

DOCUMENTOS DE TRABAJO

High Dimensional Quantile Factor Analysis

Andres Sagner

N° 886 Agosto 2020

BANCO CENTRAL DE CHILE





BANCO CENTRAL DE CHILE

CENTRAL BANK OF CHILE

La serie Documentos de Trabajo es una publicación del Banco Central de Chile que divulga los trabajos de investigación económica realizados por profesionales de esta institución o encargados por ella a terceros. El objetivo de la serie es aportar al debate temas relevantes y presentar nuevos enfoques en el análisis de los mismos. La difusión de los Documentos de Trabajo sólo intenta facilitar el intercambio de ideas y dar a conocer investigaciones, con carácter preliminar, para su discusión y comentarios.

La publicación de los Documentos de Trabajo no está sujeta a la aprobación previa de los miembros del Consejo del Banco Central de Chile. Tanto el contenido de los Documentos de Trabajo como también los análisis y conclusiones que de ellos se deriven, son de exclusiva responsabilidad de su o sus autores y no reflejan necesariamente la opinión del Banco Central de Chile o de sus Consejeros.

The Working Papers series of the Central Bank of Chile disseminates economic research conducted by Central Bank staff or third parties under the sponsorship of the Bank. The purpose of the series is to contribute to the discussion of relevant issues and develop new analytical or empirical approaches in their analyses. The only aim of the Working Papers is to disseminate preliminary research for its discussion and comments.

Publication of Working Papers is not subject to previous approval by the members of the Board of the Central Bank. The views and conclusions presented in the papers are exclusively those of the author(s) and do not necessarily reflect the position of the Central Bank of Chile or of the Board members.

Documentos de Trabajo del Banco Central de Chile
Working Papers of the Central Bank of Chile
Agustinas 1180, Santiago, Chile
Teléfono: (56-2) 3882475; Fax: (56-2) 3882231

High Dimensional Quantile Factor Analysis*

Andres Sagner
Central Bank of Chile

Abstract

In this paper, I develop a method that extends quantile regressions to high dimensional factor analysis. In this context, the quantile function of a panel of variables with crosssection and time-series dimensions N and T is endowed with a factor structure. Thus, both factors and factor loadings are allowed to be quantile-specific. I provide a set of conditions under which these objects are identified, and I propose a simple two-step iterative procedure called Quantile Principal Components (QPC) to estimate them. Uniform consistency of the estimators is established under general assumptions when $N, T \rightarrow \infty$ jointly. Lastly, under certain additional assumptions related to the density of the observations about the quantile of interest, and the relationship between N and T , I show that the QPC estimators are asymptotically normal with convergence rates similar to the ones derived in the traditional factor analysis literature. Monte Carlo simulations confirm the good performance of the QPC procedure, especially in non-linear environments, or when the factors affect higher moments of the observable variables and suggest that the proposed theory provides a good approximation to the finite sample distribution of the QPC estimators.

Resumen

En este artículo, desarrollo un método que extiende las regresiones cuantiles al análisis factorial de alta dimensión. En este contexto, la función cuantil de un panel de variables que posee N elementos observados durante T períodos se encuentra dotada de una estructura factorial. De este modo, la magnitud de los factores y las cargas factoriales pueden ser función de los cuantiles. Proporciono un conjunto de condiciones bajo las cuales estos objetos se encuentran identificados y propongo un procedimiento iterativo simple de dos pasos denominado Componentes Principales Cuantiles (CPC) para estimarlos. Bajo supuestos generales, establezco la consistencia uniforme de los estimadores cuando $N, T \rightarrow \infty$ conjuntamente. Finalmente, bajo ciertos supuestos adicionales relacionados a la densidad de observaciones en torno al cuantil de interés y a la relación entre N y T , muestro que los estimadores CPC son asintóticamente normales, con tasas de convergencia similares a aquellas derivadas en la literatura tradicional de análisis factorial. Simulaciones de Monte Carlo confirman el buen desempeño del procedimiento de estimación CPC, especialmente en ambientes no-lineales, o cuando los factores afectan momentos superiores de las variables observadas. Estos resultados sugieren que la teoría propuesta proporciona una buena aproximación a la distribución de muestras finitas de los estimadores CPC.

*I want to thank Dante Amengual, Timothy Christensen, Ivan Fernandez-Val, Jean-Jacques Forneron, Hiroaki Kaido, Oliver Linton, Pierre Perron, Zhongjun Qu, Nick Saponara, an anonymous referee, and econometrics seminar participants at Boston University for helpful discussions and insightful suggestions. Contact: Financial Stability Area, Central Bank of Chile. Email: asagner@bcentral.cl.

1 Introduction

During the last decades, high dimensional factor analysis has become an increasingly popular and useful statistical tool in many economic applications. Its popularity resides in the fact that it is a practical and easy way to summarize the information in large data sets into a small number of unobserved variables that describe a mean curve jointly. For instance, factor analysis has been used to model asset returns as a function of a small number of risk factors (Ross, 1976; Connor and Korajczyk, 1988); decompose the business cycle into common and specific shocks at the cross-country level (Gregory and Head, 1999; Forni *et al.*, 2000; Crucini *et al.*, 2011; Karadimitropoulou and Leon-Ledesma, 2013), national level (Stock and Watson, 1989; Mariano and Marasawa, 2003; Aruoba *et al.*, 2009), and industry-level (Forni and Reichlin, 1998); improve forecasting models by including the so-called diffusion indexes (Stock and Watson, 1999, 2002); and construct measures of systemic risk (Kritzman *et al.*, 2011), macroeconomic or financial uncertainty (Jurado *et al.*, 2015) and network connectedness (Billio *et al.*, 2012) which are vital for policymakers to perform macro and financial stability monitoring; among many other applications.

In this paper, I extend the quantile regression approach popularized by Koenker and Bassett (1978) to high dimensional factor analysis. I name this concept as high dimensional *Quantile Factor Analysis* (QFA). In this setup, for any scalar $\tau \in (0, 1)$, the τ -th conditional quantile function of a panel consisting of N variables y_{it} observed along T periods, $\mathcal{Q}_{y_{it}}(\tau | \lambda_i^0(\tau), f_t^0(\tau))$, is a linear function of $K(\tau) < \min\{N, T\}$ unobserved quantile-specific factors $f_t^0(\tau)$ that are known by the econometrician. Moreover, both N and T are large, and the number of quantile-specific factors, $K(\tau)$, as well as the sensitivity (or factor loading) of each variable i to each quantile factor, $\lambda_i^0(\tau)$, are also permitted to be quantile-specific. In this manner, the proposed setup captures the idea of a *quantile factor model* which has the particularity of being flexible enough to characterize several linear and nonlinear factor models available in the related literature.

Under standard assumptions, I show that both the quantile factors $f_t^0(\tau)$ and the quantile factor loadings $\lambda_i^0(\tau)$ are individually identified. The type of identification depends crucially on the rotation chosen by the econometrician. In particular, the identification of $f_t^0(\tau)$ and $\lambda_i^0(\tau)$ is local if an orthogonal rotation, extensively used in Principal Components (PC), or a recursive-type rotation is considered. In contrast, it is global once an errors-in-variables-type rotation is employed. However, the identification of the quantile common component $c_{it}^0(\tau) = \lambda_i^0(\tau)' f_t^0(\tau)$ is always a global one. Moreover, I show that if the ordering of the observable variables y_{it} is known in advance, i.e., we know which variable is affected by which quantile factor, then all previous rotations deliver observationally equivalent quantile

common components.

Then, I propose a simple two-step iterative procedure based on the minimization of the quantile loss function to obtain the *Quantile Principal Components* (QPC) estimator of $f_t^0(\tau)$ and $\lambda_i^0(\tau)$ for any $\tau \in (0, 1)$. In the first step, the estimator of the quantile factors, $\hat{f}_t(\tau)$, is computed using quantile regressions across cross-sections for each t , where the unobserved quantile loadings are replaced by an initial guess. In the second step, the estimator of the quantile factor loadings, $\hat{\lambda}_i(\tau)$, is computed using quantile regressions across periods for each i , given the previous estimates for the quantile factors. This estimation procedure offers some advantages in terms of efficiency, compared to the PC methodology, especially in nonlinear setups or in factor models where the factors impact higher moments of the observable variable y_{it} .

Also, I establish the uniform consistency of both $\hat{f}_t(\tau)$ and $\hat{\lambda}_i(\tau)$ under general assumptions. In the proof, I proceed as in [Chen et al. \(2014\)](#) and show first the uniform consistency of the QPC estimator of the quantile common component $\hat{c}_{it}(\tau) \equiv \hat{\lambda}_i(\tau)' \hat{f}_t(\tau)$ as $N, T \rightarrow \infty$ jointly, given that the objective function involved in the minimization of the quantile loss function is convex in terms of this object. This feature, together with the compactness of the parameter set, allows me to invoke a standard Uniform Law of Large Numbers argument. Then, given this intermediate result, uniform consistency of $\hat{f}_t(\tau)$ and $\hat{\lambda}_i(\tau)$ follows from the assumptions imposing a strong factor structure.

Lastly, under additional assumptions related to the sparsity of the observations around the quantile of interest $\tau \in (0, 1)$ and the relationship between N and T , I show that the QPC estimators $\hat{f}_t(\tau)$, $\hat{\lambda}_i(\tau)$ and $\hat{c}_{it}(\tau)$ are asymptotically Normal. In the latter case, the convergence rate is slower than the one found by [Bai \(2003\)](#) for the common component of standard high dimensional factor models. The proof relies on a uniform asymptotic approximation of the subgradients evaluated at the QPC estimators using an argument similar to the one employed by [Qu \(2008\)](#). Then, I show that the approximation admits a Bahadur representation, i.e., both $\sqrt{N}(\hat{f}_t(\tau) - f_t^0(\tau))$ and $\sqrt{T}(\hat{\lambda}_i(\tau) - \lambda_i^0(\tau))$ can be expressed as a normalized sum of martingale difference sequences plus an $o_p(\cdot)$ term. Asymptotic normality of these estimators follows then by a standard Uniform Central Limit Theorem argument. The limiting distribution of the common component, on its part, is derived by showing that $\hat{c}_{it}(\tau) - c_{it}^0(\tau)$ can be approximated by the sum of two random variables related to the two differences mentioned above.

This paper is related to the literature on panel data models in which the error component contains an interactive effect (a factor structure), e.g., [Koenker \(2004\)](#), [Pesaran \(2006\)](#), [Bai \(2009\)](#), [Kato et al. \(2012\)](#), [Bai and Li \(2014\)](#), [Harding and Lamarche \(2014\)](#), [Moon and Weidner \(2015, 2017\)](#), [Fernandez-Val and Weidner \(2016\)](#), among others. Albeit its similarity

with this setting, these models differ from quantile factor models in at least two key aspects. First, in the QFA context, the regressors of the model are the factors which, besides being quantile-specific, are not observable by the econometrician, entailing in this manner several estimation challenges. So, this paper contributes to the literature by providing an estimation methodology that is easy to implement, even in nonlinear environments. Second, in most of these models, the unobserved individual and time heterogeneity is treated as nuisance parameters. Consequently, a large part of the analysis is devoted to the properties of fixed or random effects estimator. In contrast, the properties of the factors and the factor loadings are barely explored. I contribute to this strand of the related literature by analyzing the asymptotic properties of these objects in a high dimensional quantile framework.

The rest of the paper is organized as follows. Section 2 starts by presenting the statistical model behind high dimensional QFA and provides some examples to illustrate this concept. Section 2.2 discusses in detail the individual identification of the quantile factors and quantile factor loadings, while Section 2.3 presents the iterative procedure to obtain their QPC estimators and highlights some of their properties. Section 3 provides the set of assumptions required to establish the uniform consistency and the asymptotic distribution for the QPC estimator of the quantile factors, quantile factor loadings, and quantile common components. Some aspects concerning the computation of consistent estimators for the asymptotic variance-covariance matrices are discussed in Section 3.2.1. Finally, Section 4 concludes and suggests additional elements that can be tackled by future research on this topic. All proofs of primary and intermediate results are given the Appendix.

2 Model and Estimation

In this section, I present the data generating process behind high dimensional Quantile Factor Analysis. Next, I provide a set of conditions under which the data identifies the relevant parameters of the model. Finally, I propose an iterative algorithm to estimate the quantile-specific factors and factor loadings, and I also discuss some of its properties in detail.

2.1 The Model

The main idea behind the traditional high dimensional factor analysis is that the behavior across T periods of a set of N observed random variables can be characterized by a linear combination of $K < \min\{N, T\}$ factors plus an error term. Formally,

$$y_{it} = \lambda_i' f_t^0 + e_{it}^0$$

where y_{it} is the i -th observed random variable at time t , f_t^0 is a vector containing K factors, λ_i^0 is a vector of factor loadings or sensitivities of the i -th variable to each factor, e_{it}^0 is an *iid* error term, and the superscript “0” stands for true or population parameters. The product $c_{it}^0 \equiv \lambda_i^0 f_t^0$ is typically known as the *common component* of y_{it} , whereas the error term is sometimes called the *idiosyncratic component*. Also, the theory underlying high dimensional factor analysis allows both N and T to be large, and it assumes that the number of factors K is known¹. Finally, note that all elements on the right-hand side of the above equation are not observable by the econometrician.

In this paper, I extend the traditional high dimensional factor analysis model by allowing the factors or the factor loadings, or both, to be a function of a random variable u_{it} distributed uniformly over the interval $[0, 1]$. To be precise, I consider that the dynamics of the observable variable y_{it} is dictated by the following data generating process

$$y_{it} = \lambda_i^0(u_{it})' f_t^0(u_{it}), \quad u_{it} \sim \mathcal{U}[0, 1] \quad (1)$$

Assumption 1. For all i , t , and $\tau \in (0, 1)$, the common component $c_{it}^0(\tau) \equiv \lambda_i^0(\tau)' f_t^0(\tau)$ is nondecreasing in τ .

Let $\tau \in (0, 1)$ and $G(\cdot | \theta_{it}^0(\tau))$ be the cumulative distribution function of y_{it} conditional on $\theta_{it}^0(\tau) \equiv [\lambda_i^0(\tau)', f_t^0(\tau)']'$. Under Assumption 1, the τ -th conditional quantile function of the observable variable y_{it} given $\theta_{it}^0(\tau)$, $\mathcal{Q}_{y_{it}}(\tau | \theta_{it}^0(\tau)) \equiv \inf \{y_{it} : G(y_{it} | \theta_{it}^0(\tau)) \geq \tau\}$, is given by

$$\mathcal{Q}_{y_{it}}(\tau | \theta_{it}^0(\tau)) = \lambda_i^0(\tau)' f_t^0(\tau), \quad \tau \in (0, 1) \quad (2)$$

where the number of factors, $K(\tau)$, is also allowed to be quantile-specific. In other words, the above equation says that all conditional quantiles of the observable random variable y_{it} have a factor model structure. So, equation (1) summarizes the idea of a *Quantile Factor Analysis* model and, consequently, we refer to $f_t^0(\tau)$ and $\lambda_i^0(\tau)$ in equation (2) as *quantile factors* and *quantile factor loadings*, respectively. At first glance, the linearity of the proposed framework may seem restrictive. However, as will be seen in the next examples, equation (1) can nest several nonlinear factor model structures.

Example 1 (Standard Factor Model). Let $y_{it} = \alpha_i^0 \beta_t^0 + v_{it}^0$, where both α_i^0 and β_t^0 are scalars, and v_{it}^0 is an *iid* random variable with cumulative distribution function $G_v(\cdot)$. By defining $v_{it}^0 \equiv G_v^{-1}(u_{it})$, where $u_{it} \sim \mathcal{U}[0, 1]$ for all i and t , the standard factor model can

¹If this assumption is relaxed, then K can be consistently estimated from the data by using the information-criteria-based tests proposed by [Bai and Ng \(2002\)](#), or the testing procedure presented in [Onatski \(2009\)](#), [Kapetanios \(2010\)](#), and [Ahn and Horenstein \(2013\)](#).

be rewritten in the form of (1) by setting $\lambda_i^0(u_{it}) = [\alpha_i^0, 1]'$ and $f_t^0(u_{it}) = [\beta_t^0, G_v^{-1}(u_{it})]'$, or $\lambda_i^0(u_{it}) = [\alpha_i^0, G_v^{-1}(u_{it})]'$ and $f_t^0(u_{it}) = [\beta_t^0, 1]'$.

Example 2 (Location-Scale Factor Model). Let $y_{it} = \alpha_i^0 \beta_t^0 + \gamma_t^0 v_{it}^0$, where $\gamma_t^0 > 0$ for all t . Similar to the previous case, this model can be rewritten in the form of (1) by defining $\lambda_i^0(u_{it}) = [\alpha_i^0, G_v^{-1}(u_{it})]'$ and $f_t^0(u_{it}) = [\beta_t^0, \gamma_t^0]'$, or $\lambda_i^0(u_{it}) = [\alpha_i^0, 1]'$ and $f_t^0(u_{it}) = [\beta_t^0, \gamma_t^0 G_v^{-1}(u_{it})]'$, where $v_{it}^0 = G_v^{-1}(u_{it})$ with $u_{it} \sim \mathcal{U}(0, 1)$ for all i and t .

Example 3 (Nonlinear Factor Model). Let $y_{it} = \alpha_i^0 \beta_t^0 e^{v_{it}^0}$, where $\alpha_i^0, \beta_t^0 > 0$ for all i and t , respectively. If $\lambda_i^0(u_{it}) = \alpha_i^0 e^{a G_v^{-1}(u_{it})}$ and $f_t^0(u_{it}) = \beta_t^0 e^{(1-a) G_v^{-1}(u_{it})}$, $a \in [0, 1]$, then this model has the form of (1).

The examples exhibited above are only a few out of many cases where a QFA model can represent a factor model structure. Example 1 corresponds to the standard linear case where both the factors and the factor loadings affect only the mean of the observable variable, i.e., the homoskedastic case. Its configuration implies that only the factors (or the loadings) are quantile-specific and that one of the factor loadings (or factors) has to be equal to 1. This configuration implies, in turn, that the quantile functions given by (2) are simply a vertical displacement of one another.

A somewhat more complicated case is considered in Example 2. In this context, the factors affect not only the mean of the observable variable but also its variance. Thus, the heteroskedasticity of this model is proportional to the square of γ_t^0 . Moreover, two key aspects of this example are worth highlighting. First, Assumption 1 imposes an additional restriction to the domain of one of the factors ($\gamma_t^0 > 0$), which suggests that equations (1) and (2) are not necessarily equivalent, the latter being the most restrictive one. Second, the number of factors depends indeed on τ . In particular, if the idiosyncratic component v_{it}^0 is symmetric about the origin, then the conditional quantile function evaluated at the median is equal to 0 and $f_t^0(0.5) = \beta_t^0$, i.e., $K(0.5) = 1$. On the contrary, for any $\tau \neq 0.5$, the quantile function is different from 0 and, consequently, $f_t^0(\tau) = [\beta_t^0, \gamma_t^0]$ and $K(\tau) = 2$.

Finally, Example 3 is a nonlinear factor model describing the behavior of a strictly positive observable variable. The data generating process implies that either the factors or the factor loadings, or both, are quantile-specific. Lastly, note that $\log y_{it} = \log \alpha_i^0 + \log \beta_t^0 + G_v^{-1}(u_{it})$, that is, the factor model is linear for $\log y_{it}$ so we can define $\tilde{\lambda}_i^0(u_{it}) = [\log \alpha_i^0 + a G_v^{-1}(u_{it}), 1]'$ and $\tilde{f}_t^0(u_{it}) = [1, \log \beta_t^0 + (1-a) G_v^{-1}(u_{it})]$, $a \in [0, 1]$, and the transformed model has the form of a QFA model.

The matrix representation of equation (2) is given by

$$\mathcal{Q}_Y(\tau | \boldsymbol{\theta}^0(\tau)) = F^0(\tau) \Lambda^0(\tau)', \quad \tau \in (0, 1)$$

where Y is a $T \times N$ matrix of observable variables, $F^0(\tau) = [f_1^0(\tau), \dots, f_T^0(\tau)]' \in \Theta_F \subset \mathbb{R}^{T \times K(\tau)}$ is a $T \times K(\tau)$ matrix of quantile factors, $\Lambda^0(\tau) = [\lambda_1^0(\tau), \dots, \lambda_N^0(\tau)]' \in \Theta_\Lambda \subset \mathbb{R}^{N \times K(\tau)}$ is a matrix of quantile factor loadings of dimension $N \times K(\tau)$, and $\boldsymbol{\theta}^0(\tau) \equiv [\Lambda^0(\tau)', F^0(\tau)']'$. The $T \times N$ matrix $C^0(\tau) \equiv F^0(\tau) \Lambda^0(\tau)'$ contains all the common components of the QFA model.

