Quantile Factor Models

Liang Chen[1] Juan J. Dolado[2,3] Jesús Gonzalo[3]

[1]University of Oxford

[2]European University Institute

[3]Universidad Carlos III de Madrid

UQAM, Montreal
(14th September, 2016)

Outline

Approximate Factor Models (AFM):

$$X_{it} = \lambda_i' F_t + e_{it}, \quad i = 1, \ldots, N; t = 1, \ldots, T. \tag{1}$$

- λ_i: $r \times 1$ factor loadings; F_t: $r \times 1$ unobservable common factors.
- Relevant Literature: Chamberlain & Rothschild (1983), Bai & Ng (2002), Stock & Watson (2002), Bai (2003), Bai and Ng (2013), etc.
- $E[e_{it}|F_t] = 0 \Rightarrow E[X_{it}|F_t] = \lambda_i' F_t$. Hence (1) is a *conditional mean* AFM.
- Principal Component (PC) estimator of the factors and the loadings:

$$(\hat{F}, \hat{\lambda}) = \arg\min_{F,\lambda} \sum_{i=1,t=1}^{N,T} (X_{it} - \lambda_i' F_t)^2$$

$$\hat{\lambda}_i(F) = \Big(\sum_{t=1}^{T} F_t F_t'\Big)^{-1} \Big(\sum_{t=1}^{T} F_t X_{it}\Big).$$

Quantile Factor Models (QFM):

$$X_{it} = \lambda_i'(U_{it})F_t \quad i = 1, \ldots, N; t = 1, \ldots, T. \tag{2}$$

- $U_{it} \sim U[0,1]; U \perp F;$
- the mapping $\tau :\mapsto \lambda_i'(\tau)F_t$ is increasing for any F_t.
- for any $\tau \in \mathcal{T} \subset (0,1)$:

  $$P[X_{it} \leq \lambda_i'(\tau)F_t|F_t] = P[\lambda_i'(U_{it})F_t \leq \lambda_i'(\tau)F_t|F_t] = P[U_{it} \leq \tau] = \tau.$$

  so

  $$Q_{X_{it}}[\tau|F_t] = \lambda_i'(\tau)F_t.$$

- or:

  $$X_{it} = \lambda_i'(U_{it})F_t = \lambda_i'(\tau)F_t + (\lambda_i'(U_{it}) - \lambda_i'(\tau))F_t = \lambda_i'(\tau)F_t + v_{it} \tag{3}$$

  where

  $$Q_{v_{it}}[\tau|F_t] = 0.$$

Motivation of QFM

A simple example with two factors affecting different parts of the conditional distribution:

$$X_{it} = \lambda_1(U_{it})F_{1t} + \lambda_2(U_{it})F_{2t}$$

such that

$$\begin{cases} \lambda_1(\tau) = 0 & \text{if } 0 \leq \tau \leq 0.5 \\ \lambda_2(\tau) = 0 & \text{if } 0.5 < \tau \leq 1 \end{cases}$$

Then

$$\begin{cases} Q_\tau(X_{it}|F_{1t}, F_{2t}) = \lambda_2 F_{2t} & \text{if } \tau \leq 0.5 \\ Q_\tau(X_{it}|F_{1t}, F_{2t}) = \lambda_1 F_{1t} & \text{if } 0.5 < \tau. \end{cases}$$

Related Literature

- Dynamic Quantile Models (a section on QFM)
  - Gourieroux and Jasiak (WP 2006) (cross-section factor quantile model for panel data)
- Quantile Panel Data
  - Koenker (2004) (quantile regression for longitudinal data)
  - Abrevaya and Dahl (2008) (effects of quantile estimation on panel data)
  - Graham, Hahn and Powell (2009) (lack of incidental paremeter problem in a non-differentable panel dara model)
  - Lamarche (2010) (robust penalized quantile regression panel data)
  - Canay (2011) (quantile regression for panel data)
  - Rosen (2012) (set identification via quantile restrictions in panel data)
  - Kato et al. (2012)(asymptotics for panel quantile regression models)
  - Harding and Lamarche (2014) (quantile regression panel models with factors estimated by averages of covariates)
  - Kristensen (2014) (LAD versus PC factors in a FAVAR)
  - Arellano and Bonhomme (WP 2016) (nonlinear panel data estimation via quantile regressions)
- Factor Augmented Quantile Regression
  - Ando and Tsay (2011) (PC factors in quantile augmented regression models).

