\mathcal{I}(\theta_0) = \mathrm{E}[S(X, U, \theta_0)S(X, U, \theta_0)^\top], \tag{12}$$

  where $\mathcal{I}(\theta_0)$ is the information matrix.
- The asymptotic variance of the MLE is the inverse of the information matrix, and the inverse of the information matrix represents the efficiency bound for this quantile problem.

Efficiency Bound

- We consider the limiting objective function derived in Lemma 4, and construct an estimator based on $Q(\theta)$.

Lemma

*Under Assumptions **A1–A7**, and*

$$\lim_{\tau\downarrow 0} \mathrm{E}\left[q_M(X)^\top f_{Y|X}\left(X^\top g(\theta_0,\tau)\right) g_\theta^\top(\theta_0,\tau) X^\top\right] \Sigma_M^{-1} \mathrm{E}\left[X f_{Y|X}\left(X^\top g(\theta_0,\tau)\right) g_\theta(\theta_0,\tau) q_M(X)\right]$$
$$= \lim_{\tau\downarrow 0} \mathrm{E}\left[q_M(X)^\top f_{Y|X}\left(X^\top g(\theta_0,1-\tau)\right) g_\theta^\top(\theta_0,1-\tau) X^\top\right] \Sigma_M^{-1} \mathrm{E}\left[X f_{Y|X}\left(X^\top g(\theta_0,1-\tau)\right) g_\theta(\theta_0,1-\tau) q_M(X)\right]$$
$$= 0.$$

As $L \to \infty$, we have that,

$$\lim_{L\to\infty} V_L \le \mathcal{I}(\theta_0).$$

Efficiency Bound

- In the next result we show that, when L diverges to infinity, the variance of the optimal GMM estimator reaches the efficiency bound.

Lemma

*Under assumptions **A1**-**A7**, and*

$$\lim_{\tau\downarrow 0} \mathrm{E}\left[\left[f_{Y|X}\left(X^{\top}g\left(\theta_0,\tau\right)\right)g_{\theta}^{\top}\left(\theta_0,\tau\right)X^{\top}\right]\cdot\left[Xf_{Y|X}\left(X^{\top}g\left(\theta_0,\tau\right)\right)g_{\theta}\left(\theta_0,\tau\right)\right]\right] =$$

$$\lim_{\tau\downarrow 0} \mathrm{E}\left[\left[f_{Y|X}\left(X^{\top}g\left(\theta_0,1-\tau\right)\right)g_{\theta}^{\top}\left(\theta_0,1-\tau\right)X^{\top}\right]\cdot\left[Xf_{Y|X}\left(X^{\top}g\left(\theta_0,1-\tau\right)\right)g_{\theta}\left(\theta_0,1-\tau\right)\right]\right]$$

$$= 0.$$

As $L \to \infty$, we have that

$$\lim_{L\to\infty} V_L^{*} = \mathcal{I}(\theta_0),$$

where $\mathcal{I}(\theta_0)$ is the Fisher-information matrix.

Smooth GMM Estimator

- We generalize the estimators developed in the paper for fixed L, M and suggest a GMM estimator that considers large L and M.
- Recall that $p_{lL}(u)$ and $q_{mM}(x)$ are scalar basis functions for u and x, respectively. In this estimator, we let $p_{lL}(u)$ be mean-zero basis functions of u.
- Remember that $U = h^{-1}(X, \theta_0, y)$ is distributed Unif[0, 1], and therefore any basis function $\tilde{p}_{lL}(u)$ can be demeaned by taking $p_{lL}(u) = \tilde{p}_{lL}(u) - \int_0^1 \tilde{p}_{lL}(v)dv$.
- Note that now we are using demeaned basis functions.
- As before, $p_L(u) = [p_{1L}(x), ..., p_{L-1,L}(x)]$ and $q_M(x) = [q_{1M}(x), ..., q_{MM}(x)]$.