2.2 Identification

In this section, I provide a set of assumptions under which the population quantile factor loadings and quantile factors, $\boldsymbol{\theta}^0(\tau)$, are identified by the data. I start by defining identification in this context.

Definition 1 (Identification). For all $\tau \in (0, 1)$, let $\boldsymbol{\theta}(\tau) = [\Lambda(\tau)', F(\tau)']'$ be a parameter matrix. We say that $\boldsymbol{\theta}(\tau)$ is identified at $\boldsymbol{\theta}^0(\tau) \in \Theta_\Lambda \times \Theta_F$ based on the quantile loss function $\rho_\tau(u) = (\tau - \mathbb{1}\{u < 0\})u$, where $\mathbb{1}\{\cdot\}$ is the indicator function, when

$$\boldsymbol{\theta}(\tau) = \arg \min_{[\Lambda', F']' \in \Theta_\Lambda \times \Theta_F} S_\tau(\Lambda, F) \quad (3)$$

where

$$S_\tau(\Lambda, F) = E \left[\sum_{i=1}^N \sum_{t=1}^T \rho_\tau(y_{it} - \lambda_i' f_t) \right] \quad (4)$$

if and only if $\boldsymbol{\theta}(\tau) = \boldsymbol{\theta}^0(\tau)$.

Definition 1 highlights the point that identification of $\boldsymbol{\theta}^0(\tau)$ depends crucially on whether we can find the minimizer of the objective function $S_\tau(\Lambda, F)$. However, this task is not as straightforward as it appears. In this sense, and as noted by [Koenker and Bassett \(1978\)](#) and [Koenker \(2005, pp. 32-33\)](#), the quantile loss function $\rho_\tau(\cdot)$ is continuous, but piecewise linear and not everywhere differentiable. So, to achieve the identification of the model's parameters, I provide below a set of conditions ensuring the existence of a minimizer.

Assumption 2 (Identification).

1. For all i, t , and $\tau \in (0, 1)$, the observable random variable y_{it} is generated by the QFA model (1) - (2) and has absolutely continuous conditional cumulative distribution functions $G_{it}(\cdot | \theta_{it}^0(\tau))$ and continuous, strictly positive conditional densities $g_{it}(\cdot | \theta_{it}^0(\tau))$.
2. For all $\tau \in (0, 1)$, $\text{rank}(C^0(\tau)) = K(\tau)$.
3. For all $\tau \in (0, 1)$, any of the following restrictions (or rotations) apply

- (a) $F^0(\tau)' F^0(\tau) / T = I_{K(\tau)}$, where $I_{K(\tau)}$ is the identity matrix of size $K(\tau)$; and $\Lambda^0(\tau)' \Lambda^0(\tau)$ is a diagonal matrix of size $K(\tau)$, whose diagonal elements being positive, distinct, and arranged in decreasing order.
- (b) $F^0(\tau)' F^0(\tau) / T = I_{K(\tau)}$; and $\Lambda^0(\tau) = [\Lambda_1^0(\tau)', \Lambda_2^0(\tau)']'$, where $\Lambda_1^0(\tau)$ is a lower triangular matrix of size $K(\tau)$ with non-zero diagonal elements.
- (c) $\Lambda^0(\tau) = [I_{K(\tau)}, \Lambda_2^0(\tau)']'$.

Assumptions 2.1 and 2.2 allow for the identification of the common component $C^0(\tau)$. In particular, a strictly positive density of y_{it} conditional on $\lambda_i^0(\tau)$ and $f_t^0(\tau)$ (i.e., $g_{it}(\cdot | \theta_{it}^0(\tau)) > 0$) implies that the quadratic approximation of the population objective function (4) centered around $\theta^0(\tau)$ attains a global minimum at $C^0(\tau)$. Given that the latter is of full rank by Assumption 2.2, i.e., the system of linear equations derived from the first-order conditions are non-degenerate, then the global minimum is unique. Assumption 2.3, on the other hand, identifies $F^0(\tau)$ and $\Lambda^0(\tau)$ separately. To see this, note that for any $K(\tau) \times K(\tau)$ invertible matrix A we have that $C^0(\tau) = F^0(\tau) \Lambda^0(\tau)' = F^0(\tau) A A^{-1} \Lambda^0(\tau)' = \tilde{F}^0(\tau) \tilde{\Lambda}^0(\tau)' = \tilde{C}^0(\tau)$, where $\tilde{F}^0(\tau) = F^0(\tau) A$ and $\tilde{\Lambda}^0(\tau) = \Lambda^0(\tau) A^{-1'}$. Because both common components are observationally equivalent, an additional structure must be imposed to determine the quantile factors and quantile loadings uniquely. Since there are many ways to restrict $F^0(\tau)$ and $\Lambda^0(\tau)$, Assumption 2.3 provides three alternative, more or less arbitrary sets of rotations that have been primarily used in traditional factor analysis models (see, for example, [Anderson and Rubin, 1956](#))². Assumption 2.3a is the default rotation in principal component analysis via maximum likelihood estimation (see [Jolliffe, 2002](#), pp. 270-274). It is, in essence, a statistical rotation since it allows us to concentrate out the factor loadings from the principal components optimization problem. Consequently, the resulting factors correspond to \sqrt{T} times the eigenvectors associated with the $K(\tau)$ largest eigenvalues of the matrix $Y'Y$. Assumption 2.3b, on its part, requires $\Lambda_1^0(\tau)$ to be an invertible lower triangular matrix. This configuration implies that the first quantile factor affects the first observable variable only; the first two quantile factors affect the first two observable variables only, and so on up to the $K(\tau)$ -th quantile factor. Afterward, all observable variables are affected by all quantile factors. Because of its similarity with a triangular system of simultaneous equations, the related literature refers to it as *recursive rotation*. It is frequently used in empirical research (see, for example, [Geweke and Zhou, 1996](#)). Finally, Assumption 2.3c is related to the measurement

²[Bai and Li \(2012\)](#), and [Bai and Ng \(2013\)](#) mention that three more related rotations can be obtained from Assumption 2.3 by switching the role of $F^0(\tau)$ and $\Lambda^0(\tau)$. For instance, in Assumption 2.3a we can alternatively consider that $\Lambda^0(\tau)' \Lambda^0(\tau) / N = I_{K(\tau)}$ and $F^0(\tau)' F^0(\tau)$ is a diagonal matrix of size $K(\tau)$ with all its diagonal elements being positive, distinct, and arranged in decreasing order. I will not consider them in this paper, but all results are straightforwardly extensible to this alternative set of rotations.

error literature, which implies that the first $K(\tau)$ observable variables are noisy measures of the corresponding quantile factors (see [Wansbeek and Meijer, 2000](#), pp. 148-150). Hence its name *errors-in-variables rotation*. Note that, unlike the two previous cases, this rotation imposes all the restrictions on the quantile loadings and, therefore, leaves the quantile factors unrestricted.

Although Assumptions 2.1 and 2.2 ensure together the existence of a unique quantile common component that minimizes (3), the choice of a particular rotation is not innocuous for the type of identification attained by $F^0(\tau)$ and $\Lambda^0(\tau)$ individually. This choice is an issue that has been discussed since [Algina \(1980\)](#) and [Bekker \(1986\)](#), among many others. In particular, Assumptions 2.3a and 2.3b are local identification conditions, whereas Assumption 2.3c is a global identification one. In the former cases, identification is only up to a column-sign change because both $F^0(\tau)$ and $-F^0(\tau)$, and $\Lambda^0(\tau)$ and $-\Lambda^0(\tau)$ satisfy the restrictions imposed by these rotations and deliver the same common component. To see this point, suppose that we have identified $C^0(\tau)$. Then, orthogonality of the quantile factors under Assumptions 2.3a, or 2.3b implies that $C^0(\tau)'C^0(\tau)/T = \Lambda^0(\tau)'\Lambda^0(\tau)$. Finally, because the common component is of full rank, we can identify the magnitude of each column of $\Lambda^0(\tau)$ but not its sign. Thus, after fixing the sign of each column of $\Lambda^0(\tau)$ (or $F^0(\tau)$), the rotations become global identification restrictions³. Note, furthermore, that there is another source of indeterminacy associated with rotations 2.3a and 2.3b. If we switch positions between the k -th and $(k+1)$ -th columns of $F^0(\tau)$ and of $\Lambda^0(\tau)$, the common component remains unchanged, implying that an ordering restriction needs to be imposed. That is exactly what the last part of Assumption 2.3a does to avoid this issue: it arranges the diagonal elements of the matrix $\Lambda^0(\tau)'\Lambda^0(\tau)$ in decreasing order. As for Assumption 2.3b, the ordering restriction is imposed in terms of specifying which variable is affected by which factors, plus a non-zero restriction to all diagonal elements of the matrix $\Lambda_1^0(\tau)$. Otherwise, the k -th and $(k+1)$ -th columns of $\Lambda^0(\tau)$ will share the same structure, and, consequently, the common component will violate Assumption 2.2. To understand why Assumption 2.3c achieves global identification of the quantile factors and the quantile loadings, consider the following partition of the quantile common component $C^0(\tau) = [C_1^0(\tau), C_2^0(\tau)]$, where $C_1^0(\tau)$ and $C_2^0(\tau)$ are of dimension $T \times K(\tau)$ and $T \times (N - K(\tau))$, respectively. Therefore, $F^0(\tau)$ and $\Lambda_2^0(\tau)$ are uniquely identified from $F^0(\tau) = C_1^0(\tau)$ and $\Lambda_2^0(\tau) = C_2^0(\tau)'F^0(\tau)(F^0(\tau)'F^0(\tau))^{-1}$, respectively. Finally, the choice of observable variables that are assumed to be noise measurements of the $K(\tau)$ underlying factors avoids the ordering indeterminacy of this rotation.

Definition 2 (Equivalence of Common Components). For all $\tau \in (0, 1)$, we say that

³An alternative way to achieve global identification under Assumption 2.3b is by normalizing to 1 all diagonal elements of the matrix $\Lambda_1^0(\tau)$.

two common components $C_1^0(\tau)$ and $C_2^0(\tau)$ with respective parameter matrices $\theta_1^0(\tau) \in \Theta_1$ and $\theta_2^0(\tau) \in \Theta_2$ are equivalent if there exists a one-to-one transformation between $\theta_1^0(\tau)$ and $\theta_2^0(\tau)$ throughout Θ_1 and Θ_2 such that $C_1^0(\tau) = C_2^0(\tau)$.

Proposition 1 (Equivalence of Rotations). *Suppose that the ordering of the observable variables Y is known and Assumption 2.2 is satisfied. Then, for all $\tau \in (0, 1)$, the rotations described in Assumption 2.3 are equivalent.*

Proposition 1 indicates that the rotations described in Assumption 2.3 yield common components that are observationally equivalent. To achieve this equivalence, we necessarily need to know the ordering of the observable variables contained in Y , a process that, in some cases, is user-specified but, in other cases, is accommodated by a structural model (see [Skrdal and Rabe-Hesketh, 2004](#), pp. 108-112).

The equivalence of rotations, as stated in Proposition 1, is an essential feature of Assumption 2.3 in at least two dimensions. First, if the interest of the econometrician is to model the τ -th quantile function of some observable variables, then the choice of rotations is irrelevant. Second, and more importantly, the equivalence is vital in the estimation of QFA models such as (2) in the sense that one can use the set of identifying restrictions that pose the less restrictive rotation in terms of computational complexity and processing time. I will discuss in detail this last point in the next section.

Next, I establish the first main result of this paper, namely the individual identification of the quantile factors and the quantile factor loadings.

Theorem 1 (Identification). *Suppose that Assumption 2 holds. Then, for every $\tau \in (0, 1)$, both $F^0(\tau)$ and $\Lambda^0(\tau)$ are identified.*

The intuition behind Theorem 1 is as follows. The quantile factors and quantile factor loadings of model (2) are individually identified as the minimizer of the population optimization problem given by (3) - (4). To achieve this goal, the theorem considers a quadratic approximation of the objective function (4) centered around $\theta^0(\tau)$. This procedure has the crucial feature that the global minimum is attained just at $\theta^0(\tau)$, for all $\tau \in (0, 1)$, subject to a particular rotation.

2.3 The QPC Estimator

In this section, I present the proposed algorithm to obtain the Quantile Principal Components (QPC) estimator of both the quantile factors and quantile factor loadings of the QFA model (2). Then, I discuss some of its properties, namely its convergence and computational complexity. I finalize the section with a finite-sample properties analysis of the QPC estimator relative to the Principal Components (PC) estimator.

I start by discussing two key issues related to the QPC estimator. First, let us consider the sample analog of the objective function (4), $V_\tau(\Lambda, F)$, defined as

$$V_\tau(\Lambda, F) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \rho_\tau(y_{it} - \lambda'_i f_t) \quad (5)$$

for all $\tau \in (0, 1)$, which is a convex function in terms of the common component $C = F\Lambda'$. Nevertheless, for any value of $\tau \in (0, 1)$, this function is not simultaneously convex in both Λ and F . But note that, when either of these two arguments is kept fixed, then the sample analog of the objective function is a convex function, i.e., if Λ is kept fixed at, say $\bar{\Lambda}$, then $V_\tau(\bar{\Lambda}, F)$ is convex in F . Similarly, if F is kept fixed at \bar{F} , then $V_\tau(\Lambda, \bar{F})$ is a convex function in Λ . Thus, this key feature of the sample analog of the objective function given by equation (5) motivates a two-step iterative procedure for obtaining the QPC estimator of $\theta^0(\tau)$.

Second, and as discussed in the previous section, the individual identification of the quantile factors and quantile factor loadings requires further restrictions on these parameters. Assumption 2.3 serves this purpose, and according to Proposition 1, all rotations considered in this assumption are equivalent if we know the ordering of the observable variables Y . This feature means that we can use any of the identifying restrictions to obtain the QPC estimator of $\theta^0(\tau)$. In this sense, Assumption 2.3a imposes nonlinear restrictions on both the quantile factors and the quantile factor loadings. In contrast, the recursive rotation (Assumption 2.3b) imposes nonlinear restrictions on the quantile factors and linear restrictions on the quantile factor loadings. Finally, the errors-in-variables rotation (Assumption 2.3c) considers linear restrictions on the quantile factor loadings only and leaves the quantile factors unrestricted. Thus, in terms of computational complexity, as we will see later in this section, the last rotation is the most convenient one to obtain the QPC estimator of $\Lambda^0(\tau)$ and $F^0(\tau)$.

Definition 3 (Quantile Principal Components Estimator). For any $\tau \in (0, 1)$, the QPC estimator $\hat{\theta}(\tau) = [\hat{\Lambda}(\tau)', \hat{F}(\tau)']'$ of $\theta^0(\tau) = [\Lambda^0(\tau)', F^0(\tau)']'$ can be obtained through the following two-step iterative procedure:

1. Start with initial matrices $\hat{\Lambda}^{(j)}(\tau) = [I_{K(\tau)}, \hat{\Lambda}_2^{(j)}(\tau)']'$ and $\hat{F}^{(j)}(\tau)$.
2. **Step 1:** Fix $\hat{\Lambda}^{(j)}(\tau)$. Then, estimate $\hat{F}^{(j+1)}(\tau)$ from

$$\mathcal{Q}_{Y_t}(\tau | \hat{\Lambda}^{(j)}(\tau)) = \hat{\Lambda}^{(j)}(\tau) f_t(\tau) \quad (6)$$

using quantile regressions for every $t = 1, \dots, T$, where $Y_t = [y_{1t}, \dots, y_{Nt}]'$ is an N -dimensional vector of observable variables.

3. **Step 2:** Fix $\hat{F}^{(j+1)}(\tau)$. Then, estimate $\hat{\Lambda}_2^{(j+1)}(\tau)$ from

$$\mathcal{Q}_{Y_i}(\tau | \hat{F}^{(j+1)}(\tau)) = \hat{F}^{(j+1)}(\tau) \lambda_i(\tau) \quad (7)$$

using quantile regressions for every $i = K(\tau) + 1, \dots, N$, where $Y_i = [y_{i1}, \dots, y_{iT}]'$ is a T -dimensional vector of observable variables.

4. For $\epsilon > 0$ small, if $\left\| \hat{\boldsymbol{\theta}}^{(j+1)}(\tau) - \hat{\boldsymbol{\theta}}^{(j)}(\tau) \right\| < \epsilon$, then $\hat{\boldsymbol{\theta}}(\tau) = \hat{\boldsymbol{\theta}}^{(j+1)}(\tau)$. Else, set $j = j + 1$ and repeat steps 1 and 2 until the previous condition is met.

The intuition behind the algorithm is straightforward. For a given τ , we start by guessing an initial matrix of quantile factor loadings. Note that because we impose the errors-in-variables rotation, the upper $K(\tau) \times K(\tau)$ partition of this guess has to be the identity matrix. Next, fix the guessed quantile factor loadings and obtain an estimate of the quantile factors using quantile regressions across cross-sections for each of the T periods (equation (6)). Now, fix the values of the estimated quantile factors and get an estimate of the quantile factor loadings using quantile regressions across periods for each of the $N - K(\tau)$ unrestricted cross-sections (equation (7)). If the discrepancy between initial guesses and quantile regressions estimates under the Euclidean norm metric is smaller than a predefined accuracy level ϵ , then the algorithm terminates, and the QPC estimator $\hat{\boldsymbol{\theta}}(\tau)$ has been found. Otherwise, repeat the above steps using the estimates of the quantile-specific factors and loadings as starting values.

The name of the QPC estimator comes from its similarity with the PC estimator computed using the EM algorithm (Dempster *et al.*, 1977)⁴. In this context, the factors are treated as the missing piece of information. Under the assumption that the common components are *iid* Normal with known variance, in the E-step of the algorithm, the factors are estimated using OLS across cross-sections given some initial values of the factor loadings. Then, in the M-step, the loadings are estimated using OLS across the time-series dimension, given the estimates of the factors⁵.

Some other algorithms available in the literature similar to my proposed procedure are Alzate and Suykens (2005) and Lim and Oh (2016). The first paper considers alternative objective functions such as Huber and quadratic epsilon intensive loss functions but under a kernel PC analysis framework. Broadly speaking, the algorithm first maps the observable variable onto a feature space using nonlinear functions induced by a kernel. In the second

⁴Although PC can be computed explicitly via the eigendecomposition of the YY' matrix in a very straightforward manner, the EM literature argues that the algorithm is an alternative that offers some attractive features when the econometrician faces high dimensional datasets or missing data.

⁵See Rubin and Thayer (1982), Roweis (1998), or Tipping and Bishop (1999) for further details.

step, it performs linear PC on the mapped data. The last paper, on the other hand, uses a composite quantile, which is a weighted linear combination (data-adaptively determined) of convex modified Huber loss functions instead of square loss functions, to better describe non-Gaussian distributed data. As a consequence, the proposed procedure is a two-step algorithm where the relevant parameters are estimated using the traditional least-squares criterion given specific values of another group of parameters.

2.3.1 Convergence and Complexity of the QPC Estimator

Because the sample analog of the objective function is convex once one of its arguments remains fixed, as highlighted at the beginning of this section, note that

$$V_{\tau}(\hat{\Lambda}^{(j)}(\tau), \hat{F}^{(j)}(\tau)) \geq V_{\tau}(\hat{\Lambda}^{(j)}(\tau), \hat{F}^{(j+1)}(\tau)) \geq V_{\tau}(\hat{\Lambda}^{(j+1)}(\tau), \hat{F}^{(j+1)}(\tau))$$

that is, $V_{\tau}(\Lambda, F)$ does not increase after each iteration. Thus, this descent property guarantees the convergence of the algorithm to a local minimum of the optimization problem given by equations (3) and (5). To ensure that the QPC estimator $\hat{\theta}(\tau)$ is not a local optimum, one could use different random starting points and keep the solution that delivers the smaller value of $V_{\tau}(\hat{\Lambda}(\tau), \hat{F}(\tau))$. Alternatively, we can use more sophisticated methodologies, such as the one based on a deterministic annealing framework proposed by [Zhou and Lange \(2010\)](#), for instance⁶.