DGPs leading to QFM representation

Example 1: Location

$$X_{it} = \alpha_i f_t + \epsilon_{it}.$$

- $\{\epsilon_{it}\}$ i.i.d; $\epsilon_{it} = Q_\epsilon(U_{it})$, where $U_{it} \sim U[0,1]$.
- $$X_{it} = \alpha_i f_t + Q_\epsilon(U_{it}) = \lambda_i'(U_{it})F_t$$

  where

  $$\lambda_i(U_{it}) = [Q_\epsilon(U_{it}) \quad \alpha_i]' \quad F_t = [1 \quad f_t]'.$$

Example 2: Location-scale (same factor, sign restricted)

$$X_{it} = \alpha_i f_t + f_t \epsilon_{it}, \quad f_t > 0.$$

- $$X_{it} = \alpha_i f_t + Q_\epsilon(U_{it}) f_t = \lambda_i'(U_{it})F_t$$

  where

  $$\lambda_i(U_{it}) = Q_\epsilon(U_{it}) + \alpha_i \quad F_t = f_t.$$

Example 3: Location-scale (same factor, sign unrestricted)

$$X_{it} = \alpha_i f_t + f_t \epsilon_{it}.$$

- As before we have
$$X_{it} = \alpha_i f_t + Q_\epsilon(U_{it}) f_t = \lambda_i'(U_{it}) F_t,$$
where
$$\lambda_i(U_{it}) = Q_\epsilon(U_{it}) + \alpha_i \quad F_t = f_t.$$
Notice that $\tau \mapsto Q_\epsilon(\tau) f_t$ is NOT increasing if f_t take both positive and negative values.

The conditional quantile function can be written as

$$Q_{X_{it}}[\tau|f_t] = \begin{cases} [Q_\epsilon(\tau) + \alpha_i] f_t & \text{if } f_t \geq 0 \\ [Q_\epsilon(1-\tau) + \alpha_i] f_t & \text{if } f_t < 0 \end{cases}$$

or

$$\begin{aligned} Q_{X_{it}}[\tau|f_t] &= \alpha_i f_t + Q_\epsilon(\tau) \cdot f_t \cdot \mathbf{1}\{f_t \geq 0\} + Q_\epsilon(1-\tau) \cdot f_t \cdot \mathbf{1}\{f_t < 0\} \\ &= [\alpha_i, Q_\epsilon(\tau), Q_\epsilon(1-\tau)][f_t, f_t^+, f_t^-]'. \end{aligned}$$

Example 3: (Cont.)

$$X_{it} = \alpha_i f_t + f_t \epsilon_{it}.$$

X_{it} is observationally equivalent to

$$\begin{aligned} X_{it} &= \alpha_i f_t + Q_\epsilon(U_{it}) \cdot f_t^+ + Q_\epsilon(1 - U_{it}) \cdot f_t^- \\ &= \alpha_i f_t + \epsilon_{it} \cdot f_t^+ + Q_\epsilon(1 - F_\epsilon(\epsilon_{it})) \cdot f_t^- \\ &= \lambda_i(U_{it})' F_t \end{aligned}$$

where

$$\lambda_i(U_{it}) = [\alpha_i, Q_\epsilon(U_{it}), Q_\epsilon(1 - U_{it})]' \quad F_t = [f_t, f_t^+, f_t^-]'.$$

For example, when the distribution of ϵ_{it} is symmetric, i.e., $Q_\epsilon(\tau) = -Q_\epsilon(1-\tau)$,

$$X_{it} = \alpha_i f_t + f_t \epsilon_{it}$$

is equivalent to

$$X_{it} = \alpha_i f_t + |f_t| \epsilon_{it}.$$

Example 4: Location-scale (different factors)

$$X_{it} = \alpha_i f_t + g_t \epsilon_{it}, \quad g_t > 0.$$

- $$X_{it} = \alpha_i f_t + Q_\epsilon(U_{it}) g_t = \lambda_i'(U_{it}) F_t$$

  where

  $$\lambda_i(U_{it}) = [\alpha_i \quad Q_\epsilon(U_{it})] \quad F_t = [f_t \quad g_t].$$

We will show how standard factor analysis (PC) fails to recover these two factors.

Example 5:

$$X_{it} = \lambda_i \cdot e^{\epsilon_{it}} f_t, \quad \lambda_i, f_t > 0.$$

- $$X_{it} = \lambda_i \cdot e^{Q_\epsilon(U_{it})} f_t = \lambda_i'(U_{it}) F_t$$

  where

  $$\lambda_i(U_{it}) = \lambda_i \cdot e^{Q_\epsilon(U_{it})} \quad F_t = f_t.$$

- Also note that:

  $$\log X_{it} = \log \lambda_i + \log f_t + \epsilon_{it}.$$

Eq. (1) is more general than location-scale shift models.

Outline

The Estimators

Analogous to PC estimators for AFM:

$$(\hat{F}, \hat{\Lambda}(\tau)) = \arg\min_{F, \Lambda(\tau)} \sum_{i=1, t=1}^{N, T} \rho_\tau(X_{it} - \lambda_i'(\tau) F_t) \quad ???$$

where $\rho_\tau(u) = u(\tau - \mathbf{1}\{u \leq 0\})$ is the check function.