Smooth GMM Estimator

- The moment conditions here are not related to individual quantile restrictions.
- All quantile restrictions are equivalent to all the moment restrictions with (for example) polynomials basis functions.
- Specifically,

$$\mathrm{E}[q_l(X)U] = 0 \quad \text{for all } \tau \text{ with } U = h^{-1}(X, \theta, Y)$$

is equivalent to

$$\mathrm{E}\left[q_l(X)\left(r(U) - \mathrm{E}\left(r(U)\right)\right)\right] = 0$$

e.g $r(U) = U^k$ with $k = 1, 2...K$

Smooth GMM Estimator

- Using these smooth basis functions for u and x, we define the new GMM estimator,

$$\widehat{\theta}^{M,L}_{SGMM} = \operatorname*{argmin}_{\theta \in \Theta} \left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{m}_L(W_i; \mathrm{g}_L(\theta)) \right)^{\top} \left(\widehat{\Sigma}_L^{-1} \otimes \widehat{\Sigma}_M^{-1} \right) \left(\frac{1}{n} \sum_{i=1}^{n} \mathrm{m}_L(W_i; \mathrm{g}_L(\theta)) \right), \tag{13}$$

where $\mathrm{m}_L(W;\theta) = p_L(h^{-1}(Y,\theta,X)) \otimes q_M(X)$, $\widehat{\Sigma}_M^{-1}$ is as defined in model (8), and $\widehat{\Sigma}_L = \frac{1}{n} \sum_{i=1}^{n} p_L(h^{-1}(X_i, \widehat{\theta}, Y_i)) p_L(h^{-1}(X_i, \widehat{\theta}, Y_i))^{\top}$, with $\widehat{\theta}$ in the last equation being a first stage root-n consistent estimator for θ.

- This implies that estimating equations correspond to moment equations instead of covariance equations as in Poirier (2017).

Smooth GMM Estimator

- Here are some additional assumptions to allow to use of Theorem 3.4 in Poirier (2017):

  B1. (Basis function) $p_{lL}(\cdot)$ and $q_{mM}(\cdot)$ are bounded and continuously differentiable everywhere;

  B2. (P-Donsker Classes) $\{q_{mM}(X_i) : m \in \mathbb{N}\}$, $\{p_{lL}(h^{-1}(X, \theta, Y)) : l \in \mathbb{N}, \theta \in \Theta\}$ and $\{p'_{lL}(h^{-1}(X, \theta, Y)) \frac{\partial}{\partial \theta} h^{-1}(X, \theta, Y) : l \in \mathbb{N}, \theta \in \Theta\}$ are uniformly bounded P-Donsker classes;

  B3. (Approximation) For any bounded continuous function $f(z, u, \theta)$ there exists $\beta^{ML}(\theta)$, a $ML \times 1$ vector such that $\sup_{x \in \mathcal{X}, u \in \mathcal{U}, \theta \in \Theta} |f(x, u, \theta) - p_L(u) \otimes q_M(x) \beta^{ML}(\theta)| \to 0$ as L and M goes to infinity;

  B4. (Eigenvalues) The minimum eigenvalues of $\mathrm{E}[p_L(h^{-1}(X, \theta, Y)) p_L(h^{-1}(X, \theta, Y))^\top]$ and $\mathrm{E}[q_M(X) q_M(X)^\top]$ are bounded above and bounded below by the function $C/\zeta(L)$ and $C/\zeta(M)$ uniformly in θ, where $\zeta(L)$ is a known function with $\zeta(L) \to \infty$ as $L \to \infty$, and $C > 0$ is a constant;

  B5. (Differentiability) $h^{-1}(X, \theta, Y)$ is twice continuously differentiable in θ a.s. - X.

Smooth GMM Estimator

- Examples of basis functions that satisfy these assumptions include polynomials, or cubic splines.
- Under some conditions, for cubic splines the corresponding function is $\zeta(m) = m$, while for power series it is $\zeta(m) = m^2$.
- We require these functions to be smooth approximating basis functions and therefore indicators are ruled out by these assumptions.
- Under these assumptions and an additional condition regarding the rate of growth of the number of basis functions, we can show that the GMM estimator in (13) is asymptotically efficient attaining the efficient bound in equation (12).

Smooth GMM Estimator

Theorem (Consistency and Asymptotic Normality)

*Let Assumptions **A1**-**A7** and **B1**-**B5** hold, and let $\widetilde{\theta}$ be a preliminary first-step estimator of θ_0 that satisfies $\|\widetilde{\theta} - \theta_0\| = O_p(\tau_n)$. Then, if $M^2\zeta(M)^2 L^2\zeta(L)^2 \left(\tau_n + \frac{1}{\sqrt{n}}\right) \to 0$ and $M, L \to \infty$ as $n \to \infty$, then*

$$\widehat{\theta}^{M,L}_{SGMM} \xrightarrow{p} \theta_0,$$

and

$$\sqrt{n}(\widehat{\theta}^{M,L}_{SGMM} - \theta_0) \xrightarrow{d} N\left(0, V^*\right),$$

where $V^* = \mathcal{I}(\theta_0)^{-1}$ is the efficient bound for this problem.