On the other hand, the computational complexity of the algorithm, which can be understood as the total iterations or total time required by an iterative procedure to achieve termination in the worst-case scenario, can be found as follows. According to Definition 3, QPC estimation involves running a series of quantile regressions, computed using interior-point methods⁷. [Portnoy and Koenker \(1997\)](#) establish that for a sample of size n and p estimated parameters, the complexity of the interior-point algorithm is $O(n^{5/2}p^3)$. In this case, the first step of the procedure computes $K(\tau)$ quantile factors from cross-sections of size N using quantile regressions T times. Thus, this first step has a complexity order of $O(N^{5/2}TK(\tau)^3)$

⁶Deterministic annealing is a statistical technique for approximating the global minimum of a given function. It consists of two iterative steps. In the first one, the objective function is flattened using a tuning parameter to eliminate (most of) its local minima. Then, optimization is performed using the transformed objective function. In the second step, the flattened objective function is warped by reverting the value of the tuning parameter with a single or handful of local minima with the hope that one of them corresponds to the global optimum.

⁷Interior-point methods, also known as barrier methods, are a particular class of algorithms designed to solve convex optimization problems that arose from the search for algorithms with better theoretical properties than the simplex method. One of its main characteristics is that they require all iterates to satisfy inequality constraints strictly. See [Nocedal and Wright \(2006\)](#), pp. 563-597 for further details about this class of algorithms.

per iteration. Analogously, the second step is of order $O((N - K(\tau))T^{5/2}K(\tau)^3)$ per iteration because it estimates $K(\tau)$ quantile factor loadings from time series of size T using quantile regressions $N - K(\tau)$ times.

Given the above results, the overall complexity of the QPC algorithm is $O(NT(K(\tau) \cdot \delta_{NT})^3)$ per iteration, where $\delta_{NT} \equiv \max\{\sqrt{N}, (1 - K(\tau)/N)^{1/3}\sqrt{T}\}$. Relative to the EM algorithm, there is an unfavorable gap, because the complexity, in this case, is limited by $O(TNK(\tau))$ per iteration (Roweis, 1998). Note also that the overall complexity depends crucially on the rotation adopted to derive the QPC estimator. If one considers the traditional or the recursive rotation instead, then the complexity of the algorithm is $O(NTK(\tau)^3 \cdot \max\{N^{3/2}, T^{3/2}\})$ in the worst-case scenario. The difference becomes apparent once we analyze the nature of the observable variables. If T grows at a much faster rate than N , as could be the case of macroeconomic data, then the complexity is of order $O(NT^{5/2}K(\tau)^3)$ under the traditional and recursive rotations, but of smaller order $O((N - K(\tau))T^{5/2}K(\tau)^3)$ in the default case. Therefore, improvements in complexity under the errors-in-variables rotation are considerable when the number of quantile factors to be estimated is large. On the contrary, if N grows faster than T , as could be the case of microeconomic data, then the complexity of the algorithm is $O(N^{5/2}TK(\tau)^3)$ under any rotation, i.e., there is no gain in selecting one rotation over another.

To get an upper bound of the number of required iterations for convergence of the algorithm, suppose that after each iteration the distance between the QPC estimator $\hat{\boldsymbol{\theta}}^{(j)}(\tau)$ and the true value of the parameters $\boldsymbol{\theta}^0(\tau)$ is reduced by a proportion $0 < \Delta_{NT} < 1$, that is $\|\hat{\boldsymbol{\theta}}^{(j)}(\tau) - \boldsymbol{\theta}^0(\tau)\| = \Delta_{NT} \cdot \|\hat{\boldsymbol{\theta}}^{(j-1)}(\tau) - \boldsymbol{\theta}^0(\tau)\|$. Therefore, after I iterations, an initial distance $\|\hat{\boldsymbol{\theta}}^{(0)}(\tau) - \boldsymbol{\theta}^0(\tau)\|$ is reduced by $(\Delta_{NT})^I \cdot \|\hat{\boldsymbol{\theta}}^{(0)}(\tau) - \boldsymbol{\theta}^0(\tau)\|$. From Definition 3 and the triangle inequality, we note that the iterative procedure stops when $(\Delta_{NT})^I \cdot \|\hat{\boldsymbol{\theta}}^{(0)}(\tau) - \boldsymbol{\theta}^0(\tau)\| < \epsilon$. Thus, the number of iterations I for termination of the algorithm is

$$I < \frac{\log \epsilon - \log \|\hat{\boldsymbol{\theta}}^{(0)}(\tau) - \boldsymbol{\theta}^0(\tau)\|}{\log \Delta_{NT}}$$

The worst-case scenario literature applied to this case suggests $\Delta_{NT} < 1 - (NT)^{-1/2}$ and assumes that the distance $\|\hat{\boldsymbol{\theta}}^{(0)}(\tau) - \boldsymbol{\theta}^0(\tau)\|$ is independent of N and T (see Cormen *et al.*, 2001, pp. 62-84). Therefore, the number of required iterations I is of order $O(\sqrt{NT} \log \epsilon)$, and the complexity of the algorithm as a whole is $O((\sqrt{NT}K(\tau) \cdot \delta_{NT})^3 \log \epsilon)$.

2.3.2 Performance of the QPC Estimator

In this section, I explore the finite sample properties of the QPC estimator with Monte Carlo simulations. In particular, I consider three data generating processes (DGP) based on the examples described at the beginning of this section:

- **DGP 1:** $y_{it} = \alpha_i^0 \beta_t^0 + v_{it}^0$, where α_i^0 , β_t^0 , and v_{it}^0 are independent draws from $\mathcal{N}(0, 1)$.
- **DGP 2:** $y_{it} = \alpha_i^0 \beta_t^0 + \gamma_t^0 v_{it}^0$, where $\gamma_t^0 = e^{x_t^0}$; and α_i^0 , β_t^0 , x_t^0 , and v_{it}^0 are independent draws from $\mathcal{N}(0, 1)$.
- **DGP 3:** $y_{it} = \alpha_i^0 \beta_t^0 e^{z_{it}^0}$, where $\alpha_i^0 = e^{z_i^0}$, $\beta_t^0 = e^{w_t^0}$; and z_i^0 , w_t^0 , and v_{it}^0 are independent draws from $\mathcal{N}(0, 1)$.

In all DGP's, I consider three different cross-section dimensions, $N = \{10, 50, 100\}$, and four different time series dimensions, $T = \{50, 100, 200, 1000\}$. Each configuration was simulated 1000 times, and in each simulation, I compute the PC estimators $\tilde{\theta} = [\tilde{\Lambda}', \tilde{F}']'$ and the QPC estimators $\hat{\theta}(\tau) = [\hat{\Lambda}(\tau)', \hat{F}(\tau)']'$ for $\tau = \{0.25, 0.50, 0.75\}$. Note that under DGP 1 and DGP 2, $K(\tau) = 1$ for $\tau = 0.5$ and $K(\tau) = 2$ for $\tau \neq 0.5$, whereas $K(\tau) = 1$ for all $\tau \in (0, 1)$ in DGP 3. I also kept track of the correlation between the estimated and true factors and factor loadings to measure the estimation precision of PC and QPC.

Table 1 shows the average correlation associated with the simulations under the standard factor model setup (DGP 1). Several findings are worth highlighting from it. Firstly, both PC and QPC do a remarkable job of estimating the simulated factors and factor loadings. In fact, in all cases considered, the average correlation between the simulated and estimated parameters is above 0.85. In the particular case of the QPC estimator with $\tau = 0.5$, it spikes to over 0.99 when both N and T are greater or equal than 100, which suggests that the estimated parameters can be effectively treated as the true ones. Secondly, it is clear from Panels A and B of Table 1 that the estimation precision of the factor and factor loading tends to improve as N and T becomes larger, respectively. This result is expected because the first and second steps of the QPC algorithm uses cross-section and time-series data to estimate the quantile factors and the quantile loadings, respectively. Thus, as N and T become larger, the estimates get closer to the corresponding true parameters for each i and t . Thirdly, when $\tau \neq 0.5$, the QPC method fails to precisely estimate α_i^0 and β_t^0 . This outcome, which worsens as both N and T becomes larger, has his root in the misspecification of the quantile factor model. Recall from Example 1 that DGP 1 can be rewritten in the form of equation (1) by setting $\lambda_i^0(u_{it}) = [\alpha_i^0, 1]'$ and $f_t^0(u_{it}) = [\beta_t^0, \Phi^{-1}(u_{it})]$, where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal distribution function. This means that, when $\tau = 0.5$, the model imposes $K(0.5) = 1$ factor, whereas when $\tau = 0.25$ or $\tau = 0.75$, it imposes $K(0.25) = K(0.75) = 2$

factors. Thus, the results showed in Table 1 support the good properties of the QPC estimator in identified quantile factor analysis models over all quantiles $\tau \in (0, 1)$.

Figure 1 plots the simulated and estimated factors obtained using the PC and QPC methodologies for the particular case where $N = 10$ and $T = 200$, as a way to quantify the estimation bias graphically. All panels of this figure show that the QPC estimators with $\tau = \{0.25, 0.50, 0.75\}$ and the PC estimator are, in effect, unbiased, which is consistent with the average correlation measure discussed previously. Note, however, that the PC estimator is more efficient than the QPC estimators. One reason that could explain this result is that the QPC estimator does not have a closed-form solution, as is the case of the PC methodology via the eigendecomposition of the $Y'Y$ matrix. Thus, the solution of the optimization problem has to be calculated numerically. Panels (b) and (d) of this figure also shows that the QPC estimator when $\tau = 0.25$ or $\tau = 0.75$ is slightly less precise than the same estimator computed at the median. This result could be because at these quantiles, the data is sparser than at $\tau = 0.5$. Consequently, both quantile factors and quantile loadings are characterized less accurately.

Table 2 shows the average correlation between the simulated and estimated factors under the second DGP. From Panel A of this table, we can see that both estimation methodologies well capture the first factor β_t^0 . In fact, in most cases, the average correlation is well above 0.90, and, as mentioned previously, it improves as N becomes larger. Another result to point out from Panel A is that the estimates of quantiles at the tails of the distribution are, on average, less accurate than those located at the center of the distribution. Note that under this DGP, $E[y_{it} | \alpha_i^0, \beta_t^0] = \mathcal{Q}_{y_{it}}(0.50 | \alpha_i^0, \beta_t^0) = \alpha_i^0 \beta_t^0$, i.e., the center of the joint conditional distribution of y_{it} is determined by one factor only. In this manner, the QPC estimator with $\tau = 0.50$ and the PC estimator does a remarkably good job in computing an estimator of β_t^0 . In particular, the average correlation is over 0.9, even for small values of N , and both estimators can be effectively treated as the true ones when $N \geq 50$. This result is also observed in the case of the QPC estimator with $\tau = 0.25$ or $\tau = 0.75$, although the average correlations are somewhat smaller. On the other hand, for any $\tau \neq 0.5$, we have that $\mathcal{Q}_{y_{it}}(\tau) = \alpha_i^0 \beta_t^0 + \gamma_t^0 \Phi^{-1}(\tau)$, which means that these quantiles contain additional information that is exploited by the QPC methodology to compute an estimator for γ_t^0 . Panel B of this table indicates that, in general, the QPC methodology with $\tau = 0.25$ or $\tau = 0.75$ delivers accurate estimators of γ_t^0 when $N = 50$ or larger. Figures 2 and 3 corroborate these results concerning the estimators of the first and second factors, respectively.

Finally, Table 3 depicts the average correlation between the simulated and estimated factor and factor loading under a nonlinear factor model (DGP 3). Here we note that, for any value of τ , the QPC methodology has a better performance relative to PC. For instance,

when $N = 100$ and $T = 1000$, the average correlation between the simulated and estimated factors via PC is about 0.85, whereas this correlation is around 0.95 for the QPC estimates (Panel A). A similar result is found in the case of the quantile factor loadings estimates. From Panel B, we see that when $N = 100$, the mean average correlation is around 0.90 under PC, whereas it is close to 0.97 under QPC. So, the estimated factor loadings can be effectively treated as the true ones for all values of τ considered. In terms of efficiency, Figure 4 points to the QPC as the clear winner. Panels (b) through (d) show that, for all values of τ , the QPC estimates are very close to their population counterparts. On the contrary, the PC estimator displays a considerable variability around the simulated series, predicting in some cases, negative values of β_t^0 . This result is at odds with the non-negativity assumption on this factor.

3 Asymptotic Theory

I start this section by presenting a set of assumptions required to establish the uniform consistency of the QPC estimator of the quantile factors, quantile factor loadings, and quantile common component of the QFA model (2). Next, I provide some additional assumptions that are used to derive an asymptotic theory for these estimators. Then, I discuss some aspects related to the computation of consistent estimators for the asymptotic variance-covariance matrices. Finally, I carry out Monte Carlo simulations intending to evaluate the adequacy of the asymptotic results for approximating the finite sample distributions of the estimators.

3.1 Consistency of the QPC Estimator

At this point, it is convenient to introduce some additional notation. For all i, t , and $\tau \in (0, 1)$, let $\varepsilon_{it}^0(\tau) \equiv y_{it} - \mathcal{Q}_{y_{it}}(\tau)$ be the quantile factor residual of model (2). I now make the following assumptions.

Assumption 3 (Uniform Consistency).

1. For a given i and $\tau \in (0, 1)$, $\psi_\tau(\varepsilon_{it}^0(\tau)) = \mathbb{1}\{\varepsilon_{it}^0(\tau) < 0\} - \tau$ is a martingale difference sequence with respect to $\lambda_i^0(\tau)$ and $f_t^0(\tau)$. Also, for all $i \neq j$, $\varepsilon_{it}^0(\tau)$ and $\varepsilon_{jt}^0(\tau)$ are independent.
2. For all i and t , the conditional densities $g_{it}(\cdot | \theta_{it}^0(\tau))$ satisfy Assumption 2.1 and

$$0 < L_g \leq g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) \leq U_g < \infty$$

3. For any $\epsilon > 0$, there exists $\sigma(\epsilon) > 0$ such that

$$\left| g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) + c \middle| \theta_{it}^0(\tau) \right) - g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) \right| < \epsilon$$

for all $|c| < \sigma(\epsilon)$, and all i and t .

4. *Quantile factors.* For all $\tau \in (0, 1)$,

(a) $T^{-1} \sum_{t=1}^T f_t^0(\tau) f_t^0(\tau)' \xrightarrow{p} \Sigma_F^0(\tau)$ as $T \rightarrow \infty$ for some $K(\tau) \times K(\tau)$ positive definite, non-random matrix.

(b) $\sup_{1 \leq t \leq T} \|f_t^0(\tau)\| = o_p(T^{1/2})$.

5. *Quantile factor loadings.* For all $\tau \in (0, 1)$,

(a) $N^{-1} \sum_{i=1}^N \lambda_i^0(\tau) \lambda_i^0(\tau)' \xrightarrow{p} \Sigma_\Lambda^0(\tau)$ as $N \rightarrow \infty$ for some $K(\tau) \times K(\tau)$ positive definite, non-random matrix.

(b) $\sup_{1 \leq i \leq N} \|\lambda_i^0(\tau)\| = o_p(N^{1/2})$.

Assumption 3.1 is familiar in the quantile regression literature (see [Koenker and Bassett, 1978](#); [Koenker, 2005](#)). For a given cross-section i and quantile indicator τ , it restricts the dependence of the dichotomous random variable $\psi_\tau(\varepsilon_{it}^0(\tau))$ with past values. It also excludes cross-sectional dependence of the quantile factor residual $\varepsilon_{it}^0(\tau)$ but allows for heteroskedasticity and dynamic models. Assumptions 3.2 and 3.3 are similar to those considered in [Oka and Qu \(2011\)](#). They are local because they impose restrictions over the conditional densities evaluated at the quantile of interest instead of over the whole conditional distribution of the observable variable y_{it} . In particular, Assumption 3.2 requires that the conditional densities evaluated at the τ -th quantile are uniformly bounded away from zero and infinity for all i and t . This requirement implies that $g_{it}(\cdot | \theta_{it}^0(\tau))$ can be unbounded at any quantile different from τ . Assumption 3.3, on its part, imposes smoothness of $g_{it}(\cdot | \theta_{it}^0(\tau))$ in a neighborhood of the τ -th quantile of y_{it} . Assumptions 3.4 and 3.5 impose some structure on the quantile factors and quantile factor loadings, respectively. The first part of these assumptions are standard in large dimensional factor models (see, for instance, [Bai and Ng, 2002](#); [Bai, 2003](#); [Bai and Ng, 2008](#); [Bai and Li, 2012](#), among others). They together imply the existence of $K(\tau)$ unobserved quantile factors, each having a non-trivial contribution to the τ -th conditional quantile function of y_{it} . The main difference with the traditional literature, however, is that in this setup, the matrices $\Sigma_F^0(\tau)$ and $\Sigma_\Lambda^0(\tau)$ are also quantile-specific. Part (b) of Assumptions 3.4 and 3.5, which is familiar in the literature of M-estimators (see [Huber and Ronchetti, 2009](#),

pp. 126-130), is required to ensure the stochastic equicontinuity of the sequential empirical processes derived from the estimation of the quantile factor residuals $\varepsilon_{it}^0(\tau)$.

I illustrate the implications of Assumptions 3.1 to 3.3 by considering the examples described in Section 2.1. Note that I do not include Assumptions 3.4 and 3.5 in this analysis because they imply restrictions over the quantile factors and quantile factor loadings, respectively, that are independent of the model structure.

Example 1 (Standard Factor Model). Because v_{it}^0 is *iid*, Assumption 3.1 is satisfied due to the independence of v_{it}^0 . Moreover, because v_{it}^0 has cumulative distribution function $G_v(\cdot)$ and density $g_v(\cdot)$, then $g_{it}(G_{it}^{-1}(\tau) | \theta_{it}^0(\tau)) = g_v(G_v^{-1}(\tau))$, where $G_v^{-1}(\tau)$ denotes the τ -th quantile function of v_{it}^0 . Thus, Assumption 3.2 is satisfied if $G_v(\cdot)$ is absolutely continuous with continuous density $g_v(\cdot)$ fulfilling $0 < L_g \leq g_v(G_v^{-1}(\tau)) \leq U_g < \infty$. Assumption 3.3 is satisfied if, additionally, $g_v(\cdot)$ is continuous in an open ball around the τ -th quantile of y_{it} .

Example 2 (Location-Scale Factor Model). Similar to the previous case, Assumption 3.1 is satisfied because of the independence of v_{it}^0 . However, in this case, $g_{it}(G_{it}^{-1}(\tau) | \theta_{it}^0(\tau)) = g_v(G_v^{-1}(\tau)) / \gamma_t^0$, $\gamma_t^0 > 0$ for all t , implying that Assumption 3.2 is met if $G_v(\cdot)$ is absolutely continuous, the density is continuous and satisfies $\delta_v < g_v(G_v^{-1}(\tau)) < \infty$, and $\gamma_t^0 < \infty$ for all t for some arbitrary strictly positive constant δ_v . If $g_v(\cdot)$ is also continuous around the quantile of interest, then Assumption 3.3 is satisfied.

Example 3 (Nonlinear Factor Model). Again, Assumption 3.1 is satisfied because of the independence of v_{it}^0 . Because $g_{it}(G_{it}^{-1}(\tau) | \theta_{it}^0(\tau)) = g_v(G_v^{-1}(\tau)) / y_{it}$ and $y_{it} > 0$ for all i and t , Assumptions 3.2 and 3.3 are satisfied by similar arguments of Example 2. However, we also require $y_{it} < \infty$ for all i and t in this case.

Next, I establish the second main result of this paper, namely the uniform consistency of the QPC estimator.