However, note that from Eq. (1) we have

$$X_{it} = \lambda_i'(U_{it}) F_t = \lambda_i' F_t + (\lambda_i(U_{it}) - \lambda_i)' F_t = \lambda_i' F_t + e_{it} \tag{4}$$

where $\lambda_i = E[\lambda_i(U_{it})]$ and $e_{it} = (\lambda_i(U_{it}) - \lambda_i)' F_t$. Notice (4) is a conditional mean AFM since $E[e_{it} | F_t] = 0$.

$E(X_t X_t') = \Lambda \Sigma_F \Lambda' + \Sigma_e$, where Σ_e is a diagonal matrix. This is similar to the covariance structure of an AFM.

Two-step estimators:

1. PC estimator for the factors: $\hat{F} = [\hat{F}_1, \ldots, \hat{F}_T]'$ the first r normalized eigenvectors of XX'.
2. QR of X_{it} on $\hat{F}$ for each i:

$$\hat{\lambda}_i(\tau) = \arg\min_{\lambda_i} \sum_{t=1}^{T} \rho_\tau(X_{it} - \lambda_i' \hat{F}_t).$$

Outline

Uniform Consistency

Assumption 1: Suppose that the observed data $\{X_{it}\}$ are generated by model (1) and

(i) The sequence $\{F_t\}$ is strictly stationary and m-dependent with $E\|F_t\|^4 < \infty$, and $\Sigma_F = E(F_t F_t') > 0$.

(ii) the random variables $\{U_{it}\}$ are uniformly distributed over $[0, 1]$ and independent across i and t, and U_{it} is independent of F_t for all i, t.

(iii) There is a compact set $\mathcal{A} \subset \mathbb{R}^r$ such that $\lambda_i(\tau) \in \mathcal{A}$ for all i and $\tau \in \mathcal{T}$, and there is a $\Sigma_\Lambda > 0$ such that $\|N^{-1} \sum_{i=1}^N \lambda_i \lambda_i' - \Sigma_\Lambda\| \to 0$ as $N \to \infty$.

(iv) The eigenvalues of $\Sigma_F \Sigma_\Lambda$ are distinct.

(v) The conditional density $f_X(x|F_t = f)$ exists, and is bounded and uniformly continuous in x for all $f \in \mathcal{F}$; $J(\lambda_i(\tau)) := E[f_X(\lambda_i(\tau)' F_t | F_t) F_t F_t']$ is positive definite for all τ.

Theorem 1 : Suppose Assumption 1 hold, then $\sup_{\tau \in \mathcal{T}} \|\hat{\lambda}_i(\tau) - H_{NT}^{-1} \lambda_i(\tau)\| = o_p(1)$ and $\sup_{\tau \in \mathcal{T}} \|\hat{\lambda}_i(\tau) - H_0^{-1} \lambda_i(\tau)\| = o_p(1)$ for all $i = 1, \ldots, N$.

$H_{NT} := (\Lambda' \Lambda / N)(F' \hat{F}_t / T) V_{NT}^{-1}$, where $\Lambda' := [\lambda_1, \ldots, \lambda_N]$, $F' := [F_1, \ldots, F_T]$, $\hat{F}' := [\hat{F}_1, \ldots, \hat{F}_T]$, and V_{NT} is a $r \times r$ diagonal matrix with the eigenvalues of $(NT)^{-1} XX'$ in decreasing order. Further, define $H_0 := \Sigma_\Lambda^{1/2} \Upsilon V^{-1/2}$, where V is a diagonal matrix with the eigenvalues of $\Sigma_\Lambda^{1/2} \Sigma_F \Sigma_\Lambda^{1/2}$ in decreasing order, and Υ is a matrix of corresponding eigenvectors.

Weak Convergence

Assumption 2 :For each $i \leq N$,

(i) $E||F_t||^8 < \infty$, and $T^{5/4}/N \to 0$ as $N, T \to \infty$;

(ii) the eigenvalues of $J_{H_0}(\lambda_i(\tau)) := H_0' J(\lambda_i(\tau)) H_0$ are bounded below by a constant $\rho^* > 0$ uniformly in τ.

Theorem 2 : Suppose Assumptions 1 and 2 hold, then for each i,

$$J_{H_0}(\lambda_i(\cdot)) \cdot \sqrt{T}[\hat{\lambda}_i(\cdot) - H_{NT}^{-1}\lambda_i(\cdot)] = -\mathbb{V}_{iT}(\cdot) + o_p(1) \text{ in } \ell^\infty(\mathcal{T}),$$

where $\mathbb{V}_{iT}(\cdot) = T^{-1/2}\sum_{t=1}^{T}\varphi_\tau(X_t - \lambda_i(\cdot)'F_t)H_0'F_t$ converges weakly to $\mathcal{B}_r(\cdot)$ in $\ell^\infty(\mathcal{T})$.

where $\varphi(u) = \mathbf{1}\{u \leq 0\} - \tau$, $\mathcal{B}_r$ denotes a vector of r independent Brownian Bridges.

