Inference in parametric models with many L-moments

Luis A. F. Alvarez

Advisors: Pedro A. Morettin and Chang Chiann

IME-USP

November 11, 2020

What are L-moments?

- L-moments are linear combinations of order statistics.
- For a random variable Y with quantile function Q_Y, Hosking (1990) defines the r-th L-moment as:

$$\lambda_r := \int_0^1 Q_Y(u) P^*_{r-1}(u) du$$

  - $P^*_r(u) = \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} u^k$ are shifted Legendre polynomials.
- L-moments provide "robust" alternatives to standard moments.
  - For $r = 2$, $\lambda_2 = \mathbb{E}|Y_1 - Y_2|$, where Y_1 and Y_2 are independent copies of Y.
  - In contrast, $\mathbb{V}[Y] = \mathbb{E}[(Y_1 - Y_2)^2]$.

Estimation of parametric models with L-moments

- Estimation of parametric models by matching L-moments has been shown to outperform MLE in finite samples from several distributions (Hosking et al., 1985; Hosking and Wallis, 1987; Hosking, 1990; Broniatowski and Decurninge, 2016).
- Let $Y_1, Y_2 \dots Y_T$ be a sample from F, where $F = F_{\theta_0}$ for some $F_{\theta_0} \in \{F_\theta : \theta \in \Theta\}$ and $\Theta \subseteq \mathbb{R}^d$.
  - $l_r(\theta)$: r-th L-moment of F_θ.
  - $\hat{l}_r$: sample estimator of the r-th L-moment.
- Conventional approach is to estimate θ_0 by solving:

$$\begin{bmatrix} l_1(\theta) \\ \vdots \\ l_d(\theta) \end{bmatrix} = \begin{bmatrix} \hat{l}_1 \\ \vdots \\ \hat{l}_d \end{bmatrix}$$

Improving upon the conventional approach

- Why not increase the number of L-moments used in estimation?
  - By increasing the number of L-moments, and weighting these properly, may achieve efficiency gains over standard approach.
  - Could construct GMM-style estimator that uses the first L L-moments in estimation and weighting matrix W^L:

$$\hat{\theta} \in \operatorname{argmin}_{\theta \in \Theta} \left[l_1(\theta) - \hat{l}_1 \quad \cdots \quad l_L(\theta) - \hat{l}_L \right] \times W^L \times \begin{bmatrix} l_1(\theta) - \hat{l}_1(\theta) \\ \vdots \\ l_L(\theta) - \hat{l}_L(\theta) \end{bmatrix}$$

  - From GMM literature, we know that increasing L "too much" with fixed T may lead to biases (Newey and Smith, 2004) $\implies$ L must be chosen "properly".
  - By appropriately choosing weights and varying L with T, may be possible that there is no asymptotic efficiency loss over MLE as $T \to \infty$.
    - L-moments characterise distributions with finite first moments (Hosking, 1990).

This thesis

- In this thesis, we propose to study the GMM-style L-moment estimator in a setting where $T, L \to \infty$.
  - Sufficient rates for inferential procedures to work.
  - Optimal weighting scheme of L-moments.
  - Study asymptotic (in)efficiency of L-moment estimator.
- We also propose to derive automatic rules for selecting the number of L-moments L.
  - Minimise higher-order expansions of MSE of L-moment estimator.
  - Similar to existing approaches in GMM literature (Donald and Newey, 2001).
- We also suggest an extension of the methodology to conditional models.
  - Application: estimation of conditional Value-at-Risk models.