Theorem 2 (Uniform Consistency of the QPC Estimator). *Suppose that Assumption 3 holds. Let $\hat{\boldsymbol{\theta}}(\tau)$ be the QPC estimator of $\boldsymbol{\theta}^0(\tau) = [\Lambda^0(\tau)', F^0(\tau)']'$ obtained using a panel consisting of $i = 1, \dots, N$ cross-sections and $t = 1, \dots, T$ periods. Then, as $N, T \rightarrow \infty$, for every $\tau \in (0, 1)$*

1. *Uniformly in i ,*

$$\sqrt{T} \left\| \hat{\lambda}_i(\tau) - \lambda_i^0(\tau) \right\| = O_p(1)$$

2. *Uniformly in t .*

$$\sqrt{N} \left\| \hat{f}_t(\tau) - f_t^0(\tau) \right\| = O_p(1)$$

The proof of Theorem 2 consists of two parts. In the first part, I show the uniform consistency of the QPC estimator of the common component $\hat{c}_{it}(\tau) = \hat{\lambda}_i(\tau)' \hat{f}_t(\tau)$, which is a strategy similar to the one used in [Chen *et al.* \(2014\)](#) in the context of nonlinear panel data models with interactive effects. The proof relies on the convexity of the objective function $V_\tau(\Lambda, F)$ for the quantile common component of y_{it} , and the compactness of the set $\Theta \subset \mathbb{R}$, which contains the difference $\hat{\phi}_{it}(\tau) = \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) - \lambda_i^0(\tau)' f_t^0(\tau)$. In particular, if consistency does not hold, then the centered objective function evaluated at a fixed $\hat{\phi}_{it}(\tau)$ is strictly positive with probability close to 1, thus implying that $\hat{c}_{it}(\tau)$ cannot be its minimizer. So, this part of the proof concludes that

$$\min \left\{ \sqrt{N}, \sqrt{T} \right\} \cdot |\hat{c}_{it}(\tau) - c_{it}^0(\tau)| = O_p(1) \quad (8)$$

for fixed i and t . Then, by a standard Uniform Weak Law of Large Numbers argument, the above result also holds uniformly in i and t . In this part, I only require Assumptions 3.1 to 3.3 since they impose restrictions over the random variable $\psi_\tau(\varepsilon_{it}^0(\tau))$, which in turn is a function of $c_{it}^0(\tau)$, and over the conditional densities $g_{it}(\cdot | \theta_{it}^0(\tau))$ evaluated at the quantile of interest, respectively.

In the second part of the proof, I show the uniform consistency of both $\hat{f}_t(\tau)$ and $\hat{\lambda}_i(\tau)$ starting from equation (8). To achieve this goal, I employ a first-order approximation of $\hat{\phi}_{it}(\tau)$ in terms of the differences $\hat{\phi}_i^\lambda(\tau) \equiv \hat{\lambda}_i(\tau) - \lambda_i^0(\tau)$ and $\hat{\phi}_t^f(\tau) \equiv \hat{f}_t(\tau) - f_t^0(\tau)$, together with the strong factor structure implied by the second part of Assumptions 3.4 and 3.5. Then, the argument exploits the fact that, as $N, T \rightarrow \infty$, equation (8) is $o_p(1)$. Therefore, both $\sqrt{T} \left\| \hat{\lambda}_i(\tau) - \lambda_i^0(\tau) \right\|$ and $\sqrt{N} \left\| \hat{f}_t(\tau) - f_t^0(\tau) \right\|$ have to be $O_p(1)$ jointly to meet this condition.

3.2 Asymptotic Distribution of the QPC Estimator

In this section, I impose the following additional assumptions to derive the limiting distribution of the QPC estimators $\hat{\lambda}_i(\tau)$, $\hat{f}_t(\tau)$, and $\hat{c}_{it}(\tau)$.

Assumption 4. *Let $g_{it}(\cdot | \theta_{it}^0(\tau))$ and $G_{it}(\cdot | \theta_{it}^0(\tau))$ be the conditional density and the conditional cumulative distribution function, respectively, of the observable random variable y_{it} in the QFA model (2). Then, for any $\tau \in (0, 1)$,*

1. $T^{-1} \sum_{t=1}^T g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) f_t^0(\tau) f_t^0(\tau)' \xrightarrow{p} H_i^0(\tau)$ uniformly in i as $T \rightarrow \infty$ for some $K(\tau) \times K(\tau)$ positive definite, non-random matrix.

2. $N^{-1} \sum_{i=1}^N g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) \lambda_i^0(\tau) \lambda_i^0(\tau)' \xrightarrow{P} J_t^0(\tau)$ uniformly in t as $N \rightarrow \infty$ for some $K(\tau) \times K(\tau)$ positive definite, non-random matrix.

Assumption 4 above imposes some restrictions on the heteroskedasticity of the model. To see this, note that both $H_i^0(\tau)$ and $J_t^0(\tau)$ are the limits of a matrix that is a function of the conditional density $g_{it}(\cdot | \theta_{it}^0(\tau))$ evaluated at the quantile of interest; the latter being the reciprocal of an object that is known in the related literature as the quantile density function⁸. In this sense, if there is a high number of observations near the quantile of interest, i.e., the data is locally dense, then $G_{it}^{-1}(\tau | \theta_{it}^0(\tau))$ can be characterized accurately. On the contrary, if the data around the quantile of interest is sparse, then the characterization of this element is less precise. Moreover, the assumptions on $g_{it}(\cdot | \theta_{it}^0(\tau))$ allow the local density of the data to vary across time-series and cross-sections, impacting the heteroskedasticity of the model.

Next, I establish the third main result of this paper, namely the asymptotic distribution of the QPC estimators of the quantile factors and quantile factor loadings.

Theorem 3 (Asymptotic Distribution of $\hat{\lambda}_i(\tau)$ and $\hat{f}_t(\tau)$). *Suppose that Assumptions 3 and 4 hold. Let $\hat{\lambda}_i(\tau)$ and $\hat{f}_t(\tau)$ be the QPC estimator of the quantile factor loadings $\lambda_i^0(\tau)$ and quantile factors $f_t^0(\tau)$, respectively. Then, as $N, T \rightarrow \infty$, for any $\tau \in (0, 1)$,*

1. Uniformly in i , if $\sqrt{T}/N \rightarrow 0$

$$\sqrt{T} \left(\hat{\lambda}_i(\tau) - \lambda_i^0(\tau) \right) \xrightarrow{d} \mathcal{N} \left(0, \tau(1 - \tau) H_i^0(\tau)^{-1} \Sigma_F^0(\tau) H_i^0(\tau)^{-1} \right)$$

2. Uniformly in t , if $\sqrt{N}/T \rightarrow 0$

$$\sqrt{N} \left(\hat{f}_t(\tau) - f_t^0(\tau) \right) \xrightarrow{d} \mathcal{N} \left(0, \tau(1 - \tau) J_t^0(\tau)^{-1} \Sigma_\Lambda^0(\tau) J_t^0(\tau)^{-1} \right)$$

where $T^{-1} \sum_{t=1}^T f_t^0(\tau) f_t^0(\tau)' \xrightarrow{P} \Sigma_F^0(\tau)$ and $N^{-1} \sum_{i=1}^N \lambda_i^0(\tau) \lambda_i^0(\tau)' \xrightarrow{P} \Sigma_\Lambda^0(\tau)$ as $T \rightarrow \infty$ and $N \rightarrow \infty$, respectively.

The proof of Theorem 3 relies on the properties of the sub-gradient of the objective function $V_\tau(\Lambda, F)$ defined in (5), evaluated at the QPC estimator $\hat{\boldsymbol{\theta}}(\tau) = [\hat{\Lambda}(\tau)', \hat{F}(\tau)']'$. Recall that the quantile loss $\rho_\tau(\cdot)$ is a piecewise linear and continuous function. Thus, $V_\tau(\Lambda, F)$ is everywhere differentiable except at the points at which the objective function is equal to zero. Therefore, the optimality conditions of the problem are defined in terms of the sub-gradient, rather than the gradient of $V_\tau(\Lambda, F)$. In this manner, since the objective

⁸Alternatively, [Tukey \(1965\)](#) refers to this object as the sparsity function.

function has two arguments, the respective sub-gradient vectors for any i , t , and $\tau \in (0, 1)$ are the following

$$R_{i,\tau}^\Lambda(\Lambda, F) = (NT)^{-1/2} \sum_{t=1}^T \psi_\tau(y_{it} - \lambda'_i f_t) f_t \quad (9)$$

$$R_{i,\tau}^F(\Lambda, F) = (NT)^{-1/2} \sum_{i=1}^N \psi_\tau(y_{it} - \lambda'_i f_t) \lambda_i \quad (10)$$

The first part of the proof provides a uniform asymptotic approximation for the sub-gradients evaluated at the QPC estimators using the uniform consistency of $\hat{\boldsymbol{\theta}}(\tau)$ and an argument similar to the one employed by [Qu \(2008\)](#) in the derivation of the limiting distribution of a test for structural change in the context of quantile regressions. Specifically, the first part of the proof demonstrates that $R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), \hat{F}(\tau))$ and $R_{i,\tau}^F(\hat{\Lambda}(\tau), \hat{F}(\tau))$ can be approximated by $R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau))$ and $R_{i,\tau}^F(\Lambda^0(\tau), F^0(\tau))$, correspondingly, an additional term capturing the difference between the QPC estimators and their actual counterparts, and an $o_p(\cdot)$ term. Based on this result, the second part shows that the previous approximation admits a Bahadur representation, i.e., both $\sqrt{T}(\hat{\lambda}_i(\tau) - \lambda_i^0(\tau))$ and $\sqrt{N}(\hat{f}_t(\tau) - f_t^0(\tau))$ can be expressed as a normalized sum of martingale difference sequences plus a random variable that converges in probability to zero⁹. Using this approximation, I invoke a standard uniform central limit theorem to show the asymptotic normality of the QPC estimator of both the quantile factors and the quantile factor loadings.

The results of Theorem 3 allow both N and T to become large simultaneously. However, additional restrictions on the relationship between these two dimensions need to be imposed. In particular, in the case of $\hat{\lambda}_i(\tau)$, the time dimension T has to grow faster than the cross-section dimension N for the result to hold. If on the contrary $\sqrt{T}/N \rightarrow \delta > 0$, then the limiting distribution will not be centered at zero because of an additional $O_p(1)$ term in the asymptotic approximation of the sub-gradient (9) when evaluated at the QPC estimator. Further, the convergence rate implied by the theorem is $\min\{N, \sqrt{T}\}$. To see this, note that the asymptotic distribution of $\hat{\lambda}_i(\tau)$ can be expressed as follows

$$\left(\frac{r_{NT}^2}{T} \tau(1-\tau) H_i^0(\tau)^{-1} \Sigma_F^0(\tau) H_i^0(\tau)^{-1} \right)^{-1/2} r_{NT} (\hat{\lambda}_i(\tau) - \lambda_i^0(\tau)) \xrightarrow{d} \mathcal{N}(0, I_{K(\tau)})$$

⁹For more details about the Bahadur representation of uniform quantile processes, see [Bahadur \(1966\)](#) and [Kiefer \(1967\)](#). For more information about the use of this representation within a quantile regression context, see [Koenker and Portnoy \(1987\)](#), [Portnoy and Koenker \(1989\)](#), [Gutenbrunner and Jureckova \(1992\)](#), [Koenker and Machado \(1999\)](#), and [Koenker and Xiao \(2002\)](#), among others.

where $r_{NT} = \min\{N, \sqrt{T}\}$. Since $\sqrt{T}/N \rightarrow 0$, the denominator of the previous expression is bounded above and below, and the convergence rate of the estimator is $\min\{N, \sqrt{T}\}$ as mentioned before, which captures the fact that to compute $\hat{\lambda}_i(\tau)$ we also need to estimate the quantile factors because the econometrician does not observe $F^0(\tau) = [f_1^0(\tau), \dots, f_T^0(\tau)]'$. If $F^0(\tau)$ is observed, then $\hat{\lambda}_i(\tau)$ is obtained simply by running quantile regressions across periods for each cross-section i , and the convergence rate is, therefore, the usual \sqrt{T} (see [Koenker, 2005](#)). Analogously, in the case of $\hat{f}_t(\tau)$, the cross-section dimension N has to grow faster than the time dimension T to avoid an asymptotic bias due to the appearance of an extra $O_p(1)$ term when approximating $R_{t,\tau}^F(\hat{\Lambda}(\tau), \hat{F}(\tau))$. Its limiting distribution can be alternatively expressed as

$$\left(\frac{\tilde{r}_{NT}^2}{N} \tau(1-\tau) J_t^0(\tau)^{-1} \Sigma_{\Lambda}^0(\tau) J_t^0(\tau)^{-1} \right)^{-1/2} \tilde{r}_{NT} (\hat{f}_t(\tau) - f_t^0(\tau)) \xrightarrow{d} \mathcal{N}(0, I_{K(\tau)})$$

where $\tilde{r}_{NT} = \min\{\sqrt{N}, T\}$. The denominator of this expression is bounded both above and below because $\sqrt{N}/T \rightarrow 0$. This result implies that the convergence rate, in this case, is $\min\{\sqrt{N}, T\}$ given that the quantile factor loadings are not observable and need to be estimated. If $\Lambda^0(\tau) = [\lambda_1^0(\tau), \dots, \lambda_N^0(\tau)]'$ is observed, then $\hat{f}_t(\tau)$ is computed from longitudinal quantile regressions for each period t and the convergence rate attained is \sqrt{N} .

Another feature of Theorem 3 is the fact that the covariance matrix of the limiting distribution of $\hat{\lambda}_i(\tau)$ and $\hat{f}_t(\tau)$ depends on $\Sigma_F^0(\tau)$ and $\Sigma_{\Lambda}^0(\tau)$, respectively. This feature is an expected result since, in the first case, the QPC algorithm computes the estimator of the quantile loadings from the observable variables y_{it} , treating $\hat{F}(\tau)$ as the actual quantile factors. Hence, the asymptotic variance of $\hat{\lambda}_i(\tau)$ reflects this feature by incorporating a term that captures the contribution of $F^0(\tau)$ to the τ -th quantile of y_{it} . A symmetric argument applies for the case of $\hat{f}_t(\tau)$ and $\Sigma_{\Lambda}^0(\tau)$. Lastly, the matrices $H_i^0(\tau)$ and $J_t^0(\tau)$ incorporate the number of observations close to the quantile of interest into the asymptotic variances in the following manner. If the density of observations around $\tau \in (0, 1)$ is high (low), the QPC estimator is computed more (less) accurately, and we would expect therefore a lower (higher) asymptotic variance.

Below, I establish the asymptotic distribution of the quantile common component, which corresponds to the last main result of this paper.

Theorem 4 (Asymptotic Distribution of $\hat{c}_{it}(\tau)$). *Suppose that Assumptions 3 and 4 hold. Let $\hat{c}_{it}(\tau) = \hat{\lambda}_i(\tau)' \hat{f}_t(\tau)$ be the QPC estimator of the quantile common component $c_{it}^0(\tau)$. Then, as $N, T \rightarrow \infty$, for any $\tau \in (0, 1)$,*

$$\left(\frac{U_{it}^0(\tau)}{T} + \frac{W_{it}^0(\tau)}{N} \right)^{-1/2} (\hat{c}_{it}(\tau) - c_{it}^0(\tau)) \xrightarrow{d} \mathcal{N}(0, 1)$$

uniformly in i and t , where

$$U_{it}^0(\tau) = \tau(1 - \tau)f_t^0(\tau)'H_i^0(\tau)^{-1}\Sigma_F^0(\tau)H_i^0(\tau)^{-1}f_t^0(\tau)$$

and

$$W_{it}^0(\tau) = \tau(1 - \tau)\lambda_i^0(\tau)'J_t^0(\tau)^{-1}\Sigma_\Lambda^0(\tau)J_t^0(\tau)^{-1}\lambda_i^0(\tau)$$

The proof of the theorem above uses the results of Theorem 3 by recalling that $c_{it}^0(\tau) = \lambda_i^0(\tau)'f_t^0(\tau)$, for all i , t , and $\tau \in (0, 1)$. Specifically, the argument is based on a uniform asymptotic approximation intended to express the difference $\hat{c}_{it}(\tau) - c_{it}^0(\tau)$ as the sum of two asymptotically independent random variables. Then, using this approximation, the final result follows straightforwardly by Slutsky's theorem and a Uniform Weak Law of Large Numbers.

One important implication of the Theorem 4 is that, as opposed to the previous theorem, it does not impose any restriction on the relationship between N and T , i.e., the ratio \sqrt{T}/N or \sqrt{N}/T can attain any limit. Further, the asymptotic distribution of $\hat{c}_{it}(\tau)$ can be rewritten as follows

$$\left(\frac{\bar{r}_{NT}^2}{T}U_{it}^0(\tau) + \frac{\bar{r}_{NT}^2}{N}W_{it}^0(\tau) \right)^{-1/2} \bar{r}_{NT} (\hat{c}_{it}(\tau) - c_{it}^0(\tau)) \xrightarrow{d} \mathcal{N}(0, 1)$$

where $\bar{r}_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$. Regardless of the relationship between N and T , the denominator of this expression is bounded both above and below. Hence, the convergence rate, in this case, is $\min\{\sqrt{N}, \sqrt{T}\}$, which means that when the true quantile factors $F^0(\tau)$ are observed, then the estimation of the quantile common component is equivalent to the estimation of $\hat{\lambda}_i(\tau)$ via quantile regressions across periods. The convergence rate is, therefore, \sqrt{T} . Conversely, if $\Lambda^0(\tau)$ is observed, then the computation of $\hat{c}_{it}(\tau)$ is equivalent to the computation of $\hat{f}_t(\tau)$ using longitudinal quantile regressions, and so the convergence rate is \sqrt{N} .

Lastly, note that Theorem 4 has two special cases. First, if $N/T \rightarrow 0$, then for any $\tau \in (0, 1)$ the asymptotic distribution of $\hat{c}_{it}(\tau)$ is given by

$$\sqrt{N} (\hat{c}_{it}(\tau) - c_{it}^0(\tau)) \xrightarrow{d} \mathcal{N} \left(0, \tau(1 - \tau)\lambda_i^0(\tau)'J_t^0(\tau)^{-1}\Sigma_\Lambda^0(\tau)J_t^0(\tau)^{-1}\lambda_i^0(\tau) \right)$$

In the second case, if $T/N \rightarrow 0$, then the limiting distribution is as follows

$$\sqrt{T} (\hat{c}_{it}(\tau) - c_{it}^0(\tau)) \xrightarrow{d} \mathcal{N} \left(0, \tau(1 - \tau)f_t^0(\tau)'H_i^0(\tau)^{-1}\Sigma_F^0(\tau)H_i^0(\tau)^{-1}f_t^0(\tau) \right)$$

for all $\tau \in (0, 1)$.

3.2.1 Estimation of Asymptotic Variance-Covariance Matrices

In this section, I propose consistent estimators for the asymptotic variance-covariance matrices of the QPC estimators derived in Theorems 3 and 4. As noted before, the main ingredients are the matrices $\Sigma_F^0(\tau)$ and $\Sigma_\Lambda^0(\tau)$. They capture the contribution of the quantile factors and quantile loadings to the total variance of y_{it} , respectively, as well as the matrices $H_i^0(\tau)$ and $J_t^0(\tau)$ that are related to the data dispersion around $\tau \in (0, 1)$.

First, the matrices $\Sigma_F^0(\tau)$ and $\Sigma_\Lambda^0(\tau)$ depend only on $F^0(\tau)$ and $\Lambda^0(\tau)$, correspondingly. Thus, in light of the results implied by Theorem 2, a consistent estimator of these matrices is given by

$$\hat{\Sigma}_F(\tau) = \frac{1}{T} \sum_{t=1}^T \hat{f}_t(\tau) \hat{f}_t(\tau)' \quad (11)$$

and

$$\hat{\Sigma}_\Lambda(\tau) = \frac{1}{N} \sum_{i=1}^N \hat{\lambda}_i(\tau) \hat{\lambda}_i(\tau)' \quad (12)$$

for each $\tau \in (0, 1)$.