To ignore the estimation effects of $\hat{F}$,

- $T^{1/2}/N \to 0$ in *linear* factor-augmented regressions. (Bai and Ng 2006)
- $T^{5/8}/N \to 0$ in *nonlinear* factor augmented regressions. (Bai and Ng 2008)

Estimation of Variance

Define (Powell 1986):

$$\hat{J}(\hat{\lambda}_i(\tau)) = \frac{1}{2h_T \cdot T} \sum_{t=1}^{T} \left\{ \mathbf{1}\{|X_{it} - \hat{\lambda}_i(\tau)'\hat{F}_t| \leq h_T\} \hat{F}_t \hat{F}_t' \right\},$$

Assumption 3: The bandwidth parameter h_T satisfies: $h_T \to 0$, $h_T \cdot T^{1/2} \to \infty$ and $\|H_{NT} - H_0\|/h_T = o_p(1)$. (e.g. $h_T = O(T^{-1/3})$)

Theorem 3 : Suppose Assumptions 1 to 3 hold, we have $\sup_{\tau \in \mathcal{T}} \left\| \hat{J}(\hat{\lambda}_i(\tau)) - J_{H_0}(\lambda_i(\tau)) \right\| = o_p(1)$, and thus for each $i \leq N$, $\hat{J}(\hat{\lambda}_i(\cdot)) \cdot \sqrt{T}[\hat{\lambda}_i(\cdot) - H_{NT}^{-1} \lambda_i(\cdot)] \Rightarrow \mathcal{B}_r(\cdot)$ in $\ell^\infty(\mathcal{T})$ as $N, T \to \infty$.

For each $i \leq N$ and each $\tau \in \mathcal{T}$,

$$[\tau(1-\tau)]^{-1/2} \cdot \hat{J}(\hat{\lambda}_i(\tau)) \cdot \sqrt{T}[\hat{\lambda}_i(\tau) - H_{NT}^{-1} \lambda_i(\tau)] \rightsquigarrow \mathcal{N}(0, I_r).$$

Imposing Identification Restrictions

Corollary 1: Suppose Assumptions 1 to 2 hold. Then

$$J_{H_0}(\lambda_i(\cdot)) \cdot \sqrt{T}[\hat{\lambda}_i(\cdot) - \lambda_i(\cdot)] = -\mathbb{V}_{iT}(\cdot) + o_p(1) \text{ in } \ell^\infty(\mathcal{T}). \tag{5}$$

if

$$T^{-1}\sum_{t=1}^{T} F_t F_t' = I_r \quad \text{and} \quad N^{-1}\sum_{i=1}^{N} \lambda_i \lambda_i' \text{ is diagonal.}$$

Other restrictions proposed by Bai and Ng (2013) can also be considered.

Simulations

DGP (Example 2):

- $X_{it} = \lambda_i f_t + f_t \epsilon_{it}$, where $\lambda_i, \epsilon_{it} \sim$ i.i.d $\mathcal{N}(0,1)$.
- $f_t = e^{\sigma Z_t}$, where Z_t are independent standard normal variables, and $\sigma = 0.7$ such that $E(X) \approx 1.28$ and $Var(X) \approx 1$.
- $\lambda_i(\tau) = \lambda_i + \Phi^{-1}(\tau)$, where $\Phi(\cdot)$ is the CDF of standard normal distribution.
- $h_T = T^{-1/3}$, $\tau = 0.25, 0.5, 0.75$, $T = 200, N = 1000$.

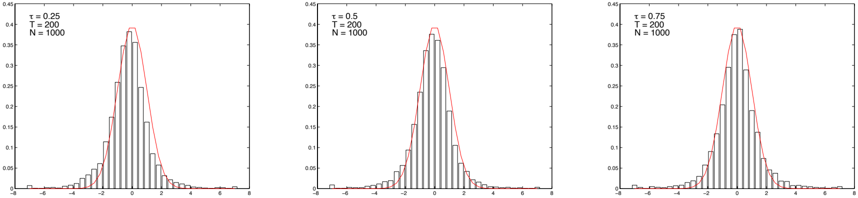


Figure: Histograms of $[\tau(1-\tau)]^{-1/2} \cdot \hat{J}(\hat{\lambda}_1(\tau)) \cdot \sqrt{T}[\hat{\lambda}_1(\tau) - H_{NT}^{-1}\lambda_1(\tau)]$ and the density function of $\mathcal{N}(0,1)$.

The N factor loadings across τ from ONE simulated sample:

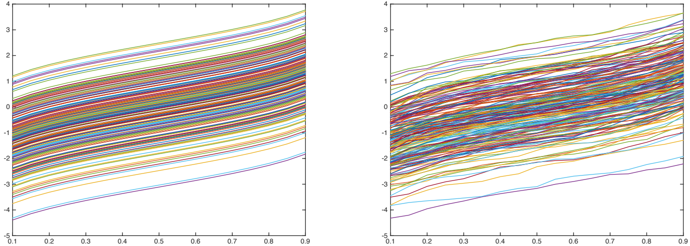


Figure: True factor loadings (left) and the estimated factor loadings (right) across τ.