Outline

Introduction

Asymptotic theory of the “many” L-moments estimator

Monte Carlo Exercise

Next steps

Setup

- Let $Y_1, Y_2 \dots Y_T$ be a sample from F, where $F = F_{\theta_0}$ for some $F_{\theta_0} \in \{F_\theta : \theta \in \Theta\}$ and $\Theta \subseteq \mathbb{R}^d$.
- We consider the estimator:

$$\hat{\theta} \in \arg\inf_{\theta \in \Theta} \left[\int_{\underline{p}}^{\bar{p}} \left(\hat{Q}_Y(u) - Q_Y(u|\theta) \right) \mathbf{P}^L(u)' du \right] W^L \left[\int_{\underline{p}}^{\bar{p}} \left(\hat{Q}_Y(u) - Q_Y(u|\theta) \right) \mathbf{P}^L(u) du \right]$$

  - $\hat{Q}_Y$ is the empirical quantile function.
  - $Q_Y(\cdot|\theta)$ is the quantile function of F_θ.
  - $\mathbf{P}^L(u) := [P_1(u), P_2(u) \dots P_L(u)]'$ is a vector of L quantile weighting functions.
  - W^L is a (possibly estimated) $L \times L$ weighting matrix.
  - $0 \le \underline{p} < \bar{p} \le 1$ are fixed trimming constants.
- L-moment-based estimator with "plug-in" sample estimator of L-moments can be obtained by setting $\mathbf{P}^L(u) = [P_0^*(u), \dots P_{L-1}^*(u)]'$, $\underline{p} = 0$ and $\bar{p} = 1$.

Consistency: assumptions

Assumption 1. $\sup_{u \in (\underline{p}, \bar{p})} |\hat{Q}_Y(u) - Q_Y(u)| \xrightarrow{P} 0.$

Assumption 2. The functions $\{P_l : l \in \mathbb{N}\}$ constitute an orthonormal sequence on $L^2[0,1]$.

Assumption 3. There exists a sequence of nonstochastic symmetric positive semidefinite matrices Ω^L such that, as $T, L \to \infty$, $\|W^L - \Omega^L\|_2 = o_P(1)$; $\|\Omega^L\|_2 = O(1)$.

Assumption 4. For each $\epsilon > 0$:

$$\liminf_{L \to \infty} \inf_{\theta \in \Theta : \|\theta - \theta_0\|_2 \geq \epsilon} \left[\int_{\underline{p}}^{\bar{p}} \left(Q_Y(u|\theta) - Q_Y(u|\theta_0) \right) \mathbf{P}^L(u)' du \right] \Omega^L \left[\int_{\underline{p}}^{\bar{p}} \left(Q_Y(u|\theta) - Q_Y(u|\theta_0) \right) \mathbf{P}^L(u) du \right] > 0.$$

Moreover, we require that $\sup_{\theta \in \Theta} \| Q_Y(\cdot|\theta) \mathbb{1}_{[\underline{p}, \bar{p}]} \|_{L^2[0,1]} < \infty$.

Consistency: discussion of assumptions

- Assumption 1 requires uniform consistency of the quantile process. This is known to be satisfied in a variety of settings, ranging from iid to weakly dependent data.
    - In our proofs, it would be sufficient to consider convergence in L^2 (Kaji, 2019).
- Assumption 2 is satisfied by shifted Legendre polynomials and orthonormal bases.
- Assumption 3 restricts the range of admissible weights. Trivially satisfied by $W^L = \mathbb{I}_L$.
- Part 1 of Assumption 4 is an identifiability condition.
    - If Θ is compact, $\theta \mapsto \|Q(\cdot|\theta)\|_{L^2[0,1]}$ is continuous, $\underline{p} = 0$, $\bar{p} = 1$, the $\{P_l\}_l$ constitute an orthonormal basis in $L^2[0,1]$ and $W^L = \mathbb{I}_L$, then part 1 is equivalent to identifiability of the parametric family $\{F_\theta\}_\theta$.
- Part 2 of Assumption 4 can be obtained by assuming compactness of Θ and continuity of $(u, \theta) \mapsto Q_Y(u|\theta)$.

Consistency: statement of result

Proposition 1

Suppose Assumptions 1-4 hold. Then, as $T, L \to \infty$, $\hat{\theta} \xrightarrow{P} \theta_0$.

Asymptotic linear representation: outline of proof

- In deriving an asymptotic linear representation of the estimator, we follow the usual argument in the standard proof for M-estimators (Newey and McFadden, 1994), but with additional steps to account for a growing number of moments ($L \to \infty$).