The matrices $H_i^0(\tau)$ and $J_t^0(\tau)$, on their part, need to be analyzed in more detail since they both depend on the conditional density $g_{it}(\cdot | \theta_{it}^0(\tau))$ evaluated at the quantile of interest, i.e., the reciprocal of the quantile density function. Because this quantity is unknown in practice, a vast literature that started with the work by [Siddiqui \(1960\)](#) has been devoted to its estimation. In particular, the approximation proposed by the author is based on the derivative of the inverse function, i.e., $dG_{it}^{-1}(\tau | \theta_{it}^0(\tau)) / d\tau = 1/g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \Big| \theta_{it}^0(\tau) \right)$, where the derivative, in turn, can be approximated numerically by the following expression

$$\frac{dG_{it}^{-1}(\tau | \theta_{it}^0(\tau))}{d\tau} \approx \frac{G_{it}^{-1}(\tau + h | \theta_{it}^0(\tau)) - G_{it}^{-1}(\tau - h | \theta_{it}^0(\tau))}{2h}$$

and $h > 0$ is a bandwidth parameter that usually depends on the sample size. For some small h , the previous approximation is preferable to the alternative $(G_{it}^{-1}(\tau + h | \theta_{it}^0(\tau)) - G_{it}^{-1}(\tau | \theta_{it}^0(\tau))) / h$ because in the former case the error is of order $O(h^4)$, whereas in the latter is of order $O(h^2)$ ¹⁰. These two elements imply that $g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \Big| \theta_{it}^0(\tau) \right)$ can be estimated from the so-called difference quotient

¹⁰By straightforward Taylor expansions we have that $\frac{G_{it}^{-1}(\tau+h|\theta_{it}^0(\tau))-G_{it}^{-1}(\tau|\theta_{it}^0(\tau))}{h} = \frac{dG_{it}^{-1}(\tau|\theta_{it}^0(\tau))}{d\tau} + \frac{h}{2} \frac{d^2G_{it}^{-1}(\tau|\theta_{it}^0(\tau))}{d\tau^2} + O(h^2)$ and $\frac{G_{it}^{-1}(\tau+h|\theta_{it}^0(\tau))-G_{it}^{-1}(\tau-h|\theta_{it}^0(\tau))}{2h} = \frac{dG_{it}^{-1}(\tau|\theta_{it}^0(\tau))}{d\tau} + \frac{h}{12} \frac{d^3G_{it}^{-1}(\tau|\theta_{it}^0(\tau))}{d\tau^3} + O(h^4)$.

$$\Delta_{it}(\tau, h) = \frac{2h}{\hat{G}_{it}^{-1}(\tau + h) - \hat{G}_{it}^{-1}(\tau - h)} \quad (13)$$

where $\hat{G}_{it}^{-1}(\cdot)$ is an estimator of $G_{it}^{-1}(\cdot | \theta_{it}^0(\tau))$.

Regarding the determination of the bandwidth parameter h , there exist several options in the related literature. One of them corresponds to the rule suggested by [Bofinger \(1975\)](#). The selected bandwidth is then an optimal one in that it minimizes the mean square error of the Gaussian density estimator

$$h_B = \left[\frac{1}{NT} \frac{4.5\phi^4(\Phi^{-1}(\tau))}{(2\Phi^{-1}(\tau)^2 + 1)^2} \right]^{1/5}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ is the standard Normal density and the standard Normal distribution function, respectively. An alternative is a bandwidth suggested by [Hall and Sheather \(1988\)](#), which is based on Edgeworth expansions of studentized quantiles with Gaussian density and can be computed by the following expression

$$h_{HS} = \left[\frac{1}{NT} \frac{1.5z_\alpha^2\phi^2(\Phi^{-1}(\tau))}{2\Phi^{-1}(\tau)^2 + 1} \right]^{1/3}$$

where z_α is such that $\Phi(z_\alpha) = 1 - \alpha/2$. Note that both h_B and h_{HS} tend to zero as $N, T \rightarrow \infty$.

Turning to the computation of $\hat{G}_{it}^{-1}(\cdot)$, the natural candidate for this task is the QPC estimator of the QFA model (2), i.e., $\hat{G}_{it}^{-1}(\tau) = \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) = \hat{c}_{it}(\tau)$. However, one potential pitfall of this approach is that there is no guarantee that the denominator of the difference quotient, $d_{it}(\tau, h) \equiv \hat{G}_{it}^{-1}(\tau + h) - \hat{G}_{it}^{-1}(\tau - h)$, is positive for all i and t , due to a potential quantile crossing problem. Fortunately, [Koenker and Machado \(1999\)](#) find that this problem occurs only infrequently and then only in extreme regions of the support of τ . Thus, in this context, we can expect that $d_{it}(\tau, h) \geq 0$ in most cases. If this difference is negative for some i and t , then the authors recommend to set $\Delta_{it}(\tau, h) = 0$. Finally, in the rare cases in which the difference is exactly equal to 0, then one can consider $d_{it}(\tau, h) - \varepsilon$ instead, where $\varepsilon > 0$ is a small tolerance parameter intended to avoid dividing by 0 when computing $\Delta_{it}(\tau, h)$. The previous observation implies that to implement this approach, we can consider a modified version of the difference quotient as follows

$$\tilde{\Delta}_{it}(\tau, h) = \max \left\{ 0, \frac{2h}{\hat{c}_{it}(\tau + h) - \hat{c}_{it}(\tau - h) - \varepsilon} \right\}, \quad \varepsilon > 0 \quad (14)$$

Thus, given the above, the matrices $H_i^0(\tau)$ and $J_t^0(\tau)$ can be consistently estimated by

$$\hat{H}_i(\tau) = \frac{1}{T} \sum_{t=1}^T \tilde{\Delta}_{it}(\tau, h) \hat{f}_t(\tau) \hat{f}_t(\tau)' \quad (15)$$

and

$$\hat{J}_t(\tau) = \frac{1}{N} \sum_{i=1}^N \tilde{\Delta}_{it}(\tau, h) \hat{\lambda}_i(\tau) \hat{\lambda}_i(\tau)' \quad (16)$$

for any i , t , and $\tau \in (0, 1)$, and a given choice of the bandwidth parameter h .

3.2.2 Monte Carlo Simulations

In this section, I explore the adequacy of the asymptotic distributions derived previously to approximate the finite sample distribution of the QPC estimators $\hat{\lambda}_i(\tau)$, $\hat{f}_t(\tau)$, and $\hat{c}_{it}(\tau)$ using Monte Carlo simulations. To this end, the DGP considered is a location-scale factor model given by $y_{it} = \alpha_i^0 \beta_t^0 + \beta_t^0 v_{it}^0$, where $\beta_t^0 = e^{z_t^0} > 0$ for all t . In particular, α_i^0 , z_t^0 , and v_{it}^0 are independent draws from a standard Normal distribution with distribution function $\Phi(\cdot)$ for all i , t . Moreover, I consider two alternative time dimensions $T = \{50, 100\}$ and four alternative cross-section dimensions $N = \{25, 50, 100, 1000\}$, which yields a total of eight different setups. Each DGP configuration was simulated 1,000 times and at each simulation I compute the QPC estimator of the quantile factors $f_t^0(\tau) = \beta_t^0$, the quantile factor loadings $\lambda_i^0(\tau) = \alpha_i^0 + \Phi^{-1}(\tau)$, and the quantile common components $c_{it}^0(\tau) = \lambda_i^0(\tau) f_t^0(\tau)$ for $\tau = \{0.25, 0.50, 0.75\}$.

Next, the QPC estimators were standardized in the following manner

$$\begin{aligned} \tilde{\lambda}_i(\tau) &= \left(\frac{\tau(1-\tau) \hat{H}_i(\tau)^{-1} \hat{\Sigma}_F(\tau) \hat{H}_i(\tau)^{-1}}{T} \right)^{-1/2} \left(\hat{\lambda}_i(\tau) - \lambda_i^0(\tau) \right) \\ \tilde{f}_t(\tau) &= \left(\frac{\tau(1-\tau) \hat{J}_t(\tau)^{-1} \hat{\Sigma}_\Delta(\tau) \hat{J}_t(\tau)^{-1}}{N} \right)^{-1/2} \left(\hat{f}_t(\tau) - f_t^0(\tau) \right) \\ \tilde{c}_{it}(\tau) &= \left(\frac{\hat{U}_{it}(\tau)}{T} + \frac{\hat{W}_{it}(\tau)}{N} \right)^{-1/2} \left(\hat{c}_{it}(\tau) - c_{it}^0(\tau) \right) \end{aligned}$$

for all i and t , where the elements of the variance-covariance matrices were computed according to the methodology described in the previous section. To preserve space, I report only the results using the Bofinger bandwidth to compute the matrices $\hat{H}_i(\tau)$ and $\hat{J}_t(\tau)$. The results under h_{HS} are similar and are, therefore, omitted. If the asymptotic theory provided

in Section 3.2 is a suitable approximation to the finite sample distribution of the QPC estimators, then the distribution of the standardized estimators $\tilde{\lambda}_i(\tau)$, $\tilde{f}_t(\tau)$ and $\tilde{c}_{it}(\tau)$ should be close to a standard Normal distribution for any i and t .

Tables 4 and 5 show the sample mean and the standard deviation, respectively, of the standardized QPC estimators. These statistics were computed across simulations and, to facilitate the exposition, I only report the results for $i = \lfloor N/2 \rfloor$ and $t = \lfloor T/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of x ¹¹. The results show that, in general, the means are close to zero, and the standard deviations are close to one, getting even closer to these values as N and T become large. For example, when $T = 50$ and $N = 25$, the sample mean and standard deviation of $\tilde{f}_t(0.25)$ is 0.158 and 1.380, respectively. When $T = 100$ and $N = 1000$, the results approach the desired values, in absolute terms, to 0.004 and 1.078, correspondingly. Also, note that the standard deviation of all standardized estimators computed at the median tends to be lower than the QPC estimates at either $\tau = 0.25$ or $\tau = 0.75$. This result could be attributable to the reciprocal of the quantile density function proxied by the difference quotient (14). In particular, under a normal distribution, the observations are generally dense around the median, whereas they are rather sparse at quantiles close to the tails. Thus, QPC estimators at $\tau = 0.5$ are more accurate relative to those computed at $\tau = 0.25$ or $\tau = 0.75$.

Figures 5 and 6 display the histogram of $\tilde{f}_t(0.25)$ and $\tilde{f}_t(0.50)$ for $T = 50$ and $T = 100$, respectively, contrasted with the standard normal density¹². Figures 7 and 8 show the same information for $\tilde{\lambda}_i(0.25)$ and $\tilde{\lambda}_i(0.50)$, as well as Figures 9 and 10 for $\tilde{c}_{it}(0.25)$ and $\tilde{c}_{it}(0.50)$. In all cases, to make the comparison with the Normal density, the histograms were scaled so that the total area of the columns add up to one. Overall, the figures suggest that the asymptotic distributions established in Theorems 3 and 4 provide a good approximation to the finite sample distributions of the QPC estimators. In the case of $\tilde{f}_t(\tau)$, for a given time series dimension T and irrespectively of the value of τ , the approximation tends to improve as N becomes larger in the sense that its histogram tends to stay more and more within the boundaries defined by the standard Normal density. A similar pattern can be noticed for the standardized quantile factor loading $\tilde{\lambda}_i(\tau)$ and the standardized quantile common component $\tilde{c}_{it}(\tau)$. In the former case, for a given value of the cross-section dimension N and independent of the value of τ , the asymptotic approximation becomes more accurate when T grows from 50 to 100.

To summarize, the evidence that emerges from the Monte Carlo simulations, although limited to a few particular cases, seems to support the asymptotic theory presented in this section. More precisely, the results show that the limiting distributions yield good approxi-

¹¹The results for any observation (i, t) , $i = 1, \dots, N$, and $t = 1, \dots, T$, are similar and hence do not change the general conclusions of the paper.

¹²The histograms for $\tau = 0.75$ are similar, thus were omitted to preserve space.

mations to the finite sample distributions of the QPC estimators $\hat{\lambda}_i(\tau)$, $\hat{f}_t(\tau)$, and $\hat{c}_{it}(\tau)$.

4 Conclusions

In this paper, I propose a novel concept, high dimensional quantile factor analysis, where the τ -th conditional quantile function of a set of observable variables has a factor structure. Also, both factors and factor loadings, as well as the number of factors, are allowed to be quantile-specific. Then, I provide a set of conditions under which these objects are identified but highlighting that the type of identification, namely local or global, depends crucially on the rotation considered by the econometrician. I propose a simple two-step iterative procedure to obtain the QPC estimators of the quantile factors and quantile factor loadings, which resembles the EM algorithm employed in PC estimation via maximum likelihood. Monte Carlo simulations highlight the good performance of the procedure in small to moderate sample sizes. In particular, the QPC estimator is more efficient than the PC estimator in nonlinear settings, and can satisfactorily recover factors affecting higher moments of the observable variables when PC estimator cannot. Lastly, under general assumptions, I establish uniform consistency, and I provide an asymptotic theory to derive the rates of convergence and the limiting distribution of the QPC estimators of the quantile factors, quantile factor loadings, and quantile common components when both N and T grows large jointly.

Admittedly, there are several aspects of this context that deserve further attention. First, the potential of the proposed framework can be illustrated with an interesting empirical application. In this sense, [Sagner \(2020\)](#) employs the QFA methodology to propose and estimate a new measure of systemic risk for the US from the information contained in asset returns. The author shows that in the context of the external habits formation model of [Campbell and Cochrane \(1999\)](#), and under the assumption that stock returns are heteroskedastic, the equilibrium risk premium has a location-scale factor structure where the factors are a monotonic transformation of the surplus consumption ratio, a state variable that captures the systemic risk in the structural model. Another area of research within this context is the development of an asymptotic theory for QFA models endowed with an approximate factor structure of traditional factor models. This development will help to build a statistical test in the spirit of [Connor and Korajczyk \(1993\)](#) and [Kapetanios \(2010\)](#) for determining the number of quantile factors. This issue is essential, especially in empirical research. Recall that this paper builds on the crucial assumption that the number of quantile factors $K(\tau)$ is known in advance. The intuition behind the proposed test is that, if the number of quantile factors is misspecified, then the quantile errors $\varepsilon_{it}^0(\tau)$ will be cross-correlated. In contrast, if the statistical model is correctly specified, the errors will be weakly correlated. Finally, another appealing area

of research, especially for finance and risk management, is the extension of the inferential theory further into the tails of the conditional distribution. As can be noticed from the results, as τ approaches the boundaries of the $(0, 1)$ interval, the proposed theory does not hold anymore since the asymptotic variance-covariance matrices are not well defined when $\tau \rightarrow 0$ or $\tau \rightarrow 1$. The incorporation of the asymptotics developed in the extreme quantile regression literature (see, for example, [Smith, 1994](#); [Portnoy and Jureckova, 1999](#); [Chernozhukov, 2005](#); [Chernozhukov and Fernandez-Val, 2011](#), among many others) into the high dimensional QFA framework is assuredly a promising starting point to tackle this issue.

A Appendix

This appendix provides detailed proofs for all theorems and propositions established in the main text. The notation used is as follows. $\|A\|$ stands for the Euclidean norm of a matrix (or a vector), i.e., $\|A\|^2 = \text{trace}(A'A)$. The symbols “ \xrightarrow{p} ” and “ \xrightarrow{d} ” denote convergence in probability and convergence in distribution, respectively, whereas $O_p(\cdot)$ and $o_p(\cdot)$ are the usual symbols for the order of convergence in probability.

A.1 Proof of Proposition 1

Proof. The proof consists of three parts.

First, consider Assumption 2.3a \Rightarrow Assumption 2.3b. For all $\tau \in (0, 1)$, let $F^0(\tau)$ and $\Lambda^0(\tau)$ satisfy Assumption 2.3a. Consider the partition of the quantile factor loadings matrix $\Lambda^0(\tau) = [\Lambda_1^0(\tau)', \Lambda_2^0(\tau)']$, where $\Lambda_1^0(\tau)$ is a $K(\tau) \times K(\tau)$ matrix. Next, consider the Gram-Schmidt decomposition $\Lambda_1^0(\tau)' = Q^0(\tau) R^0(\tau)$, where $Q^0(\tau)$ is an orthogonal matrix, and $R^0(\tau)$ is a non-singular upper triangular matrix, both of them of size $K(\tau) \times K(\tau)$. Define $\tilde{F}^0(\tau) = F^0(\tau) Q^0(\tau)$ and $\tilde{\Lambda}^0(\tau) = \Lambda^0(\tau) Q^0(\tau)$. Note that

$$\tilde{\Lambda}^0(\tau) = \begin{bmatrix} \tilde{\Lambda}_1^0(\tau) \\ \tilde{\Lambda}_2^0(\tau) \end{bmatrix} = \begin{bmatrix} R(\tau)' \\ \Lambda_2^0(\tau) Q^0(\tau) \end{bmatrix}$$

and

$$\frac{\tilde{F}^0(\tau)' \tilde{F}^0(\tau)}{T} = Q^0(\tau)' \left(\frac{F^0(\tau)' F^0(\tau)}{T} \right) Q^0(\tau) = I_{K(\tau)}$$

Hence, $\tilde{F}^0(\tau)$ and $\tilde{\Lambda}^0(\tau)$ satisfy Assumption 2.3b. Finally, note that $\tilde{C}^0(\tau) = \tilde{F}^0(\tau) \tilde{\Lambda}^0(\tau)' = F^0(\tau) Q^0(\tau) Q^0(\tau)' \Lambda^0(\tau)' = C^0(\tau)$. Therefore, Assumptions 2.3a and 2.3b are equivalent according to Definition 2.

Next, consider Assumption 2.3b \Rightarrow Assumption 2.3c. For all $\tau \in (0, 1)$, let $\tilde{F}^0(\tau)$ and $\tilde{\Lambda}^0(\tau)$ satisfy Assumption 2.3b. Consider the partition of the quantile factor loadings matrix $\tilde{\Lambda}(\tau) = [\tilde{\Lambda}_1^0(\tau)', \tilde{\Lambda}_2^0(\tau)']'$, where $\tilde{\Lambda}_1^0(\tau)$ is a $K(\tau) \times K(\tau)$ lower triangular matrix. Because all diagonal elements of $\tilde{\Lambda}_1^0(\tau)$ are non-zero, its inverse exists, and we can then define $\bar{\Lambda}^0(\tau) = \tilde{\Lambda}^0(\tau) \tilde{\Lambda}_1^0(\tau)^{-1} = [I_{K(\tau)}, (\tilde{\Lambda}_2^0(\tau) \tilde{\Lambda}_1^0(\tau)^{-1})']$ and $\bar{F}^0(\tau) = \tilde{F}^0(\tau) \tilde{\Lambda}_1^0(\tau)$. Thus, $\bar{F}^0(\tau)$ and $\bar{\Lambda}^0(\tau)$ satisfy Assumption 2.3c. Moreover, since $\bar{C}^0(\tau) = \bar{F}^0(\tau) \bar{\Lambda}^0(\tau) = \tilde{F}^0(\tau) \tilde{\Lambda}_1^0(\tau) \tilde{\Lambda}_1^0(\tau)^{-1} \tilde{\Lambda}^0(\tau) = \tilde{C}^0(\tau)$, then Assumptions 2.3b and 2.3c are equivalent according to Definition 2.

Lastly, consider Assumption 2.3c \Rightarrow Assumption 2.3a. For all $\tau \in (0, 1)$, let $\bar{F}^0(\tau)$ and $\bar{\Lambda}^0(\tau)$ satisfy Assumption 2.3c. Because $\text{rank}(\bar{C}^0(\tau)) = K(\tau)$ by Assumption 2.2, consider

the singular value decomposition

$$\bar{C}^0(\tau) = \bar{U}^0(\tau) \begin{bmatrix} \bar{D}^0(\tau) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \bar{V}^0(\tau)'$$

where $\bar{U}^0(\tau)$ is a $T \times T$ orthogonal matrix, $\bar{D}^0(\tau)$ is a diagonal matrix of size $K(\tau)$, and $\bar{V}^0(\tau)$ is an $N \times N$ orthogonal matrix. The partitions $\bar{U}^0(\tau) = [\bar{U}_1^0(\tau), \bar{U}_2^0(\tau)]$ and $\bar{V}^0(\tau) = [\bar{V}_1^0(\tau), \bar{V}_2^0(\tau)]$, where $\bar{U}_1^0(\tau)' \bar{U}_1^0(\tau) = \bar{V}_1^0(\tau)' \bar{V}_1^0(\tau) = I_{K(\tau)}$, imply that $\bar{C}^0(\tau) = \bar{U}_1^0(\tau) \bar{D}^0(\tau) \bar{V}_1^0(\tau)'$. With these elements at hand, define $F^0(\tau) = \sqrt{T} \bar{U}_1^0(\tau)$ and $\Lambda^0(\tau) = \bar{V}_1^0(\tau) \bar{D}^0(\tau) / \sqrt{T}$, and note that

$$\Lambda^0(\tau)' \Lambda^0(\tau) = \frac{\bar{D}^0(\tau) \bar{V}_1^0(\tau)' \bar{V}_1^0(\tau) \bar{D}^0(\tau)}{T} = \frac{\bar{D}^0(\tau)^2}{T}$$

which is a diagonal matrix, and

$$\frac{F^0(\tau)' F^0(\tau)}{T} = \bar{U}_1^0(\tau)' \bar{U}_1^0(\tau) = I_{K(\tau)}$$

Hence, $F^0(\tau)$ and $\Lambda^0(\tau)$ satisfy Assumption 2.3a, and because $\bar{C}^0(\tau) = C^0(\tau)$ by construction, then we conclude that Assumptions 2.3c and 2.3a are equivalent according to Definition 2. The proof is complete. \blacksquare

A.2 Proof of Theorem 1

The following definitions will be used extensively in this section. For all $\tau \in (0, 1)$, let $e_{it}(\tau, \lambda_i, f_t)$ and $\varepsilon_{it}^0(\tau)$ be the quantile factor error and the quantile factor residual, respectively, which are given by the following expressions

$$e_{it}(\tau, \lambda_i, f_t) = \lambda_i' f_t - \mathcal{Q}_{y_{it}}(\tau | \theta_{it}^0(\tau))$$

and

$$\varepsilon_{it}^0(\tau) = y_{it} - \mathcal{Q}_{y_{it}}(\tau | \theta_{it}^0(\tau))$$

where the τ -th quantile function of the observable variable y_{it} conditional on $\theta_{it} \equiv [\lambda_i', f_t']'$, $\mathcal{Q}_{y_{it}}(\tau | \theta_{it}^0(\tau))$, is given by equation (2). Before turning to the proof of Theorem 1, let us first consider two useful lemmas.