Inference

Based on the identification restriction and Corollary 1, for each i and a $p \times r$ matrix R we can test

1. $R\lambda_i(\tau) = c$ for some known c at any given $\tau \in \mathcal{T}$;
2. $R\lambda_i(\tau) = c$ for some known c for all $\tau \in \mathcal{T}$;
3. $R\lambda_i(\tau) = R\lambda_i(0.5)$ for all $\tau \in \mathcal{T}$ — Durbin's problem because $\lambda_i(0.5)$ needs to be estimated.

For problems 1 and 2, under the null we have

$$v_T(\tau) = [RJ_{H_0}^{-2}(\lambda_i(\tau))R']^{-1/2} \cdot \sqrt{T}[R\hat{\lambda}_i(\tau) - c] \Rightarrow \mathcal{B}_p(\tau),$$

so

$$\sup_{\tau \in \mathcal{T}} \|v_T(\tau)\| \rightsquigarrow \sup_{\tau \in \mathcal{T}} \|\mathcal{B}_p(\tau)\|.$$

Inference – Durbin's problem

For problem 3, let $\hat{c} - R\lambda_i(0.5) = O_p(T^{-1/2})$, and

$$\begin{aligned} \hat{v}_T(\tau) &= [RJ_{H_0}^{-2}(\lambda_i(\tau))R']^{-1/2} \cdot \sqrt{T}[R\hat{\lambda}_i(\tau) - \hat{c}] \\ &= v_T(\tau) - G(\tau)' \cdot \sqrt{T}(\hat{c} - R\lambda_i(0.5)) \end{aligned}$$

where $G(\tau) = [RJ_{H_0}^{-2}(\lambda_i(\tau))R']^{-1/2}$.

Durbin's problem: the limiting distribution of $v_T(\tau)$ is parameter-free, but the limiting distribution of $\hat{v}_T(\tau)$ is contaminated by $\sqrt{T}(\hat{c} - R\lambda_i(0.5))$.

Inference – Khamaladze's transformation

When $p = 1$, define

$$g(\tau) = [\tau, G(\tau)']' \quad \dot{g}(\tau) = dg(\tau)/d\tau \quad C(\tau) = \int_{\tau}^{1} \dot{g}(s)\dot{g}(s)'ds$$

then under certain regularity conditions, the transformed process

$$\tilde{v}_T(\tau) = \hat{v}_T(\tau) - \int_0^{\tau} \left(C(\tau)^{-1} \int_{s=1}^{1} \dot{g}(t) d\hat{v}_T(t) \right) ds \Rightarrow \mathcal{W}(\tau),$$

a standard Brownian motion process. This is called the Khamaladze's transformation, which purges the estimation effects of $\sqrt{T}(\hat{c} - R\lambda_i(0.5))$ from $\hat{v}_T(\tau)$.

Outline

Extension: Problem – When the Two Step Procedure does not Work?

Consider **Example 4:**

$$X_{it} = \alpha_i f_t + g_t \epsilon_{it},$$

and assume that f_t, g_t and ϵ_{it} are mutually independent, and $E[\epsilon_{it}] = 0$.

- Note that the 1st step of our 2-step procedure can only capture f_t, because g_t only shift the scale but not the mean, i.e., g_t is not a mean factor.
- However, g_t may contain important information about the quantiles of the X, thus are useful for factor-augmented QR regressions (Ando and Tsay 2011).
- The question is: how to extract both f_t and g_t from X. And in general, how to extract all the relevant factors that affect the quantiles of X?

▸ Explanation

Extension: A Solution – Iterative Procedure (Single Step)

To find the r factors that affect the $\tau-$quantile of X, we seek to solve the following problem:

$$(\hat{F}, \hat{\Lambda}) = \underset{F \in \mathbb{R}^{T \times r}, \Lambda \in \mathbb{R}^{N \times r}}{\arg\min} \sum_{i=1}^{N} \sum_{t=1}^{T} \rho_\tau (X_{it} - \lambda_i' F_t).$$

The following simple iterative procedure is proposed:

1. Given $\hat{F}^{(k)} = [\hat{F}_1^{(k)}, \ldots, \hat{F}_T^{(k)}]$, using QR of X_{it} on $\hat{F}^{(k)}$ to estimate $\hat{\Lambda}_i^{(k)}$ for $i = 1, \ldots, N$.
2. Given $\hat{\Lambda}^{(k+1)} = [\hat{\Lambda}_1^{(k+1)}, \ldots, \hat{\Lambda}_N^{(k+1)}]$, using QR of X_{it} on $\hat{\Lambda}^{(k+1)}$ to estimate $\hat{F}_t^{(k+1)}$ for $t = 1, \ldots, T$.
3. Repeat the above steps until $\hat{F}^{(k)}$ and $\hat{F}^{(k+1)}$ are close enough.