- Define $h^L(\theta) := \int_{\underline{p}}^{\bar{p}} \left(\hat{Q}_Y(u) - Q_Y(u|\theta) \right) \mathbf{P}^L(u) du$.
  - If h is differentiable on $\text{int}(\Theta)$ and $\theta_0 \in \text{int}(\Theta)$, then, using that $\hat{\theta} \xrightarrow{P} \theta_0$, we get, with probability approaching 1:
  $$\nabla_{\theta'} h^L(\hat{\theta})' W^L h^L(\hat{\theta}) = 0.$$

- Idea is then to perform a line-by-line mean-value expansion of $h^L(\hat{\theta})$ around $h^L(\theta_0)$, and "solve for" $\sqrt{T}(\hat{\theta} - \theta_0)$.
  - Since L grows, we require additional assumptions to bound the estimation error in $\nabla_{\theta'} h^L(\hat{\theta})$.
  - We will also need to bound the eigenvalues of $\nabla_{\theta'} h^L(\theta_0)' \Omega^L \nabla_{\theta'} h^L(\theta_0)$ uniformly from below to invert $\nabla_{\theta'} h^L(\hat{\theta})' W^L \nabla_{\theta'} h^L(\hat{\theta})$ with high probability.

Asymptotic linear representation: assumptions

Assumption 5. $\theta_0 \in \text{int}(\Theta_0)$. $Q_Y(u|\theta)$ is continuously differentiable on $\text{int}(\Theta)$, uniformly in $u \in [\underline{p}, \bar{p}]$. Moreover, for each $\theta \in \Theta$, $\nabla_{\theta'} Q_Y(\cdot|\theta)$ is square integrable on $[\underline{p}, \bar{p}]$.

Assumption 6. $\sqrt{T}(\hat{Q}_Y(\cdot) - Q_Y(\cdot))$ converges weakly in $L^\infty(\underline{p}, \overline{p})$.

Assumption 7. $Q_Y(u|\theta)$ is **twice** continuously differentiable on $\text{int}(\Theta)$, uniformly in $u \in [\underline{p}, \bar{p}]$. For each $\theta \in \Theta$, the Hessian $\nabla_{\theta\theta'} Q_Y(u|\cdot)$ is bounded in a neighbourhood of θ, uniformly in $u \in [\underline{p}, \bar{p}]$.

Assumption 8. The smallest eigenvalue of $\nabla_{\theta'} h^L(\theta_0)' \Omega^L \nabla_{\theta'} h^L(\theta_0)$ is bounded away from 0, uniformly in L.

Assumption 9. $T, L \to \infty$ with $\frac{L}{T} \to 0$.

Asymptotic linear representation: discussion of assumptions

- Assumptions 5 and 7 are needed for the mean-value expansions used in the proof.
- Assumption 6 is weak convergence (in $L^\infty(\underline{p}, \bar{p})$) of the empirical quantile process.
  - It would be sufficient to assume $\|\sqrt{T}(Q_Y(\cdot) - \hat{Q}_Y(\cdot))\mathbb{1}_{[\underline{p},\bar{p}]}\|^2_{L^2[0,1]} = O_P(1)$, which is implied by weak convergence in L^2.
- Assumption 8 is required to invert $\nabla_{\theta'} h^L(\hat{\theta})' W^L \nabla_{\theta'} h^L(\hat{\theta})$ with high probability.
  - It is related to the rank condition used in the proof of asymptotic normality of M-estimators, which is known to be equivalent to a local identification condition under rank-regularity assumptions (Rothenberg, 1971).
- Assumption 9 is a rate requirement on the growth of L.
  - It is used to control the estimation error in $\nabla_{\theta'} h^L(\hat{\theta})$.
  - The rate condition is sufficient, but by no means necessary.
  - It may be possible to weaken this assumption, possibly at the cost of more stringent assumptions.