Lemma 1. *Suppose that Assumption 2.1 holds. Then, the function*

$$\mathcal{A}_\tau(\lambda_i, f_t) = E \left[\rho_\tau(\varepsilon_{it}^0(\tau) - e_{it}(\tau, \lambda_i, f_t)) - \rho_\tau(\varepsilon_{it}^0(\tau)) \mid \theta_{it}^0(\tau) \right]$$

can be alternatively expressed as

$$\begin{aligned} \mathcal{A}_\tau(\lambda_i, f_t) &= \frac{1}{2} g_{it} \left(G_{it}^{-1}(\tau \mid \theta_{it}^0(\tau)) \mid \theta_{it}^0(\tau) \right) \left[(\lambda_i - \lambda_i^0(\tau))' f_t^0(\tau) \right. \\ &\quad \left. + (f_t - f_t^0(\tau))' \lambda_i^0(\tau) \right]^2 + o_p \left(\|\theta_{it} - \theta_{it}^0(\tau)\|^2 \right) \end{aligned} \quad (\text{A.1})$$

for all i, t , and $\tau \in (0, 1)$.

Proof. For any $\tau \in (0, 1)$, using the definition of the quantile loss function $\rho_\tau(\cdot)$ and the conditional expectation yields

$$\begin{aligned} \mathcal{A}_\tau(\lambda_i, f_t) &= [G_{it}(\lambda_i' f_t \mid \theta_{it}^0(\tau)) - \tau] \lambda_i' f_t \\ &\quad + \tau \left[\int_{\lambda_i' f_t}^{\infty} u_{it} g_{it}(u_{it} \mid \theta_{it}^0(\tau)) du_{it} - \int_{\lambda_i^0(\tau)' f_t^0(\tau)}^{\infty} u_{it} g_{it}(u_{it} \mid \theta_{it}^0(\tau)) du_{it} \right] \\ &\quad + (\tau - 1) \left[\int_{-\infty}^{\lambda_i' f_t} u_{it} g_{it}(u_{it} \mid \theta_{it}^0(\tau)) du_{it} - \int_{-\infty}^{\lambda_i^0(\tau)' f_t^0(\tau)} u_{it} g_{it}(u_{it} \mid \theta_{it}^0(\tau)) du_{it} \right] \end{aligned}$$

Based on the previous expression, the gradient and the Hessian of $\mathcal{A}_\tau(\lambda_i, f_t)$ is given, respectively, by

$$\nabla_\tau^{\mathcal{A}}(\lambda_i, f_t) = (G_{it}(\lambda_i' f_t \mid \theta_{it}^0(\tau)) - \tau) \begin{bmatrix} f_t \\ \lambda_i \end{bmatrix}$$

and

$$H_\tau^{\mathcal{A}}(\lambda_i, f_t) = g_{it}(\lambda_i' f_t \mid \theta_{it}^0(\tau)) \begin{bmatrix} f_t f_t' & f_t \lambda_i' + h_\tau^{\mathcal{A}}(\lambda_i, f_t) I_{K(\tau)} \\ \lambda_i f_t' + h_\tau^{\mathcal{A}}(\lambda_i, f_t) I_{K(\tau)} & \lambda_i \lambda_i' \end{bmatrix}$$

where $h_\tau^{\mathcal{A}}(\lambda_i, f_t) \equiv (G_{it}(\lambda_i' f_t \mid \theta_{it}^0(\tau)) - \tau) / g_{it}(\lambda_i' f_t \mid \theta_{it}^0(\tau))$ is a scalar that depends on the conditional cumulative distribution function and the conditional density of y_{it} . Thus, a

second-order Taylor expansion of the function $\mathcal{A}_\tau(\lambda_i, f_t)$ around $\theta_{it}^0(\tau)$ is as follows

$$\begin{aligned}
\mathcal{A}_\tau(\lambda_i, f_t) &= \mathcal{A}_\tau(\lambda_i^0(\tau), f_t^0(\tau)) + [(\lambda_i - \lambda_i^0(\tau))', (f_t - f_t^0(\tau))'] \nabla_\tau^{\mathcal{A}}(\lambda_i^0(\tau), f_t^0(\tau)) \\
&\quad + [(\lambda_i - \lambda_i^0(\tau))', (f_t - f_t^0(\tau))'] H_\tau^{\mathcal{A}}(\lambda_i^0(\tau), f_t^0(\tau)) \begin{bmatrix} \lambda_i - \lambda_i^0(\tau) \\ f_t - f_t^0(\tau) \end{bmatrix} \\
&\quad + o_p\left(\|\theta_{it} - \theta_{it}^0(\tau)\|^2\right) \\
&= \frac{1}{2} g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \Big| \theta_{it}^0(\tau) \right) [(\lambda_i - \lambda_i^0(\tau))' f_t^0(\tau) + (f_t - f_t^0(\tau))' \lambda_i^0(\tau)]^2 \\
&\quad + o_p\left(\|\theta_{it} - \theta_{it}^0(\tau)\|^2\right)
\end{aligned}$$

for all i, t , and $\tau \in (0, 1)$. This last result completes the proof. \blacksquare

Next, the following Lemma provides a first-order approximation of the quantile factor error $e_{it}(\tau, \lambda_i, f_t)$.

Lemma 2. *For any i, t , and $\tau \in (0, 1)$, the quantile factor error $e_{it}(\tau, \lambda_i, f_t)$ can be alternatively expressed as*

$$e_{it}(\tau, \lambda_i, f_t) = [\lambda_i - \lambda_i^0(\tau)]' f_t^0(\tau) + [f_t - f_t^0(\tau)]' \lambda_i^0(\tau) + o_p(\|\theta_{it} - \theta_{it}^0(\tau)\|) \quad (\text{A.2})$$

Proof. For any i, t , and $\tau \in (0, 1)$, a first-order Taylor expansion of $e_{it}(\tau, \lambda_i, f_t)$ around $\theta_{it}^0(\tau)$ results in

$$\begin{aligned}
e_{it}(\tau, \lambda_i, f_t) &= e_{it}(\tau, \lambda_i^0, f_t^0) + [(\lambda_i - \lambda_i^0(\tau))', (f_t - f_t^0)'] \begin{bmatrix} f_t^0(\tau) \\ \lambda_i^0(\tau) \end{bmatrix} \\
&\quad + o_p(\|\theta_{it} - \theta_{it}^0(\tau)\|) \\
&= (\lambda_i - \lambda_i^0(\tau))' f_t^0(\tau) + (f_t - f_t^0(\tau))' \lambda_i^0(\tau) + o_p(\|\theta_{it} - \theta_{it}^0(\tau)\|)
\end{aligned}$$

The last result completes the proof. \blacksquare

Proof of Theorem 1. For any $\tau \in (0, 1)$, let $S_\tau(\Lambda, F)$ be the objective function defined in expression (4). Using the definition of the quantile factor errors $e_{it}(\tau, \lambda_i, f_t)$ and the quantile factor residuals $\varepsilon_{it}^0(\tau)$, we have that $S_\tau(\Lambda, F) = E[\tilde{S}_\tau(\Lambda, F)]$, where

$$\tilde{S}_\tau(\Lambda, F) = E \left[\sum_{i=1}^N \sum_{t=1}^T \rho_\tau(\varepsilon_{it}^0(\tau) - e_{it}(\tau, \lambda_i, f_t)) \Big| \boldsymbol{\theta}^0(\tau) \right]$$

and $\boldsymbol{\theta}^0(\tau) \equiv [\Lambda^0(\tau)', F^0(\tau)']'$. Note that minimizing $S_\tau(\Lambda, F)$ for $\boldsymbol{\theta} = [\Lambda', F']'$ is equivalent to minimizing $\tilde{S}_\tau(\Lambda, F)$ for the same argument. Moreover, because the quantile factor residual

is independent of both Λ and F , the latter is, in turn, equivalent to minimizing $\tilde{S}_\tau^*(\Lambda, F)$ for $\boldsymbol{\theta}$, where

$$\tilde{S}_\tau^*(\Lambda, F) = E \left[\sum_{i=1}^N \sum_{t=1}^T \left\{ \rho_\tau(\varepsilon_{it}^0(\tau) - e_{it}(\tau, \lambda_i, f_t)) - \rho_\tau(\varepsilon_{it}^0(\tau)) \right\} \middle| \boldsymbol{\theta}^0(\tau) \right]$$

Using Lemma 1 followed by Lemma 2, the above equation can be expressed as

$$\begin{aligned} \tilde{S}_\tau^*(\Lambda, F) &= \frac{1}{2} \sum_{i=1}^N \sum_{t=1}^T g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) \left[(\lambda_i - \lambda_i^0(\tau))' f_t^0(\tau) \right. \\ &\quad \left. + (f_t - f_t^0(\tau))' \lambda_i^0(\tau) \right]^2 + o_p(1) \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{t=1}^T g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) [\lambda_i' f_t - \lambda_i^0(\tau)' f_t^0(\tau)]^2 + o_p(1) \\ &= \text{trace} \left[(F\Lambda' - F^0(\tau)\Lambda^0(\tau)')' W(\tau) (F\Lambda' - F^0(\tau)\Lambda^0(\tau)') \right] \end{aligned}$$

where $W(\tau)$ is a matrix of weights whose elements $w_{it} \equiv g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) / 2$ are strictly greater than 0, for all i, t , and $\tau \in (0, 1)$, by Assumption 2.1.

Let $\bar{\theta}_{it}(\tau) = [\bar{\lambda}_i(\tau)', \bar{f}_t(\tau)']'$ be the values of the quantile factors and quantile factor loadings, respectively, that minimize $S_\tau(\Lambda, F)$, and thus also $\tilde{S}_\tau^*(\Lambda, F)$. Then

$$\begin{aligned} \tilde{S}_\tau^*(\bar{\Lambda}(\tau), \bar{F}(\tau)) &= \arg \min_{[\Lambda', F']' \in \Theta} \tilde{S}_\tau^*(\Lambda, F) \\ &= \tilde{S}_\tau^*(\Lambda^0(\tau), F^0(\tau)) \\ &= 0 \end{aligned}$$

since $\tilde{S}_\tau^*(\Lambda^0(\tau), F^0(\tau)) = 0$ by construction. The above results imply that

$$\begin{aligned} \text{trace} \left[(\bar{F}(\tau)\bar{\Lambda}(\tau)' - F^0(\tau)\Lambda^0(\tau)')' W(\tau) (\bar{F}(\tau)\bar{\Lambda}(\tau)' - F^0(\tau)\Lambda^0(\tau)') \right] &= 0 \\ \text{trace} \left[(\bar{C}(\tau) - C^0(\tau))' W(\tau) (\bar{C}(\tau) - C^0(\tau)) \right] &= 0 \end{aligned}$$

Because of the properties of the trace operator and Assumption 2.2, the above equation implies that $\bar{C}(\tau) = C^0(\tau)$ for any $\tau \in (0, 1)$, i.e., the common component of the QFA model is identified.

Finally, to see how the quantile factors $\bar{F}(\tau)$ and the quantile factor loadings $\bar{\Lambda}(\tau)$ are

individually identified, consider the singular value decomposition of $\bar{C}(\tau)$

$$\bar{C}(\tau) = \bar{U}(\tau) \begin{bmatrix} \bar{D}(\tau) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \bar{V}(\tau)'$$

where $\bar{U}(\tau)$ is a $T \times T$ orthogonal matrix, \bar{D} is a $K(\tau) \times K(\tau)$ diagonal matrix, and $\bar{V}(\tau)$ is an $N \times N$ orthogonal matrix. Consider the partitions $\bar{U}(\tau) = [\bar{U}_1(\tau), \bar{U}_2(\tau)]$ and $\bar{V}(\tau) = [\bar{V}_1(\tau), \bar{V}_2(\tau)]$ such that $\bar{C}(\tau) = \bar{U}_1(\tau) \bar{D}(\tau) \bar{V}_1(\tau)'$. Define

$$\begin{aligned} \bar{F}(\tau) &= \sqrt{T} \bar{U}_1(\tau) \\ \bar{\Lambda}(\tau) &= \frac{1}{\sqrt{T}} \bar{V}_1(\tau) \bar{D}(\tau) \end{aligned}$$

and note that $\bar{F}(\tau)' \bar{F}(\tau) / T = \bar{U}_1(\tau)' \bar{U}_1(\tau) = I_{K(\tau)}$ and $\bar{\Lambda}(\tau)' \bar{\Lambda}(\tau) = \bar{D}(\tau)^2 / T$ is a diagonal matrix. Hence, both $\bar{F}(\tau)$ and $\bar{\Lambda}(\tau)$ are individually identified (up to a column-sign change) under Assumption 2.3a, and by Proposition 1 all rotations considered in Assumption 2.3 are equivalent. The proof is complete. \blacksquare

A.3 Proof of Theorem 2

Definition 3 entails that the QPC estimator $\hat{\boldsymbol{\theta}}(\tau)$ is the one that solves the following optimization problem

$$\hat{\boldsymbol{\theta}}(\tau) = \arg \min_{\{\Lambda, F\}} V_\tau(\Lambda, F), \quad \tau \in (0, 1)$$

where the objective function $V_\tau(\Lambda, F)$ was defined in equation (5). One aspect to note is that because $V_\tau(\Lambda^0(\tau), F^0(\tau))$ does not depend on both Λ and F , then the previous expression is equivalent to

$$\hat{\boldsymbol{\theta}}(\tau) = \arg \min_{\{\Lambda, F\}} V_\tau^*(\Lambda, F), \quad \tau \in (0, 1)$$

where $V_\tau^*(\Lambda, F) \equiv V_\tau(\Lambda, F) - V_\tau(\Lambda^0(\tau), F^0(\tau))$ is just the objective function centered about $V_\tau(\Lambda^0(\tau), F^0(\tau))$. Moreover, note that because $\hat{\boldsymbol{\theta}}(\tau)$ is the minimizer of $V_\tau(\Lambda, F)$, and hence of $V_\tau^*(\Lambda, F)$, the previous equation evaluated at the QPC estimator is equal to 0 with probability close to 1 for all $\tau \in (0, 1)$.

The following definitions will be used extensively in this section. For all i, t , and $\tau \in (0, 1)$, we define $\hat{\phi}_{it}(\tau) \equiv \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) - \lambda_i^0(\tau)' f_t^0(\tau)$ as the difference between the estimated and the actual quantile common component of the observable variable y_{it} , and $\hat{\Phi}(\tau)$ a $T \times N$ matrix whose elements are $\hat{\phi}_{it}(\tau)$. By using these definitions, the centered objective function

evaluated at the QPC estimator can be characterized by the following expression

$$V_\tau^*(\hat{\Phi}(\tau)) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left[\rho_\tau(\varepsilon_{it}^0(\tau) - \hat{\phi}_{it}(\tau)) - \rho_\tau(\varepsilon_{it}^0(\tau)) \right] \quad (\text{A.3})$$

where $\varepsilon_{it}^0(\tau) \equiv y_{it} - \mathcal{Q}_{y_{it}}(\tau)$ is the quantile factor residual defined in Section 3.1. Using the identity by [Knight \(1998\)](#), we have that $V_\tau^*(\hat{\Phi}(\tau)) = W_\tau^*(\hat{\Phi}(\tau)) + Z_\tau^*(\hat{\Phi}(\tau))$, where

$$W_\tau^*(\hat{\Phi}(\tau)) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \int_0^{\hat{\phi}_{it}(\tau)} [\mathbb{1}\{\varepsilon_{it}^0(\tau) < s\} - \mathbb{1}\{\varepsilon_{it}^0(\tau) < 0\}] ds \quad (\text{A.4})$$

$$Z_\tau^*(\hat{\Phi}(\tau)) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \psi_\tau(\varepsilon_{it}^0(\tau)) \hat{\phi}_{it}(\tau) \quad (\text{A.5})$$

Lastly, for all i, t , and $\tau \in (0, 1)$, we define

$$\begin{aligned} b_{it}(\tau, \phi) &\equiv \mathbb{1}\{\varepsilon_{it}^0(\tau) < \phi\} \\ \xi_{it}(\tau, \phi) &\equiv [b_{it}(\tau, \phi) - b_{it}(\tau, 0)] - E[b_{it}(\tau, \phi) - b_{it}(\tau, 0)] \end{aligned}$$

The following lemmas will be useful in providing an upper and lower bound for $W_\tau^*(\hat{\Phi}(\tau))$, as well as in deriving their asymptotic properties. They use an argument similar to the ones exposed in [Oka and Qu \(2011\)](#).

Lemma 3. *For every $\tau \in (0, 1)$,*

$$\begin{aligned} 0 &\leq \frac{1}{2NT} \sum_{i=1}^N \sum_{t=1}^T \left[b_{it}(\tau, \hat{\phi}_{it}(\tau)/2) - b_{it}(\tau, 0) \right] \hat{\phi}_{it}(\tau) \leq W_\tau^*(\hat{\Phi}(\tau)) \\ &\leq \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left[b_{it}(\tau, \hat{\phi}_{it}(\tau)) - b_{it}(\tau, 0) \right] \hat{\phi}_{it}(\tau) \end{aligned}$$

Proof. See Lemma A.1 of [Oka and Qu \(2011\)](#). The details are omitted. ■

Lemma 4. *Suppose that Assumptions 3.1 to 3.3 hold. For all i, t and $\tau \in (0, 1)$, let $\Theta = \left\{ \hat{\phi}_{it}(\tau) \in \mathbb{R} : \left| \hat{\phi}_{it}(\tau) \right| = \left(\min \left\{ \sqrt{N}, \sqrt{T} \right\} \right)^{-1} \cdot B \right\}$ be a compact set, where $B < \infty$ is an arbitrary positive constant. Then*

$$\sup_{1 \leq s \leq T} \sup_{\hat{\phi}_{it}(\tau) \in \Theta} \left| (NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^s \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right| = o_p(1)$$

Proof. The proof considers fixed $\hat{\phi}_{it}(\tau)$. Uniform convergence over Θ is guaranteed by the

compactness of this set and the monotonicity of $b_{it}(\tau, \hat{\phi}_{it}(\tau))$ in $\hat{\phi}_{it}(\tau)$.

First, for any $\hat{\phi}_{it}(\tau) \in \Theta$, $\xi_{it}(\tau, \hat{\phi}_{it}(\tau))$ satisfies

$$\begin{aligned}
E \left[\left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right|^2 \middle| \theta_{it}^0(\tau) \right] &\leq \left| G_{it} \left(\lambda_i^0(\tau)' f_t^0(\tau) + \hat{\phi}_{it}(\tau) \middle| \theta_{it}^0(\tau) \right) \right. \\
&\quad \left. - G_{it} \left(\lambda_i^0(\tau)' f_t^0(\tau) \middle| \theta_{it}^0(\tau) \right) \right| \\
&= g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) \cdot \left| \hat{\phi}_{it}(\tau) \right| \\
&\leq \left(\min \left\{ \sqrt{N}, \sqrt{T} \right\} \right)^{-1} BU_g
\end{aligned} \tag{A.6}$$

where in the second inequality, I used the Mean Value Theorem and Assumption 3.3, and in the last inequality, I used Assumption 3.2.