Comparison of Both Procedures: Some Simulations

DGP (**Example 4**):

$$X_{it} = \lambda_i f_t + g_t \epsilon_{it},$$

$f_t \sim i.i.d\, \mathcal{N}(0,1)$, is the **mean-shift factor**. $g_t = e^{h_t}$ where $h_t \sim i.i.d\, \mathcal{N}(0,0.5)$, is the **scale-shift factor**. $\epsilon_{it} \sim i.i.d\, \mathcal{N}(0,1)$.

$\hat{F}_{PC}$ is the $T \times 2$ matrix of the 2 PC estimators. $\hat{F}_{QR}$ is the $T \times 2$ matrix of the 2 estimated factors using the iteration procedure. The table reports the R^2 (averaged from 100 replications) in the regressions of f_t on $\hat{F}_{PC}$, g_t on $\hat{F}_{PC}$, f_t on $\hat{F}_{QR}$ and g_t on $\hat{F}_{QR}$.

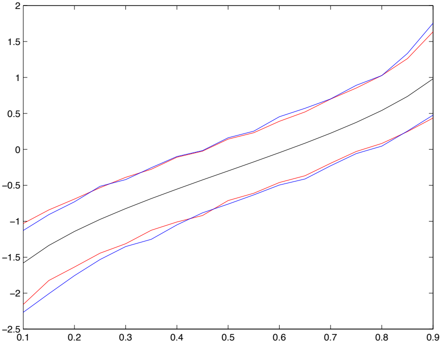
Table: 1-step procedure v.s PC

	$f,\hat{F}_{PC}$	$g,\hat{F}_{PC}$	$f,\hat{F}_{QR}$	$g,\hat{F}_{QR}$
$N,T = 20, \tau = 0.25$	.8940	.2705	.7823	.8307
$N,T = 50, \tau = 0.25$	.9596	.1662	.9255	.9256
$N,T = 100, \tau = 0.25$	.9802	.1058	.9618	.9635
$N,T = 20, \tau = 0.5$	.8940	.2705	.8120	.2361
$N,T = 50, \tau = 0.5$	.9596	.1662	.9600	.1541
$N,T = 100, \tau = 0.5$	.9802	.1058	.9650	.1110
$N,T = 20, \tau = 0.75$	.8940	.2705	.6946	.8149
$N,T = 50, \tau = 0.75$	.9596	.1662	.9345	.9251
$N,T = 100, \tau = 0.75$	.9802	.1058	.9570	.9635

Comparison of Both Procedures: Confidence Bands of the Loadings

DGP (Example 2):

- $X_{it} = \lambda_i f_t + f_t \epsilon_{it}$, where $\lambda_i, \epsilon_{it} \sim$ i.i.d $\mathcal{N}(0,1)$.
- $f_t = e^{\sigma Z_t}$, where Z_t are independent standard normal variables, and $\sigma = 0.7$ such that $E(X) \approx 1.28$ and $Var(X) \approx 1$.
- The 95% (pointwise) confidence interval of the $\lambda_1(\tau)$ (black) using the 2-step procedure (red) and the iterative procedure (blue).



Extension: Asymptotic Theory for the Iterative Procedure

Difficulties:

- Non smooth object function.
- $N + T$ incidental parameters.

Related studies:

- Ivan-Fernandez and Weidner (2015): $N + T$ incidental parameters but smooth object function.
- Kato et. al. (2012): Non smooth object function but only N incidental parameters.
- Kristensen (2011): Smooth the check function when $\tau = 1/2$ but no asymptotic theory.
- Kato and Galvao (2011): Smooth the check function for any τ but only N incidental parameters.

More Precisely the Question on Consistency is:

In AFM, let $\theta = [\lambda_1, \ldots, \lambda_N, F_1, \ldots, F_T]$, and

$$S_{NT}(\theta) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i' F_t)^2$$

be the object function. Then:

$$S_{NT}(\theta) - S_{NT}(\theta_0) = \underbrace{2 \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} e_{it} (\lambda_i' F_t - \lambda_{i0}' F_{t0})}_{I} + \underbrace{\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\lambda_i' F_t - \lambda_{i0}' F_{t0})^2}_{II}.$$

Note that: (i) $I = o_P(1)$ uniformly, (ii) $II \geq 0$ for any θ. Thus, from $S_{NT}(\hat{\theta}) - S_{NT}(\theta_0) \leq 0$ we have

$$\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\hat{\lambda}_i' \hat{F}_t - \lambda_{i0}' F_{t0})^2 = o_p(1),$$

and it follows easily that

$$\|P_{\hat{F}} - P_{F_0}\| \xrightarrow{P} 0,$$

where P_A denotes the projection matrix $A(A'A)^{-1}A'$.

(Cont.)