Asymptotic linear representation: statement of result

Proposition 2

Suppose Assumptions 1-9 hold. Then the estimator admits the asymptotic linear representation:

$$\sqrt{T}(\hat{\theta} - \theta_0) = -(\nabla_{\theta'} h^L(\theta_0)' \Omega^L \nabla_{\theta'} h^L(\theta_0))^{-1} \nabla'_\theta h^L(\theta_0)' \Omega^L (\sqrt{T} h^L(\theta_0)) + o_P(1)$$

Asymptotic approximation to the distribution

- We would like to provide an approximation to the distribution of the leading term in the asymptotic linear representation.

- Problem: growing dimensionality ($L \to \infty$) makes it difficult to establish weak convergence of this term.

  - We cannot work directly with $\sqrt{T} h^L(\theta_0)$, as we would do in the fixed-L case.

  - Could try to apply a CLT of linear order statistics to $\nabla_{\theta_0} h^L(\theta_0)' \Omega^L(\sqrt{T} h^L(\theta_0))$.
    - Not trivial to analyse and (appears) restricted to the iid case.
    - Not easy to extend to conditional models (later on).

- Our approach: work with a strong approximation concept.

  - Idea is to define, in the same probability space, a sequence of rvs X_T, $T \in \mathbb{N}$, with known distribution (possibly up to an estimable term) that "approximates" $\sqrt{T} h^L(\theta_0)$.

  - This approach to inference has been successfully applied in other areas (e.g. Belloni et al. (2019) in sieve estimation of conditional quantile models).

Asymptotic approximation: Gaussian approximation

- One possibility is to work with strong approximations of the empirical quantile process $\sqrt{T}(\hat{Q}_Y(\cdot) - Q_Y(\cdot))$ to a Gaussian process.

- Under some conditions, it is possible to define a sequence of Brownian bridges B_T, $T \in \mathbb{N}$; with known (up to an estimable term) covariance kernel Γ, such that:

$$\sup_{\underline{p} \leq u \leq \bar{p}} |\sqrt{T}(\hat{Q}_Y(u) - Q_Y(u)) - f(Q_Y(u))^{-1} B_T(u)| = o_P(1)$$

  - f is the Lebesgue density of Y.
  - in the project, we reproduce results in the literature for the iid (Csorgo and Revesz, 1978) and strictly stationary strongly mixing case (Fotopoulos and Ahn, 1994; Yoshihara, 1995).

- Using this strong approximation, we are able to show that:

$$\sqrt{T}(\hat{\theta} - \theta_0) = -(\nabla_{\theta'} h^L(\theta_0)' \Omega^L \nabla_{\theta'} h^L(\theta_0))^{-1} \nabla_{\theta'} h^L(\theta_0)' \Omega^L \left[\int_{\underline{p}}^{\bar{p}} \frac{B_T(u)}{f_{\theta_0}(Q_Y(u))} \mathbf{P}^L(u) du \right] + o_P(1)$$

Asymptotic approximation: Bahadur-Kiefer representation

- Yet another possibility is to work with Bahadur-Kiefer representations of the empirical quantile process.

- Under some assumptions, it is possible to show that:

$$\sup_{\underline{p}\leq u\leq\bar{p}} \left|\sqrt{T}(\hat{Q}_Y(u) - Q_Y(u)) - f(Q_Y(u))^{-1}\sqrt{T}(\hat{F}_Y(Q_Y(u)) - F(Q_Y(u))\right| = o_P(1)$$

  - $\hat{F}_Y(u) = \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\{Y_t \leq u\}$ is the empirical cdf.

- If F is continuous and strictly increasing, then $\hat{F}_Y(Q_Y(u)) = \frac{1}{T}\sum_{t=1}^{T} \mathbb{1}\{U_t \leq u\}$, where $U_t := F(Y_t)$ is a Uniform[0,1] rv.

- Both points could be combined to form the basis for an inferential procedure, at least in the iid case.
  - Though there are Bahadur-Kiefer representations in the weakly dependent case, difficult to see how to use it for inference in this context.