- The first-step, consistent estimator can be for example $\widehat{\theta}^{M,L}_{GMM}$ which converges to θ_0 at the rate $\tau_n = n^{-1/2}$.

Outline

Monte Carlo: Design

- We draw random samples from the DGP:

$$x \sim Unif[0,1]$$
$$y|x \sim N(\beta_0 + \beta_1 x, (\gamma_0 + \gamma_1 x)^2)$$
$$\beta_0 = 0.5,\ \beta_1 = 1,\ \gamma_0 = 0.2,\ \gamma_1 = 0.8$$

We estimate the linear specification

$$X^\top g(\tau;\theta),\ X = [1, x]$$
$$g(\tau;\theta) = [\theta_1 + \theta_2 \quad (\theta_3 + \theta_4)\Phi^{-1}(\tau)]^\top.$$

Monte Carlo: Design

- Sample size $N \in \{500, 1000, 3000\}$.
- Number of partitions $L \in \{10, 50, 100\}$.
- Number of moments $M \in \{1, 2, ..., 5\}$.

Monte Carlo: Results

Table: Results (non-smooth) – N = 1000 L = 10

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Standard GMM					
$\theta_1 = 0.5$	-0.00596	-0.00606	0.02781	0.94914	0.02843
$\theta_2 = 0.2$	-0.00085	-0.00061	0.02291	0.95117	0.02291
$\theta_3 = 1$	0.02345	0.02597	0.07612	0.94608	0.07961
$\theta_4 = 0.8$	0.00752	0.00500	0.06604	0.94812	0.06644
Optimal GMM					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00931	-0.00920	0.02468	0.93149	0.02637
$\theta_2 = 0.2$	-0.00437	-0.00512	0.01880	0.94683	0.01930
$\theta_3 = 1$	0.03297	0.03134	0.07363	0.91922	0.08064
$\theta_4 = 0.8$	0.01335	0.01187	0.05760	0.94172	0.05910
MLE					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00047	-0.00093	0.02035	0.95100	0.02034
$\theta_2 = 0.2$	-0.00077	-0.00159	0.01415	0.95300	0.01417
$\theta_3 = 1$	0.00194	0.00225	0.05775	0.94200	0.05775
$\theta_4 = 0.8$	0.00047	0.00199	0.04038	0.94900	0.04037

Monte Carlo: Results

Table: Results (non-smooth) – N = 1000 L = 50

			Standard GMM		
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00702	-0.00614	0.03253	0.95166	0.03326
$\theta_2 = 0.2$	0.00178	0.00211	0.02866	0.94361	0.02871
$\theta_3 = 1$	0.03182	0.03208	0.08100	0.93051	0.08699
$\theta_4 = 0.8$	0.01916	0.01565	0.08581	0.93958	0.08788
			Optimal GMM		
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00989	-0.00922	0.02627	0.93519	0.02805
$\theta_2 = 0.2$	-0.00343	-0.00387	0.01900	0.94239	0.01930
$\theta_3 = 1$	0.03917	0.03776	0.08003	0.91975	0.08906
$\theta_4 = 0.8$	0.01773	0.01497	0.06460	0.93519	0.06696
			MLE		
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00047	-0.00093	0.02035	0.95100	0.02034
$\theta_2 = 0.2$	-0.00077	-0.00159	0.01415	0.95300	0.01417
$\theta_3 = 1$	0.00194	0.00225	0.05775	0.94200	0.05775
$\theta_4 = 0.8$	0.00047	0.00199	0.04038	0.94900	0.04037

Monte Carlo: Results

Table: Results (non-smooth) – N = 1000 L = 100

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Standard GMM					
$\theta_1 = 0.5$	-0.00495	-0.00601	0.03567	0.94828	0.03600
$\theta_2 = 0.2$	0.00510	0.00389	0.03127	0.94016	0.03167
$\theta_3 = 1$	0.03223	0.03367	0.08349	0.93813	0.08946
$\theta_4 = 0.8$	0.02230	0.01533	0.09109	0.93712	0.09374
Optimal GMM					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00917	-0.00817	0.02722	0.93977	0.02871
$\theta_2 = 0.2$	-0.00273	-0.00361	0.01967	0.94912	0.01985
$\theta_3 = 1$	0.03948	0.03604	0.08112	0.91900	0.09018
$\theta_4 = 0.8$	0.02101	0.01601	0.07027	0.94496	0.07331
MLE					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00047	-0.00093	0.02035	0.95100	0.02034
$\theta_2 = 0.2$	-0.00077	-0.00159	0.01415	0.95300	0.01417
$\theta_3 = 1$	0.00194	0.00225	0.05775	0.94200	0.05775
$\theta_4 = 0.8$	0.00047	0.00199	0.04038	0.94900	0.04037