In QFM, the object function is:

$$H_{NT}(\theta) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\rho_\tau(X_{it} - \lambda_i' F_t).$$

Similarly,

$$H_{NT}(\theta) - H_{NT}(\theta_0) = \underbrace{2\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}[\tau - \mathbf{1}\{e_{it} \le 0\}](\lambda_i' F_t - \lambda_{i0}' F_{t0})}_{I}$$
$$+ \underbrace{\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}|X_{it} - \hat{\lambda}_i'\hat{F}_t| \cdot \mathbf{1}\{(X_{it} - \hat{\lambda}_i'\hat{F}_t)(X_{it} - \lambda_{i0}' F_{t0}) \le 0.\}}_{II}$$

where (i) $I = o_P(1)$ uniformly, (ii) $II \ge 0$ for any θ.

(Cont.)

Thus, from $H_{NT}(\hat{\theta}) - H_{NT}(\theta_0) \leq 0$ we have

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}|X_{it} - \hat{\lambda}_i'\hat{F}_t| \cdot \mathbf{1}\{(X_{it} - \hat{\lambda}_i'\hat{F}_t)(X_{it} - \lambda_{i0}'F_{t0}) \leq 0\} = o_P(1). \tag{6}$$

The question is how to show $\|P_{\hat{F}} - P_{F_0}\| \xrightarrow{P} 0$ from Eq. (6).

Cross-sectional Quantiles

Consider the following model:

$$X_{it} = \alpha_i f_t + g_t \epsilon_{it}.$$

If we treat α_i as i.i.d random variables, and assume that

$$\begin{pmatrix} \alpha_i \\ \epsilon_{it} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_\alpha \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \\ & \sigma_\epsilon^2 \end{pmatrix} \right), \tag{7}$$

and treat f_t and g_t as fixed parameters (i.e., every thing is conditional on them), then X_{it} are i.i.d across N, and

$$P[X_{it} \leq x] = P[\alpha_i f_t + g_t \epsilon_{it} \leq x] = \Phi\left(\frac{x - \mu_\alpha f_t}{\sqrt{\sigma_\alpha^2 f_t^2 + \sigma_\epsilon^2 g_t^2}} \right) \tag{8}$$

where Φ is the CDF of a standard normal distribution. Thus, we have the cross-sectional quantile of X_{it} at time t as:

$$Q_{X_t}(\tau) = \Phi^{-1}(\tau)\sqrt{\sigma_\alpha^2 f_t^2 + \sigma_\epsilon^2 g_t^2} + \mu_\alpha f_t. \tag{9}$$

- Problem: It is not possible to consistently estimate the space of $[f, g]$ and the factor loadings: $\lambda_i(\tau)$.

Outline

Stock Returns – 2-step procedure

- Monthly US common stock returns (CRSP), 1980-2014, ($N = 475, T = 420$), $r = 1+$ intercept.

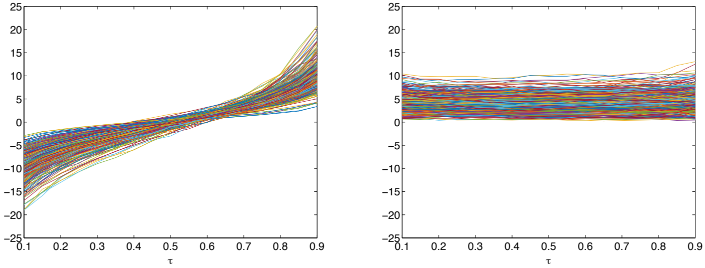


Figure: Factor loadings for the constant (left) and $\hat{F}_t$ (right).

Mutual Funds Returns – 2-step procedure

- Monthly US mutual funds returns (CRSP), 2000-2014, $(N = 2419, T = 180)$, $r = 1+$ intercept.

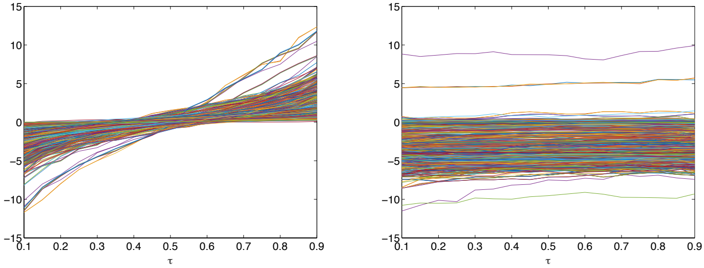


Figure: Estimated quantile factor loading processes for the constant (left) and $\hat{F}_t$ (right).

FF Portfolio Returns – 2-step procedure

- Monthly excess returns in FF portfolios (Fama & French, 1993), 1985-2012, $(N = 100, T = 324)$, $r = 3+$ intercept.

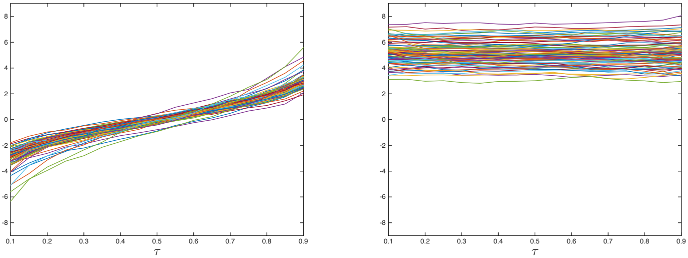


Figure: FF portfolios: estimated quantile factor loading processes for the constant (left), $\hat{F}_{1t}$ (right).