Outline

Monte Carlo Exercise: setup

- We consider random samples $Y_1, Y_2, \ldots, Y_T \sim F_{\theta_0}$, $F_{\theta_0} \in \{F_\theta : \theta \in \Theta\}$ and $\Theta \subseteq \mathbb{R}^d$.

- Our goal is to estimate θ_0 and use it in predicting "extreme" quantiles $Q_Y(u)$ via the "plug-in" estimator $Q(u|\tilde{\theta})$.

- We consider three estimators for θ_0:

  (i) MLE

  (ii) The L-moment estimator with identity weights.

  (iii) A two-step estimator where we first estimate θ_0 by (ii), and then plug it into the optimal L-moment weighting matrix formula derived under the Gaussian strong approximation.

- Estimator (ii) with $L = d$ is the conventional approach (Hosking, 1990).

- In the project, we also consider using a linearly interpolated estimator for $Q_Y(\cdot)$; and the unbiased (for the iid case) sample L-moment estimator of Hosking (1990).

Monte Carlo Exercise: Generalised Extreme Value Distribution

- In this exercise, the family $\{F_\theta : \theta \in \theta_0\}$ corresponds to the Generalised Extreme Value Distribution, and $\theta_0 := (\text{loc}, \text{scale}, \text{shape})' = (0, 1, 0.2)'$.

Rescaled Root Mean Squared Error $\mathrm{RMSE}(\hat{Q}(u))/Q(u)$: $T = 70$.

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.1779	0.3299	0.6181	1.1479
Method of L-moments: identity weights				
L= 3	0.1727	0.2848	0.4791	0.7902
L= 4	0.1730	0.2894	0.5011	0.8658
L= 5	0.1737	0.2898	0.5057	0.8905
L= 6	0.1744	0.2902	0.5090	0.9098
⋮				
L= 10	0.1769	0.2884	0.4977	0.8781
L= 11	0.1775	0.2877	0.4920	0.8566
⋮				
L= 15	0.1798	0.2846	0.4665	0.7627
L= 16	0.1803	0.2838	0.4607	0.7415

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.1779	0.3299	0.6181	1.1479
Method of L-moments: optimal weights				
-				
L= 4	0.1729	0.2820	0.4623	0.7391
L= 5	0.1751	0.2805	0.4465	0.6966
L= 6	0.1760	0.2834	0.4543	0.7133
⋮				
L= 10	0.1752	0.2807	0.4470	0.6952
L= 11	0.1745	0.2788	0.4430	0.6881
⋮				
L= 15	0.1736	0.2761	0.4377	0.6750
L= 16	0.1739	0.2749	0.4333	0.6638

Monte Carlo Exercise: Generalised Extreme Value Distribution (cont.)

- In this exercise, the family $\{F_\theta : \theta \in \theta_0\}$ corresponds to the Generalised Extreme Value Distribution, and $\theta_0 := (\text{loc}, \text{scale}, \text{shape})' = (0, 1, 0.2)'$.

Rescaled Root Mean Squared Error $\text{RMSE}(\hat{Q}(u))\,/\,Q(u)$: $T = 500$.

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.0625	0.1008	0.1561	0.2221
Method of L-moments: identity weights				
L= 3	0.0627	0.1076	0.1735	0.2536
L= 4	0.0624	0.1120	0.1871	0.2803
L= 5	0.0623	0.1140	0.1942	0.2954
L= 6	0.0622	0.1157	0.2001	0.3082
⋮				
L= 10	0.0624	0.1185	0.2117	0.3354
L= 11	0.0625	0.1188	0.2132	0.3393
⋮				
L= 15	0.0628	0.1195	0.2169	0.3494
L= 16	0.0628	0.1196	0.2174	0.3508

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.0625	0.1008	0.1561	0.2221
Method of L-moments: optimal weights				
-				
L= 4	0.0621	0.0999	0.1541	0.2174
L= 5	0.0624	0.0997	0.1524	0.2132
L= 6	0.0628	0.1007	0.1540	0.2158
⋮				
L= 10	0.0632	0.1013	0.1542	0.2154
L= 11	0.0634	0.1014	0.1542	0.2149
⋮				
L= 15	0.0633	0.1008	0.1531	0.2134
L= 16	0.0632	0.1005	0.1525	0.2124

Monte Carlo Exercise: Generalised Pareto Distribution

- In this exercise, the family $\{F_\theta : \theta \in \theta_0\}$ corresponds to the Generalised Pareto Distribution, and $\theta_0 := (\text{loc}, \text{scale}, \text{shape})' = (0, 1, 0.2)'$.