Next, using the Doob inequality followed by the Rosenthal inequality (Hall and Heyde, 1980, pp. 15 and 23), we have that

$$\begin{aligned}
P \left[\sup_{1 \leq s \leq T} \left| (NT)^{-1/2} \sum_{i=1}^N \sum_{t=1}^s \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right| > \epsilon \right] \\
\leq \frac{M}{\epsilon^{2\gamma}} \left\{ (NT)^{-\gamma} E \left[\left(\sum_{i=1}^N \sum_{t=1}^T E \left[\left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right|^2 \middle| \theta_{it}^0(\tau) \right] \right)^\gamma \right] \right. \\
\left. + (NT)^{-\gamma} \sum_{i=1}^N \sum_{t=1}^T E \left[\left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right|^{2\gamma} \right] \right\}
\end{aligned}$$

where M is a positive constant that depends only on $\gamma > 2$. By using equation (A.6), the first term inside of the curly brackets is given by

$$\begin{aligned}
(NT)^{-\gamma} E \left[\left(\sum_{i=1}^N \sum_{t=1}^T E \left[\left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right|^2 \middle| \theta_{it}^0(\tau) \right] \right)^\gamma \right] \\
\leq (NT)^{-\gamma} E \left[\left(\sum_{i=1}^N \sum_{t=1}^T \left(\min \left\{ \sqrt{N}, \sqrt{T} \right\} \right)^{-1} BU_g \right)^\gamma \right] \\
= \left(\min \left\{ \sqrt{N}, \sqrt{T} \right\} \right)^{-\gamma} B^\gamma U_g^\gamma \rightarrow 0 \text{ as } N, T \rightarrow \infty
\end{aligned} \tag{A.7}$$

Similarly, the second term inside of the curly brackets can be rewritten as

$$\begin{aligned}
& (NT)^{-\gamma} \sum_{i=1}^N \sum_{t=1}^T E \left[\left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right|^{2\gamma} \right] \\
&= (NT)^{-\gamma} \sum_{i=1}^N \sum_{t=1}^T E \left[E \left[\left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right|^{2\gamma-2} \left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right|^2 \left| \theta_{it}^0(\tau) \right| \right] \right] \\
&\leq (NT)^{-\gamma} \sum_{i=1}^N \sum_{t=1}^T \left(\min \left\{ \sqrt{N}, \sqrt{T} \right\} \right)^{-1} BU_g \\
&= (NT)^{1-\gamma} \left(\min \left\{ \sqrt{N}, \sqrt{T} \right\} \right)^{-1} BU_g \rightarrow 0 \text{ as } N, T \rightarrow \infty \tag{A.8}
\end{aligned}$$

where in the second inequality, I used the fact that $\left| \xi_{it}(\tau, \hat{\phi}_{it}(\tau)) \right| \leq 1$. Equations (A.7) and (A.8) together provide the result stated in this lemma. \blacksquare

Proof of Theorem 2. The proof consists of two parts. In the first part, I show the uniform consistency of the quantile common component QPC estimator $\hat{C}(\tau)$ by employing a proof by contradiction. In particular, I show that if uniform consistency does not hold, then the objective function centered about $V_\tau(\Lambda^0(\tau), F^0(\tau))$ and evaluated at $\hat{\theta}(\tau)$ is strictly positive with probability close to 1, implying that $\hat{C}(\tau)$ is not its minimizer. In the second part, I exploit the uniform consistency of the quantile common component to show the consistency of both the quantile factors and the quantile factor loadings using an argument similar to Lemma 1 in [Chen et al. \(2014\)](#).

First, consider the proof for the estimated quantile common component $\hat{c}_{it}(\tau) = \hat{\lambda}_i(\tau)' \hat{f}_t(\tau)$. Because the centered objective function $V_\tau^*(\hat{\Phi}(\tau))$ given in (A.3) is convex in $\hat{\phi}_{it}(\tau) = \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) - \lambda_i^0(\tau)' f_t^0(\tau)$, it suffices to consider its property over this argument satisfying $\min \left\{ \sqrt{N}, \sqrt{T} \right\} \cdot \left| \hat{\phi}_{it}(\tau) \right| = B$, where B is some arbitrary positive constant. By [Knight \(1998\)](#) identity, $V_\tau^*(\hat{\Phi}(\tau)) = W_\tau^*(\hat{\Phi}(\tau)) + Z_\tau^*(\hat{\Phi}(\tau))$ (equations A.4 and A.5), so that we can analyze each term separately.

Start with $W_\tau^*(\hat{\Phi}(\tau))$. By Lemma 3 and the triangle inequality, the term $\min \{N, T\} \cdot W_\tau^*(\hat{\Phi}(\tau))$ is bounded below by the following expression

$$\begin{aligned}
& \frac{\min \{N, T\}}{2NT} \left\{ \sum_{i=1}^N \sum_{t=1}^T E \left[\left(b_{it}(\tau, \hat{\phi}_{it}(\tau)/2) - b_{it}(\tau, 0) \right) \hat{\phi}_{it}(\tau) \left| \theta_{it}^0(\tau) \right| \right] \right. \\
& \quad \left. - \left| \sum_{i=1}^N \sum_{t=1}^T \xi_{it}(\tau, \hat{\phi}_{it}(\tau)/2) \hat{\phi}_{it}(\tau) \right| \right\}
\end{aligned}$$

Using the Mean Value Theorem and Assumption 3.3, the first term inside of the curly

brackets can be rewritten as

$$\begin{aligned} \frac{\min\{N, T\}}{4NT} \sum_{i=1}^N \sum_{t=1}^T g_{it} (G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) | \theta_{it}^0(\tau)) \cdot \left| \hat{\phi}_{it}(\tau) \right|^2 &\geq \frac{\min\{N, T\}}{4NT} L_g \sum_{i=1}^N \sum_{t=1}^T \left| \hat{\phi}_{it}(\tau) \right|^2 \\ &= \frac{1}{4} B^2 L_g \end{aligned}$$

where in the derivation of the result, I invoke Assumption 3.2. By Lemma 4, the second term inside of the curly brackets is of order $o_p((\max\{\sqrt{N}, \sqrt{T}\})^{-1}B)$. Hence,

$$\min\{N, T\} \cdot W_\tau^*(\hat{\Phi}(\tau)) \geq \frac{1}{4} B^2 L_g \quad (\text{A.9})$$

For $Z_\tau^*(\hat{\Phi}(\tau))$, we have that

$$\min\{N, T\} \cdot \left| Z_\tau^*(\hat{\Phi}(\tau)) \right| \leq \min\{N^{-1/2}T^{-1}, N^{-1}T^{-1/2}\} \cdot B \left| \sum_{i=1}^N \sum_{t=1}^T \psi_\tau(\varepsilon_{it}^0(\tau)) \right| \quad (\text{A.10})$$

Applying the Hájek-Rényi inequality for martingales (see [Chow and Teicher, 1997](#), p. 255) to the previous expression yields

$$\begin{aligned} P \left[\sup_{1 \leq s \leq T} \left| \min\{N^{-1/2}s^{-1}, N^{-1}s^{-1/2}\} \sum_{i=1}^N \sum_{t=1}^s \psi_\tau(\varepsilon_{it}^0(\tau)) \right| > C \right] \\ \leq \frac{1}{C^2} \sum_{t=1}^T \min\{t^{-2}, (Nt)^{-1}\} \cdot E \left[\left| \psi_\tau(\varepsilon_{it}^0(\tau)) \right|^2 \right] \end{aligned}$$

where C is an arbitrary constant. Note that $E[|\psi_\tau(\varepsilon_{it}^0(\tau))|^2] = \tau(1-\tau) < \infty$ for all i and t . Moreover, $\sum_{t=1}^T t^{-1}$ and $\sum_{t=1}^T t^{-2}$ are known as the generalized harmonic numbers of order $s = 1$ and $s = 2$, respectively, which converge to the Riemann zeta function $\zeta(s)$ as $T \rightarrow \infty$. Thus, in the case of $s = 1$, $\sum_{t=1}^T t^{-1} = \log T + \bar{\gamma}$, where $\bar{\gamma} \approx 0.577$ is the Euler-Mascheroni constant, whereas in the case of $s = 2$, $\sum_{t=1}^T t^{-2} = \pi^2/6$. Both results imply that the right-hand side of the previous inequality can be made arbitrarily small by choosing a large C . Consequently, if B is large, expression (A.9) is the dominant term asymptotically. The previous argument implies that $V_\tau^*(\hat{\Phi}(\tau))$ is strictly positive with probability close to 1 for large N and T . However, this is a contradiction since $\hat{\phi}_{it}(\tau)$ is the minimizer of the centered objective function $V_\tau^*(\hat{\Phi}(\tau))$. In other words, this function has to be equal to zero

with probability close to 1 as $N, T \rightarrow \infty$. Therefore, we conclude that

$$\min \left\{ \sqrt{N}, \sqrt{T} \right\} \cdot \left| \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) - \lambda_i^0(\tau)' f_t^0(\tau) \right| = O_p(1) \quad (\text{A.11})$$

For the second part of the proof, note that as $N, T \rightarrow \infty$, the previous result implies that $\left| \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) - \lambda_i^0(\tau)' f_t^0(\tau) \right| = o_p(1)$. Let $\hat{\phi}_i^\lambda(\tau) \equiv \hat{\lambda}_i(\tau) - \lambda_i^0(\tau)$ and $\hat{\phi}_t^f(\tau) \equiv \hat{f}_t(\tau) - f_t^0(\tau)$ be the difference between the QPC estimator and the actual quantile factor loading and quantile factor, respectively, for all i, t , and $\tau \in (0, 1)$. Hence, by Lemma 2, we have that

$$\begin{aligned} \left| \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) - \lambda_i^0(\tau)' f_t^0(\tau) \right| &= \left| \hat{\phi}_i^\lambda(\tau)' f_t^0(\tau) + \lambda_i^0(\tau)' \hat{\phi}_t^f(\tau) \right| \\ &\leq \left\| \hat{\phi}_i^\lambda(\tau) \right\| \cdot \left\| f_t^0(\tau) \right\| + \left\| \hat{\phi}_t^f(\tau) \right\| \cdot \left\| \lambda_i^0(\tau) \right\| \\ &= o_p \left(\sqrt{T} \left\| \hat{\phi}_i^\lambda(\tau) \right\| \right) + o_p \left(\sqrt{N} \left\| \hat{\phi}_t^f(\tau) \right\| \right) \end{aligned} \quad (\text{A.12})$$

where in the second inequality, I used the Cauchy-Schwarz inequality (see [Lütkepohl, 1996](#), p. 111), and the second part of Assumptions 3.4 and 3.5 supports the result (A.12). Thus, the only way that the previous sum is $o_p(1)$, as shown previously, occurs when

$$\sqrt{T} \left\| \hat{\phi}_i^\lambda(\tau) \right\| = O_p(1)$$

and

$$\sqrt{N} \left\| \hat{\phi}_t^f(\tau) \right\| = O_p(1)$$

uniformly in i and t , respectively, and for any $\tau \in (0, 1)$. The last two results complete the proof. \blacksquare

A.4 Proof of Theorem 3

Start by noting that, for all $\tau \in (0, 1)$, both subgradients given by equations (9) and (10) can be alternatively rewritten as follows

$$R_{i,\tau}^\Lambda(\Lambda, F) = \tilde{R}_{i,\tau}^\Lambda(\Lambda, F) + (NT)^{-1/2} \sum_{t=1}^T \left[G_{it}(\lambda_i' f_t | \theta_{it}^0(\tau)) - \tau \right] f_t \quad (\text{A.13})$$

where

$$\tilde{R}_{i,\tau}^\Lambda(\Lambda, F) = (NT)^{-1/2} \sum_{t=1}^T \left[\mathbf{1}\{y_{it} - \lambda_i' f_t < 0\} - G_{it}(\lambda_i' f_t | \theta_{it}^0(\tau)) \right] f_t$$

Similarly,

$$R_{t,\tau}^F(\Lambda, F) = \tilde{R}_{t,\tau}^F(\Lambda, F) + (NT)^{-1/2} \sum_{i=1}^N [G_{it}(\lambda'_i f_t | \theta_{it}^0(\tau)) - \tau] \lambda_i \quad (\text{A.14})$$

with

$$\tilde{R}_{t,\tau}^F(\Lambda, F) = (NT)^{-1/2} \sum_{i=1}^N [\mathbb{1}\{y_{it} - \lambda'_i f_t < 0\} - G_{it}(\lambda'_i f_t | \theta_{it}^0(\tau))] \lambda_i$$

Next, for all i, t , and $\tau \in (0, 1)$, define $\hat{\phi}_i^\lambda(\tau) \equiv \hat{\lambda}_i(\tau) - \lambda_i^0(\tau)$ and $\hat{\phi}_t^f(\tau) \equiv \hat{f}_t(\tau) - f_t^0(\tau)$ as the difference between the QPC estimator and the true quantile factor loading and quantile factor, respectively. Also, let $\hat{\Phi}^\Lambda(\tau)$ be an $N \times K(\tau)$ matrix whose i -th row is $\hat{\phi}_i^\lambda(\tau)'$ and $\hat{\Phi}^F(\tau)$ a $T \times K(\tau)$ matrix whose t -th row is $\hat{\phi}_t^f(\tau)'$.

The following lemma will be useful in the derivation of the asymptotic normality of $\hat{\Lambda}(\tau) = [\hat{\lambda}_1(\tau), \dots, \hat{\lambda}_N(\tau)]'$ and $\hat{F}(\tau) = [\hat{f}_1(\tau), \dots, \hat{f}_T(\tau)]'$. It uses an argument similar to the one exposed in [Qu \(2008\)](#).

Lemma 5. *Suppose that Assumptions 3 and 4 hold. Then, as $N, T \rightarrow \infty$, for any $\tau \in (0, 1)$,*

1. *Uniformly in i , if $\sqrt{T}/N \rightarrow 0$*

$$R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), \hat{F}(\tau)) = R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau)) + \sqrt{\frac{T}{N}} H_i^0(\tau) \hat{\phi}_i^\lambda(\tau) + o_p(T/\sqrt{N})$$

2. *Uniformly in t , if $\sqrt{N}/T \rightarrow 0$*

$$R_{t,\tau}^F(\hat{\Lambda}(\tau), \hat{F}(\tau)) = R_{t,\tau}^F(\Lambda^0(\tau), F^0(\tau)) + \sqrt{\frac{N}{T}} J_t^0(\tau) \hat{\phi}_t^f(\tau) + o_p(N/\sqrt{T})$$

Proof. I only prove part 1 of the lemma, since the second part can be derived from similar arguments. Thus, consider the sub-gradient $R_{i,\tau}^\Lambda(\Lambda, F)$ evaluated at the QPC estimators $\hat{\boldsymbol{\theta}}(\tau) = [\hat{\Lambda}(\tau)', \hat{F}(\tau)']'$. Then, for any $\tau \in (0, 1)$, we have that

$$R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), \hat{F}(\tau)) = \underbrace{R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), \hat{F}(\tau)) - R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), F^0(\tau))}_{(a)} + \underbrace{R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), F^0(\tau))}_{(b)} \quad (\text{A.15})$$

Term (a) on the right-hand side of the above expression satisfies

$$\begin{aligned}
\|(a)\| &= (NT)^{-1/2} \left\| \sum_{t=1}^T \psi_\tau(y_{it} - \hat{\lambda}_i(\tau)' \hat{f}_t(\tau)) \hat{f}_t(\tau) - \sum_{t=1}^T \psi_t(y_{it} - \hat{\lambda}_i(\tau)' f_t^0(\tau)) f_t^0(\tau) \right\| \\
&\leq (NT)^{-1/2} \sum_{t=1}^T \left| \mathbf{1}\{y_{it} - \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) < 0\} - \mathbf{1}\{y_{it} - \hat{\lambda}_i(\tau)' f_t^0(\tau) < 0\} \right| \cdot \|f_t^0(\tau)\| \\
&\quad + (NT)^{-1/2} \sum_{t=1}^T \left| \mathbf{1}\{y_{it} - \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) < 0\} - \tau \right| \cdot \|\hat{\phi}_t^f(\tau)\|
\end{aligned}$$

The first term is bounded above by

$$\begin{aligned}
(NT)^{-1/2} \sum_{t=1}^T \|f_t^0(\tau)\| &\leq N^{-1/2} \sum_{t=1}^T \sup_{1 \leq t \leq T} T^{-1/2} \|f_t^0(\tau)\| \\
&= o_p(T/\sqrt{N})
\end{aligned} \tag{A.16}$$

where in the last part, I used Assumption 3.4. By proceeding similarly, the second term is bounded above by

$$(NT)^{-1/2} \sum_{t=1}^T \|\hat{\phi}_t^f(\tau)\| = O_p(\sqrt{T}/N) \tag{A.17}$$

because of Theorem 2. Therefore, if $\sqrt{T}/N \rightarrow 0$, results (A.16) and (A.17) imply $(a) = o_p(1) + o_p(T/\sqrt{N}) = o_p(T/\sqrt{N})$. Next, consider the second term on the right-hand side of (A.15), which can be rewritten in the following manner using expression (A.13)

$$\begin{aligned}
(b) &= \underbrace{\tilde{R}_{i,\tau}^\Lambda(\Lambda^0(\tau) + \hat{\Phi}^\Lambda(\tau), F^0(\tau)) - \tilde{R}_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau))}_{(b)'} + \underbrace{\tilde{R}_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau))}_{(b)''} \\
&\quad + \underbrace{(NT)^{-1/2} \sum_{t=1}^T \left[G_{it} \left((\lambda_i^0(\tau) + \hat{\phi}_i^\lambda(\tau))' f_t^0(\tau) \middle| \theta_{it}^0(\tau) \right) - \tau \right] f_t^0(\tau)}_{(b)'''}
\end{aligned}$$

First, $\sqrt{N}(b)' = o_p(1)$ by Theorem A.3 in Bai (1996), which means that $(b)' = o_p(N^{-1/2})$. Next, using equation (A.13), the second term of the previous expression is equal to

$$\begin{aligned}
(b)'' &= R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau)) \\
&\quad - (NT)^{-1/2} \sum_{t=1}^T [G_{it}(\lambda_i^0(\tau)' f_t^0(\tau) | \theta_{it}^0(\tau)) - \tau] f_t^0(\tau) \\
&= R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau))
\end{aligned} \tag{A.18}$$

Finally, for the last term on the right-hand side of (b), we have that

$$\begin{aligned}
(b)''' &= (NT)^{-1/2} \sum_{t=1}^T \left[G_{it} \left((\lambda_i^0(\tau) + \hat{\phi}_i^\lambda(\tau))' f_t^0(\tau) \middle| \theta_{it}^0(\tau) \right) \right. \\
&\quad \left. - G_{it}(\lambda_i^0(\tau)' f_t^0(\tau) | \theta_{it}^0(\tau)) \right] f_t^0(\tau) \\
&= (NT)^{-1/2} \left(\sum_{t=1}^T g_{it} \left(G_{it}^{-1}(\tau | \theta_{it}^0(\tau)) \middle| \theta_{it}^0(\tau) \right) f_t^0(\tau) f_t^0(\tau)' \right) \hat{\phi}_i^\lambda(\tau) \\
&= \sqrt{\frac{T}{N}} H_i^0(\tau) \hat{\phi}_i^\lambda(\tau) \text{ as } T \rightarrow \infty
\end{aligned} \tag{A.19}$$

uniformly in i , where in the second equality, I used the Mean Value Theorem and Assumption 3.1, and in the last equality I, used Assumption 4.2. Hence, all previous results imply that for any $\tau \in (0, 1)$, as $N, T \rightarrow \infty$ and if $\sqrt{T}/N \rightarrow 0$, then expression (A.15) is given by

$$R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), \hat{F}(\tau)) = R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau)) + \sqrt{\frac{T}{N}} H_i^0(\tau) \hat{\phi}_i^\lambda(\tau) + o_p(T/\sqrt{N}) \tag{A.20}$$

uniformly in i . This last result completes the proof ■

Proof of Theorem 3. For proving part 1 of the theorem, apply Lemma 5 and use the next result

$$\begin{aligned}
\left\| R_{i,\tau}^\Lambda(\hat{\Lambda}(\tau), \hat{F}(\tau)) \right\| &\leq (NT)^{-1/2} \sum_{t=1}^T \left| \mathbb{1}\{y_{it} - \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) < 0\} - \tau \right| \cdot \left\| f_t^0(\tau) + \hat{\phi}_t^f(\tau) \right\| \\
&\leq (NT)^{-1/2} \sum_{t=1}^T \|f_t^0(\tau)\| + (NT)^{-1/2} \sum_{t=1}^T \left\| \hat{\phi}_t^f(\tau) \right\| \\
&= o_p(T/\sqrt{N}) + O_p(\sqrt{T}/N)
\end{aligned} \tag{A.21}$$

which follows from Assumption 3.4 and Theorem 2. Hence, if $\sqrt{T}/N \rightarrow 0$, then the above expression is $o_p(T/\sqrt{N})$, implying that

$$\sqrt{T}\hat{\phi}_i^\lambda(\tau) = -\sqrt{N}H_i^0(\tau)^{-1}R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau)) + o_p(T)$$

and by Slutsky's theorem

$$\sqrt{T}(\hat{\lambda}_i(\tau) - \lambda_i^0(\tau)) \xrightarrow{d} \mathcal{N}(0, \tau(1-\tau)H_i^0(\tau)^{-1}\Sigma_F^0(\tau)H_i^0(\tau)^{-1}) \quad (\text{A.22})$$

uniformly in i and for $N, T \rightarrow \infty$. The proof of part 2 is similar and is, therefore, omitted. The proof is complete. \blacksquare

A.5 Proof of Theorem 4

The limiting distribution of the estimated quantile common components $\hat{c}_{it}(\tau)$ of the QFA model (2) can be derived from Theorems 2 and 3. The proof utilizes an argument similar to the one considered in Bai (2003).