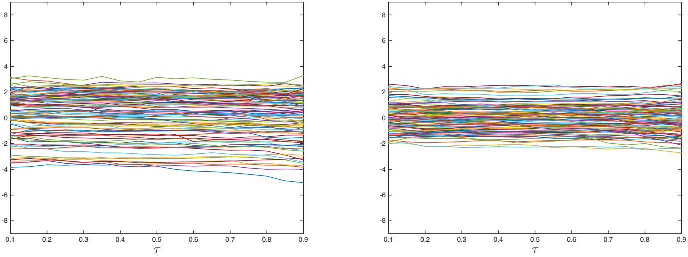


Figure: FF portfolios: estimated quantile factor loading processes for $\hat{F}_{2t}$ (left) and $\hat{F}_{3t}$ (right).

Testing for Constancy of QFM loadings

This is Case 3 in the Inference slide.

Table: Testing for constant quantile factor loading processes.

Critical Values	10%(1.94)	5%(2.22)	1%(2.80)
Common Stocks	303(70.63%)	341(79.49%)	382(89.04%)
Mutual Funds	1894(80.05%)	1994(84.28%)	2086(88.17%)
FF portfolios	266(82.10%)	287(88.58%)	295(91.04%)

Model Checks

Two set of R^2 to check the models:

1. If the two-step procedure works, the iterative procedure should also work, and R^2 in the regressions of $\hat{F}_{PC}$ on $\hat{F}_{QR}(\tau)$ should be close to 1 for all τs.
2. If the extra factor captured by $\hat{F}_{QR}(\tau)$ but not by $\hat{F}_{PC}$ is a constant factor, R^2 in the regressions of a constant term on $\hat{F}_{QR}(\tau)$ should be close to 1 for most τs. Otherwise there may exist extra quantile factors that can not captured by $\hat{F}_{PC}$.

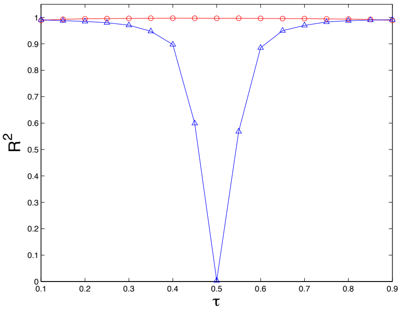


Figure: R^2 of regressing $\hat{F}_{PC}$ on $\hat{F}_{QR}$ (red) and regressing a constant term on $\hat{F}_{QR}$ (blue) for $\tau = 0.1, 0.15, \ldots, 0.9$ from a simulated location-shift factor model with $r = 1$ and $N = T = 100$.

Model Checks

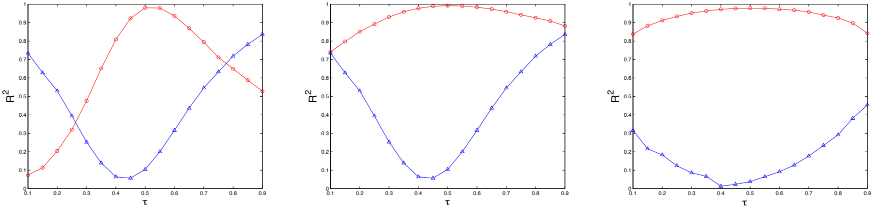


Figure: R^2 of regressing $\hat{F}_{PC}$ on $\hat{F}_{QR}$ (red) and regressing a constant term on $\hat{F}_{QR}$ (blue) for $\tau = 0.1, 0.15, \ldots, 0.9$. **Left**: stock returns; **Middle**: mutual fund returns; **Right**: FF portfolios.

Further Research Agenda

- Let the number of factors vary by τ in the iterative procedure.
- Construction of an IC to select the number of factors for each τ.
- Applications where extra quantile factors ($\tau \neq 0.5$) imply better forecasts in Factor Augmented Regression Models.
- Derivation of the asymptotic distribution of estimated factors by the iterative procedure.
- ....

Thank you !

$$X_{it} = \lambda_i f_t + g_t \epsilon_{it} = \lambda_i f_t + Q_\epsilon(U_{it}) g_t = \lambda_i'(U_{it}) F_t$$

where

$$\lambda_i(U_{it}) = (\lambda_i, Q_\epsilon(U_{it}))' \quad F_t = (f_t, g_t)'.$$

However,

$$\lambda_i = E[\lambda_i(U_{it})] = E[\lambda_i, Q_\epsilon(U_{it})] = (\lambda_i, 0),$$

thus Assumption 1(iii) is violated because

$$\lim_{N\to\infty} N^{-1} \sum_{i=1}^{N} \lambda_i \lambda_i' \quad \text{is not full rank.}$$

Back