Rescaled Root Mean Squared Error $\text{RMSE}(\hat{Q}(u)) / Q(u)$: $T = 70$.

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.1957	0.3770	0.8737	2.3422
Method of L-moments: identity weights				
L= 3	0.1587	0.2918	0.5393	0.9827
L= 4	0.1606	0.2921	0.5444	1.0153
L= 5	0.1623	0.2920	0.5444	1.0264
L= 6	0.1638	0.2918	0.5430	1.0282
		⋮		
L= 10	0.1680	0.2902	0.5234	0.9535
L= 11	0.1689	0.2897	0.5165	0.9241
		⋮		
L= 15	0.1718	0.2877	0.4881	0.8062
L= 16	0.1724	0.2873	0.4814	0.7797

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.1957	0.3770	0.8737	2.3422
Method of L-moments: optimal weights				
		-		
L= 4	0.1572	0.2948	0.5374	0.9651
L= 5	0.1592	0.2975	0.5224	0.9082
L= 6	0.1598	0.2977	0.5199	0.8995
		⋮		
L= 10	0.1603	0.2951	0.4990	0.8276
L= 11	0.1594	0.2928	0.4915	0.8070
		⋮		
L= 15	0.1738	0.2918	0.4613	0.7024
L= 16	0.1533	0.2649	0.3927	0.5484

Monte Carlo Exercise: Generalised Pareto Distribution (cont.)

- In this exercise, the family $\{F_\theta : \theta \in \theta_0\}$ corresponds to the Generalised Pareto Distribution, and $\theta_0 := (\text{loc}, \text{scale}, \text{shape})' = (0, 1, 0.2)'$.

Rescaled Root Mean Squared Error $\text{RMSE}(\hat{Q}(u))/Q(u)$: $T = 500$.

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.0652	0.1156	0.1961	0.2974
Method of L-moments:identity weights				
L= 3	0.0573	0.1123	0.2000	0.3118
L= 4	0.0575	0.1146	0.2105	0.3359
L= 5	0.0578	0.1160	0.2176	0.3531
L= 6	0.0581	0.1168	0.2224	0.3653
		$\vdots$		
L= 10	0.0592	0.1183	0.2323	0.3923
L= 11	0.0594	0.1185	0.2336	0.3961
		$\vdots$		
L= 15	0.0601	0.1188	0.2362	0.4047
L= 16	0.0602	0.1187	0.2362	0.4051

	Q(0.9)	Q(0.99)	Q(0.999)	Q(0.9999)
MLE	0.0652	0.1156	0.1961	0.2974
Method of L-moments: optimal weights				
		-		
L= 4	0.0569	0.1076	0.1845	0.2774
L= 5	0.0569	0.1076	0.1826	0.2717
L= 6	0.0570	0.1087	0.1842	0.2733
		$\vdots$		
L= 10	0.0569	0.1091	0.1837	0.2700
L= 11	0.0569	0.1092	0.1838	0.2697
		$\vdots$		
L= 15	0.0569	0.1062	0.1751	0.2536
L= 16	0.0511	0.0901	0.1396	0.1938

Outline

Next steps: selecting the number of L-moments

- In light of the theory and Monte Carlo evidence, we propose to construct a semiautomatic method for selecting the number of L-moments.
  - Idea is to minimise higher-order expansions of the MSE of the object of interest.
  - We'll build on Donald and Newey (2001) and Newey and Smith (2004), who provide higher-order expansions of the MSE of GMM estimators.
    - More recent applications of the methodology (in the GMM context): Okui (2009), Cheng et al. (2019) and Abadie et al. (2019).