Monte Carlo: Results

Table: Results (smooth) – N = 1000 L = 10, part 1

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Standard GMM					
$\theta_1 = 0.5$	-0.00670	-0.00670	0.02751	0.94159	0.02830
$\theta_2 = 0.2$	-0.00197	-0.00207	0.02196	0.95166	0.02204
$\theta_3 = 1$	0.02543	0.02674	0.07627	0.94260	0.08036
$\theta_4 = 0.8$	0.00946	0.00714	0.06314	0.94260	0.06381
Smooth GMM – M = 1 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00014	-0.00024	0.02327	0.94100	0.02326
$\theta_2 = 0.2$	0.00061	0.00007	0.01618	0.94700	0.01619
$\theta_3 = 1$	0.00170	0.00369	0.06404	0.94700	0.06403
$\theta_4 = 0.8$	-0.00056	0.00207	0.04556	0.95200	0.04554
Smooth GMM – M = 2 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00005	-0.00076	0.02193	0.95200	0.02192
$\theta_2 = 0.2$	0.00031	0.00007	0.01526	0.94800	0.01525
$\theta_3 = 1$	0.00288	0.00287	0.06070	0.94500	0.06074
$\theta_4 = 0.8$	0.00273	0.00629	0.04277	0.95500	0.04284

Monte Carlo: Results

Table: Results (smooth) – N = 1000 L = 10, part 2

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Smooth GMM – M = 4 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00038	-0.00030	0.02195	0.94900	0.02194
$\theta_2 = 0.2$	0.00146	0.00129	0.01554	0.94300	0.01560
$\theta_3 = 1$	0.00301	0.00176	0.06186	0.94900	0.06191
$\theta_4 = 0.8$	0.00501	0.00623	0.04354	0.94800	0.04381
Smooth GMM – M = 5 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00047	0.00017	0.02184	0.94700	0.02184
$\theta_2 = 0.2$	0.00217	0.00184	0.01571	0.94100	0.01585
$\theta_3 = 1$	0.00348	0.00403	0.06220	0.94500	0.06227
$\theta_4 = 0.8$	0.00582	0.00711	0.04406	0.94800	0.04442
MLE					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00047	-0.00093	0.02035	0.95100	0.02034
$\theta_2 = 0.2$	-0.00077	-0.00159	0.01415	0.95300	0.01417
$\theta_3 = 1$	0.00194	0.00225	0.05775	0.94200	0.05775
$\theta_4 = 0.8$	0.00047	0.00199	0.04038	0.94900	0.04037

Monte Carlo: Results

Table: Results (smooth) – N = 1000 L = 50, part 1

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Standard GMM					
$\theta_1 = 0.5$	-0.00726	-0.00712	0.03272	0.95459	0.03350
$\theta_2 = 0.2$	0.00122	0.00146	0.02791	0.94450	0.02793
$\theta_3 = 1$	0.03394	0.03591	0.08237	0.93037	0.08905
$\theta_4 = 0.8$	0.02033	0.01644	0.08286	0.93946	0.08527
Smooth GMM – M = 1 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00076	0.00087	0.02479	0.94400	0.02479
$\theta_2 = 0.2$	0.00098	0.00049	0.01733	0.94100	0.01735
$\theta_3 = 1$	0.00558	0.00577	0.06801	0.94000	0.06821
$\theta_4 = 0.8$	0.00278	0.00491	0.04865	0.95100	0.04871
Smooth GMM – M = 2 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00182	0.00152	0.02261	0.95100	0.02268
$\theta_2 = 0.2$	0.00164	0.00064	0.01625	0.93700	0.01633
$\theta_3 = 1$	0.00643	0.00639	0.06288	0.94900	0.06318
$\theta_4 = 0.8$	0.00594	0.00822	0.04557	0.94600	0.04593

Monte Carlo: Results

Table: Results (smooth) – N = 1000 L = 50, part 2

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Smooth GMM – M = 4 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00306	0.00273	0.02321	0.94500	0.02340
$\theta_2 = 0.2$	0.00340	0.00320	0.01674	0.93800	0.01707
$\theta_3 = 1$	0.00962	0.01120	0.06495	0.94300	0.06563
$\theta_4 = 0.8$	0.01073	0.01208	0.04696	0.94700	0.04815
Smooth GMM – M = 5 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00389	0.00292	0.02320	0.93900	0.02351
$\theta_2 = 0.2$	0.00451	0.00449	0.01724	0.93800	0.01782
$\theta_3 = 1$	0.01061	0.01295	0.06583	0.95200	0.06664
$\theta_4 = 0.8$	0.01241	0.01341	0.04826	0.94300	0.04981
MLE					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00047	-0.00093	0.02035	0.95100	0.02034
$\theta_2 = 0.2$	-0.00077	-0.00159	0.01415	0.95300	0.01417
$\theta_3 = 1$	0.00194	0.00225	0.05775	0.94200	0.05775
$\theta_4 = 0.8$	0.00047	0.00199	0.04038	0.94900	0.04037

Monte Carlo: Results

Table: Results (smooth) – N = 1000 L = 100, part 1

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Standard GMM					
$\theta_1 = 0.5$	-0.00516	-0.00595	0.03595	0.94679	0.03631
$\theta_2 = 0.2$	0.00467	0.00328	0.03084	0.93775	0.03118
$\theta_3 = 1$	0.03444	0.03392	0.08408	0.92871	0.09083
$\theta_4 = 0.8$	0.02255	0.01540	0.08984	0.93675	0.09259
Smooth GMM – M = 1 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00182	0.00159	0.02566	0.95000	0.02571
$\theta_2 = 0.2$	0.00212	0.00200	0.01780	0.94300	0.01792
$\theta_3 = 1$	0.00661	0.00636	0.07062	0.94600	0.07090
$\theta_4 = 0.8$	0.00358	0.00567	0.04984	0.94800	0.04994
Smooth GMM – M = 2 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00280	0.00250	0.02296	0.94300	0.02312
$\theta_2 = 0.2$	0.00307	0.00234	0.01647	0.94400	0.01675
$\theta_3 = 1$	0.00910	0.00783	0.06424	0.94800	0.06485
$\theta_4 = 0.8$	0.00743	0.00882	0.04712	0.94300	0.04767

Monte Carlo: Results

Table: Results (smooth) – N = 1000 L = 100, part 2

	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Standard GMM					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
Smooth GMM – M = 4 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00535	0.00505	0.02376	0.94500	0.02435
$\theta_2 = 0.2$	0.00582	0.00443	0.01771	0.93100	0.01864
$\theta_3 = 1$	0.01266	0.01415	0.06555	0.95000	0.06672
$\theta_4 = 0.8$	0.01329	0.01564	0.04946	0.93800	0.05119
Smooth GMM – M = 5 (Analytical Σ_L)					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	0.00624	0.00626	0.02382	0.94200	0.02461
$\theta_2 = 0.2$	0.00712	0.00604	0.01788	0.91800	0.01924
$\theta_3 = 1$	0.01590	0.01645	0.06729	0.94400	0.06911
$\theta_4 = 0.8$	0.01694	0.02001	0.05025	0.94000	0.05300
MLE					
	Mean Bias	Median Bias	Std. Dev	Coverage (95 % CI)	RMSE
$\theta_1 = 0.5$	-0.00047	-0.00093	0.02035	0.95100	0.02034
$\theta_2 = 0.2$	-0.00077	-0.00159	0.01415	0.95300	0.01417
$\theta_3 = 1$	0.00194	0.00225	0.05775	0.94200	0.05775
$\theta_4 = 0.8$	0.00047	0.00199	0.04038	0.94900	0.04037

Outline

Conclusion

- Develop new GMM-QR estimators for simultaneous multiple quantiles.
- Estimation allows imposing restrictions across quantiles.
- Easy-to-implement efficient estimation and inference.

Conclusion

- Develop new GMM-QR estimators for simultaneous multiple quantiles.
- Estimation allows imposing restrictions across quantiles.
- Easy-to-implement efficient estimation and inference.
- Next Step:
  - Empirical application.
  - Optimal choice of L.