Next steps: optimality in the "many L-moment" environment

- Under a Gaussian approximation, the optimal choice of weighting matrix is:

$$\Omega^L = \mathbb{E}\left[\left(\int_{\underline{p}}^{\bar{p}} \frac{B_T(u)}{f_{\theta_0}(Q_y(u))}\mathbf{P}^L(u)du\right)\left(\int_{\underline{p}}^{\bar{p}} \frac{B_T(u)}{f_{\theta_0}(Q_y(U))}\mathbf{P}^L(u)du\right)'\right]^{-1}$$

- Under such choice, the variance of the leading term is:

$$(\nabla_{\theta'} h^L(\theta_0)' \Omega^L \nabla_{\theta'} h^L(\theta_0))^{-1}$$

- Does the method of L-moments achieve efficiency? Under which choise of quantile weighting function?
    - Perhaps work with Bahadur-Kiefer representation?

Next steps: extension to conditional models

- Gourieroux and Jasiak (2008) defined a conditional version of the r-th L-moment by replacing $Q_Y(\cdot)$ in the definition with $Q_{Y|X}(\cdot|x)$, the quantile function of the conditional distribution $F_{Y|X}(y|X = x)$, where X is a vector of covariates.

- Estimation is conducted by considering parametric restrictions on the first L conditional L-moments.

- We propose to revisit their approach in the framework of our project, where L varies with sample size.
    - Strong approximations of Belloni et al. (2019), which hold conditionally on X, will be useful in this context.

- Application: estimation of dynamic quantile models under flexible parametrisations, with an emphasis on conditional Value-at-Risk models.

Obrigado!

References I

Abadie, A., J. Gu, and S. Shen (2019). Instrumental variable estimation with first stage heterogeneity. Technical report, MIT Economics Department.

Belloni, A., V. Chernozhukov, D. Chetverikov, and I. Fernández-Val (2019). Conditional quantile processes based on series or many regressors. *Journal of Econometrics 213*(1), 4 – 29. Annals: In Honor of Roger Koenker.

Broniatowski, M. and A. Decurninge (2016). Estimation for models defined by conditions on their L-moments. *IEEE Transactions on Information Theory 62*(9), 5181–5198.

Cheng, X., Z. Liao, and R. Shi (2019). On uniform asymptotic risk of averaging gmm estimators. *Quantitative Economics 10*(3), 931–979.

Csorgo, M. and P. Revesz (1978, 07). Strong approximations of the quantile process. *Ann. Statist. 6*(4), 882–894.

Donald, S. G. and W. K. Newey (2001). Choosing the number of instruments. *Econometrica 69*(5), 1161–1191.

References II

Fotopoulos, S. and S. Ahn (1994). Strong approximation of the quantile processes and its applications under strong mixing properties. *Journal of Multivariate Analysis 51*(1), 17–45.

Gourieroux, C. and J. Jasiak (2008). Dynamic quantile models. *Journal of Econometrics 147*(1), 198–205.

Hosking, J. R. M. (1990, sep). L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *Journal of the Royal Statistical Society: Series B (Methodological) 52*(1), 105–124.

Hosking, J. R. M. and J. R. Wallis (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics 29*(3), 339–349.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics 27*(3), 251–261.

Kaji, T. (2019). Asymptotic theory of l-statistics and integrable empirical processes.

References III

Newey, W. K. and D. McFadden (1994). Chapter 36 large sample estimation and hypothesis testing. In *Handbook of Econometrics*, Volume 4, pp. 2111 – 2245. Elsevier.

Newey, W. K. and R. J. Smith (2004). Higher order properties of gmm and generalized empirical likelihood estimators. *Econometrica 72*(1), 219–255.

Okui, R. (2009). The optimal choice of moments in dynamic panel data models. *Journal of Econometrics 151*(1), 1–16.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica 39*(3), 577–91.

Yoshihara, K. (1995). The bahadur representation of sample quantiles for sequences of strongly mixing random variables. *Statistics & Probability Letters 24*(4), 299 – 304.