Proof. From the definitions of $\hat{c}_{it}(\tau)$ and $c_{it}^0(\tau)$, we have that for any $\tau \in (0, 1)$,

$$\begin{aligned} \hat{c}_{it}(\tau) - c_{it}^0(\tau) &= \hat{\lambda}_i(\tau)' \hat{f}_t(\tau) - \lambda_i^0(\tau)' f_t^0(\tau) \\ &= \hat{\phi}_i^\lambda(\tau)' \hat{\phi}_t^f(\tau) + f_t^0(\tau)' \hat{\phi}_i^\lambda(\tau) + \lambda_i^0(\tau)' \hat{\phi}_t^f(\tau) \end{aligned} \quad (\text{A.23})$$

Using the Cauchy-Schwarz inequality and Theorem 2, the first term on the right-hand side of the previous expression satisfies

$$\begin{aligned} \left| \hat{\phi}_i^\lambda(\tau)' \hat{\phi}_t^f(\tau) \right| &\leq \left\| \hat{\phi}_i^\lambda(\tau) \right\| \cdot \left\| \hat{\phi}_t^f(\tau) \right\| \\ &= O_p((NT)^{-1/2}) \end{aligned}$$

uniformly over i and t . Next, using this result and Theorem 3, expression (A.23) can be rewritten as follows

$$\begin{aligned} \hat{c}_{it}(\tau) - c_{it}^0(\tau) &= -\sqrt{\frac{N}{T}} f_t^0(\tau)' H_i^0(\tau)^{-1} R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau)) \\ &\quad - \sqrt{\frac{T}{N}} \lambda_i^0(\tau)' J_t^0(\tau)^{-1} R_{t,\tau}^F(\Lambda^0(\tau), F^0(\tau)) \\ &\quad + O_p((NT)^{-1/2}) + o_p(\max\{\sqrt{N}, \sqrt{T}\}) \end{aligned}$$

Let $\zeta_{it}^0(\tau) \equiv \sqrt{N}f_t^0(\tau)'H_i^0(\tau)^{-1}R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau))$. Then, by Assumptions 3.1, 3.4 and 4.2, $\zeta_{it}^0(\tau) \xrightarrow{d} \mathcal{N}(0, U_{it}^0(\tau))$ uniformly over i and t , where

$$U_{it}^0(\tau) \equiv \tau(1-\tau)f_t^0(\tau)'H_i^0(\tau)^{-1}\Sigma_F^0(\tau)H_i^0(\tau)^{-1}f_t^0(\tau), \quad \tau \in (0, 1)$$

Similarly, let $\vartheta_{it}^0(\tau) \equiv \sqrt{T}\lambda_i^0(\tau)'J_t^0(\tau)^{-1}R_{t,\tau}^F(\Lambda^0(\tau), F^0(\tau))$. Thus, by Assumptions 3.1, 3.5 and 4.1, $\vartheta_{it}^0(\tau) \xrightarrow{d} \mathcal{N}(0, W_{it}^0(\tau))$ uniformly over i and t , with

$$W_{it}^0(\tau) \equiv \tau(1-\tau)\lambda_i^0(\tau)'J_t^0(\tau)^{-1}\Sigma_\Lambda^0(\tau)J_t^0(\tau)^{-1}\lambda_i^0(\tau), \quad \tau \in (0, 1)$$

Thereby, for any $\tau \in (0, 1)$,

$$\hat{c}_{it}(\tau) - c_{it}^0(\tau) = -\left(\frac{\zeta_{it}^0(\tau)}{\sqrt{T}} + \frac{\vartheta_{it}^0(\tau)}{\sqrt{N}}\right) + o_p(\max\{\sqrt{N}, \sqrt{T}\}) \quad (\text{A.24})$$

as $N, T \rightarrow \infty$. Note that $\zeta_{it}^0(\tau)$ and $\vartheta_{it}^0(\tau)$ are asymptotically independent. This observation occurs because both random variables depend on the sub-gradients $R_{i,\tau}^\Lambda(\Lambda^0(\tau), F^0(\tau))$ and $R_{t,\tau}^F(\Lambda^0(\tau), F^0(\tau))$, correspondingly, which in turn are sums of martingale difference sequences across time-series and cross-sections. Thus, the previous result implies that $(\zeta_{it}^0(\tau), \vartheta_{it}^0(\tau))$ converges uniformly in i and t to a bivariate normal distribution as $N, T \rightarrow \infty$. Hence, for any $\tau \in (0, 1)$,

$$-\left(\frac{\zeta_{it}^0(\tau)}{\sqrt{T}} + \frac{\vartheta_{it}^0(\tau)}{\sqrt{N}}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{U_{it}^0(\tau)}{T} + \frac{W_{it}^0(\tau)}{N}\right)$$

which using expression (A.24) implies that

$$\left(\frac{U_{it}^0(\tau)}{T} + \frac{W_{it}^0(\tau)}{N}\right)^{-1/2} (\hat{c}_{it}(\tau) - c_{it}^0(\tau)) \xrightarrow{d} \mathcal{N}(0, 1) \quad (\text{A.25})$$

uniformly over i and t . This last result completes the proof. ■

References

- Ahn, S. C. and A. R. Horenstein (2013), “Eigenvalue Ratio Test for the Number of Factors”, *Econometrica* 81(3): 1203–1227.
- Algina, J. (1980), “A Note on Identification in the Oblique and Orthogonal Factor Analysis Models”, *Psychometrika* 45(3): 393–396.
- Alzate, C. and J. A. K. Suykens (2005), “Extending Kernel Principal Components Analysis to General Underlying Loss Functions”, in “Proceedings of the International Joint Conference on Neuronal Networks”, pages 214–219.
- Anderson, T. W. and H. Rubin (1956), “Statistical Inference in Factor Analysis”, in “Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability”, volume 5, pages 111–150, University of California Press.
- Aruoba, S. B., F. X. Diebold, and C. Scotti (2009), “Real-Time Measurement of Business Conditions”, *Journal of Business & Economic Statistics* 27(4): 417–427.
- Bahadur, R. R. (1966), “A Note on Quantiles in Large Samples”, *The Annals of Mathematical Statistics* 37(3): 577–580.
- Bai, J. (1996), “Testing for Parameter Constancy in Linear Regressions: An Empirical Distribution Function Approach”, *Econometrica* 64(3): 597–622.
- Bai, J. (2003), “Inferential Theory for Factor Models of Large Dimensions”, *Econometrica* 71(1): 135–171.
- Bai, J. (2009), “Panel Data Models with Interactive Fixed Effects”, *Econometrica* 77(4): 1229–1279.
- Bai, J. and K. Li (2012), “Statistical Analysis of Factor Models of High Dimension”, *The Annals of Statistics* 40(1): 436–465.
- Bai, J. and K. Li (2014), “Theory and Methods of Panel Data Models with Interactive Effects”, *The Annals of Statistics* 42(1): 142–170.
- Bai, J. and S. Ng (2002), “Determining the Number of Factors in Approximate Factor Models”, *Econometrica* 70(1): 191–221.
- Bai, J. and S. Ng (2008), “Large Dimensional Factor Analysis”, *Foundations and Trends in Econometrics* 3(2): 89–63.

- Bai, J. and S. Ng (2013), “Principal Components Estimation and Identification of Static Factors”, *Journal of Econometrics* 176(1): 18–29.
- Bekker, P. A. (1986), “A Note on the Identification of Restricted Factor Loading Matrices”, *Psychometrika* 51(4): 607–611.
- Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon (2012), “Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors”, *Journal of Financial Economics* 104(3): 535–559.
- Bofinger, E. (1975), “Estimation of a Density Function Using Order Statistics”, *Australian & New Zealand Journal of Statistics* 17(1): 1–7.
- Campbell, J. Y. and J. H. Cochrane (1999), “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior”, *Journal of Political Economy* 107(2): 205–251.
- Chen, M., I. Fernandez-Val, and M. Weidner (2014), “Nonlinear Panel Models with Interactive Effects”, ArXiv e-prints.
- Chernozhukov, V. (2005), “Extremal Quantile Regression”, *The Annals of Statistics* 33(2): 806–839.
- Chernozhukov, V. and I. Fernandez-Val (2011), “Inference for Extremal Conditional Quantile Models with an Application to Market and Birthweight Risks”, *The Review of Economic Studies* 78(2): 559–589.
- Chow, Y. S. and H. Teicher (1997), *Probability Theory: Independence, Interchangeability, Martingales*, Springer, third edition.
- Connor, G. and R. A. Korajczyk (1988), “Risk and Return in an Equilibrium APT: Application of a New Test Methodology”, *Journal of Financial Economics* 21(2): 255–289.
- Connor, G. and R. A. Korajczyk (1993), “A Test for the Number of Factors in an Approximate Factor Model”, *The Journal of Finance* 48(4): 1263–1291.
- Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein (2001), *Introduction to Algorithms*, The MIT Press, 2nd edition.
- Crucini, M. J., M. A. Kose, and C. Otrok (2011), “What Are the Driving Forces of International Business Cycles?”, *Review of Economic Dynamics* 14(1): 156–175.

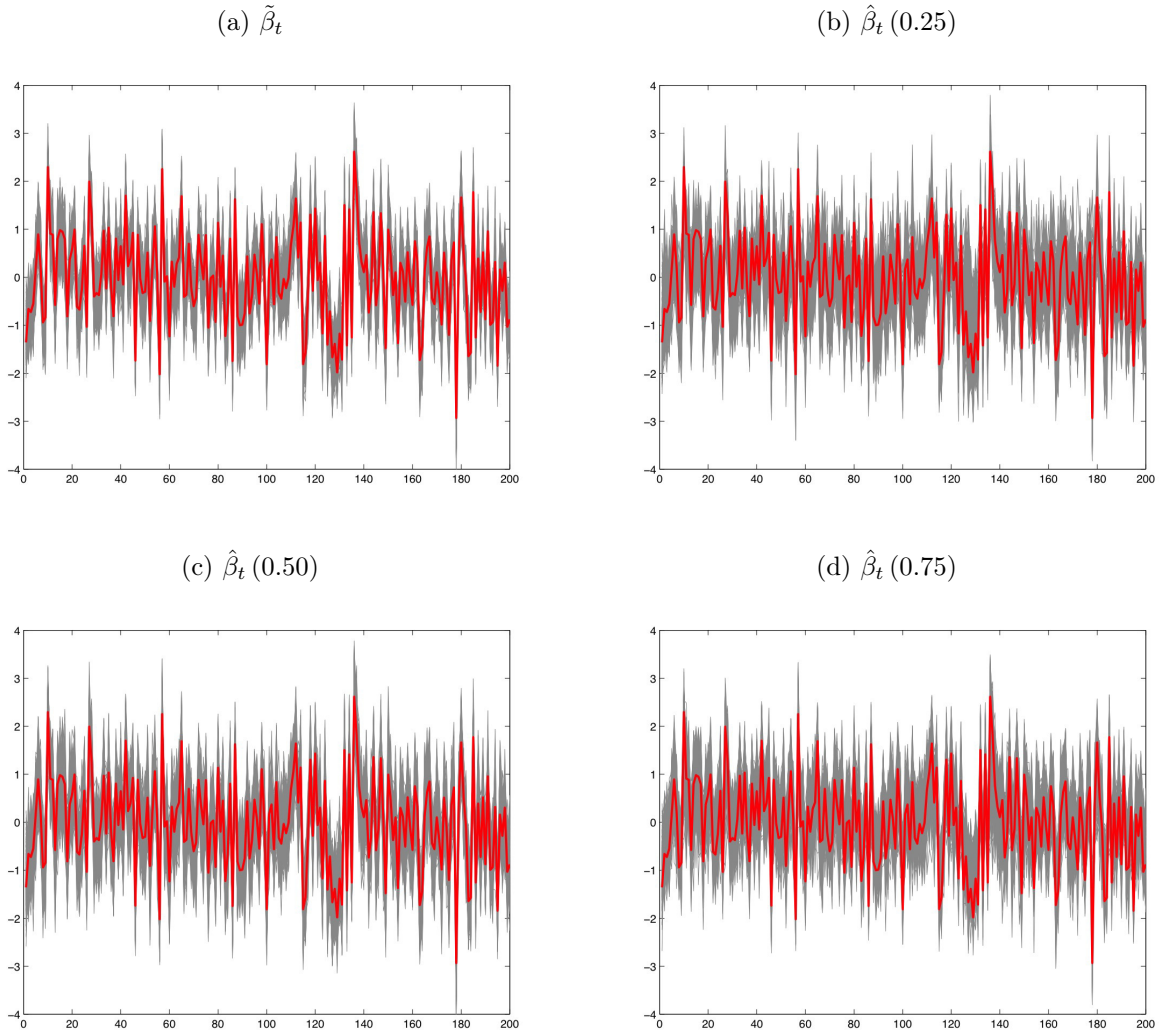
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), “Maximum Likelihood from Incomplete Data via the EM Algorithm”, *Journal of the Royal Statistical Society* 39(1): 1–38.
- Fernandez-Val, I. and M. Weidner (2016), “Individual and Time Effects in Nonlinear Panel Models with Large N, T”, *Journal of Econometrics* 192(1): 291–312.
- Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000), “The Generalized Dynamic Factor Model: Identification and Estimation”, *The Review of Economics and Statistics* 82(4): 540–554.
- Forni, M. and L. Reichlin (1998), “Let’s Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics”, *The Review of Economic Studies* 65(3): 453–473.
- Geweke, J. and G. Zhou (1996), “Measuring the Price of the Arbitrage Pricing Theory”, *The Review of Financial Studies* 9(2): 557–587.
- Gregory, A. W. and A. C. Head (1999), “Common and Country-Specific Fluctuations in Productivity, Investment and the Current Account”, *Journal of Monetary Economics* 44(3): 423–451.
- Gutenbrunner, C. and J. Jureckova (1992), “Regression Rank Scores and Regression Quantiles”, *The Annals of Statistics* 20(1): 305–330.
- Hall, P. and C. C. Heyde (1980), *Martingale Limit Theory and Its Applications*, Probability and Mathematical Statistics, Academic Press.
- Hall, P. and S. J. Sheather (1988), “On the Distribution of a Studentized Quantile”, *Journal of the Royal Statistical Society* 50(3): 381–391.
- Harding, M. and C. Lamarche (2014), “Estimating and Testing a Quantile Regression Model with Interactive Effects”, *Journal of Econometrics* 178(1): 101–113.
- Huber, P. J. and E. M. Ronchetti (2009), *Robust Statistics*, John Wiley & Sons, 2nd edition.
- Jolliffe, I. (2002), *Principal Component Analysis*, Springer, 2nd edition.
- Jurado, K., S. C. Ludvigson, and S. Ng (2015), “Measuring Uncertainty”, *American Economic Review* 105(3): 1177–1216.
- Kapetanios, G. (2010), “A Testing Procedure for Determining the Number of Factors in Approximate Factor Models with Large Datasets”, *Journal of Business & Economic Statistics* 28(3): 397–409.

- Karadimitropoulou, A. and M. Leon-Ledesma (2013), “World, Country and Sector Factors in International Business Cycles”, *Journal of Economic Dynamics & Control* 37(12): 2913–2927.
- Kato, K., A. F. Galvao, and G. V. Montes-Rojas (2012), “Asymptotics for Panel Quantile Regression Models with Individual Effects”, *Journal of Econometrics* 170(1): 76–91.
- Kiefer, J. (1967), “On Bahadur’s Representation of Sample Quantiles”, *The Annals of Mathematical Statistics* 38(5): 1323–1342.
- Knight, K. (1998), “Limiting Distributions for L1 Regression Estimators Under General Conditions”, *The Annals of Statistics* 26(2): 755–770.
- Koenker, R. (2004), “Quantile Regression for Longitudinal Data”, *Journal of Multivariate Analysis* 91(1): 74–89.
- Koenker, R. (2005), *Quantile Regression*, Cambridge University Press.
- Koenker, R. and G. Bassett (1978), “Regression Quantiles”, *Econometrica* 46(1): 33–50.
- Koenker, R. and J. A. F. Machado (1999), “Goodness of Fit and Related Inference Processes for Quantile Regression”, *Journal of the American Statistical Association* 94(448): 1296–1310.
- Koenker, R. and S. Portnoy (1987), “L-Estimation for Linear Models”, *Journal of the American Statistical Association* 82(399): 851–857.
- Koenker, R. and Z. Xiao (2002), “Inference on the Quantile Regression Process”, *Econometrica* 70(4): 1583–1612.
- Kritzman, M., Y. Li, S. Page, and R. Rigobon (2011), “Principal Components as a Measure of Systemic Risk”, *The Journal of Portfolio Management* 37(4): 112–126.
- Lim, Y. and H.-S. Oh (2016), “A Data-Adaptive Principal Component Analysis: Use of Composite Asymmetric Huber Function”, *Journal of Computational and Graphical Statistics* 25(4): 1230–1247.
- Lütkepohl, H. (1996), *Handbook of Matrices*, John Wiley & Sons.
- Mariano, R. S. and Y. Marasawa (2003), “A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series”, *Journal of Applied Econometrics* 18(4): 427–443.

- Moon, H. R. and M. Weidner (2015), “Linear Regression for Panel with Unknown Number of Factors and Interactive Fixed Effects”, *Econometrica* 83(4): 1543–1579.
- Moon, H. R. and M. Weidner (2017), “Dynamic Linear Panel Regression Models with Interactive Fixed Effects”, *Econometric Theory* 33(1): 158–195.
- Nocedal, J. and S. J. Wright (2006), *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering, Springer, 2nd edition.
- Oka, T. and Z. Qu (2011), “Estimating Structural Changes in Regression Quantiles”, *Journal of Econometrics* 162(2): 248–267.
- Onatski, A. (2009), “Testing Hypotheses About the Number of Factors in Large Factor Models”, *Econometrica* 77(5): 1447–1479.
- Pesaran, M. H. (2006), “Estimation and Inference iarge Heterogeneous Panels with a Multi-factor Error Structure”, *Econometrica* 74(4): 967–1012.
- Portnoy, S. and J. Jureckova (1999), “On Extreme Regression Quantiles”, *Extremes* 2(3): 227–243.
- Portnoy, S. and R. Koenker (1989), “Adaptive L-Estimation for Linear Models”, *The Annals of Statistics* 17(1): 362–381.
- Portnoy, S. and R. Koenker (1997), “The Gaussian Hare and the Laplacian Tortoise: Computability of Squared-Error versus Absolute-Error Estimators”, *Statistical Science* 12(4): 279–300.
- Qu, Z. (2008), “Testing for Structural Change in Regression Quantiles”, *Journal of Econometrics* 146(1): 170–184.
- Ross, S. A. (1976), “The Arbitrage Theory of Capital Asset Pricing”, *Journal of Economic Theory* 13(3): 341–360.
- Roweis, S. (1998), “EM Algorithms for PCA and SPCA”, in “Advances in Neural Information Processing Systems”, volume 10, pages 626–632.
- Rubin, D. B. and D. T. Thayer (1982), “EM Algorithms for ML Factor Analysis”, *Psychometrika* 47(1): 69–76.
- Sagner, A. (2020), “Measuring Systemic Risk: A Quantile Factor Analysis Approach”, Working Paper 874, Central Bank of Chile.

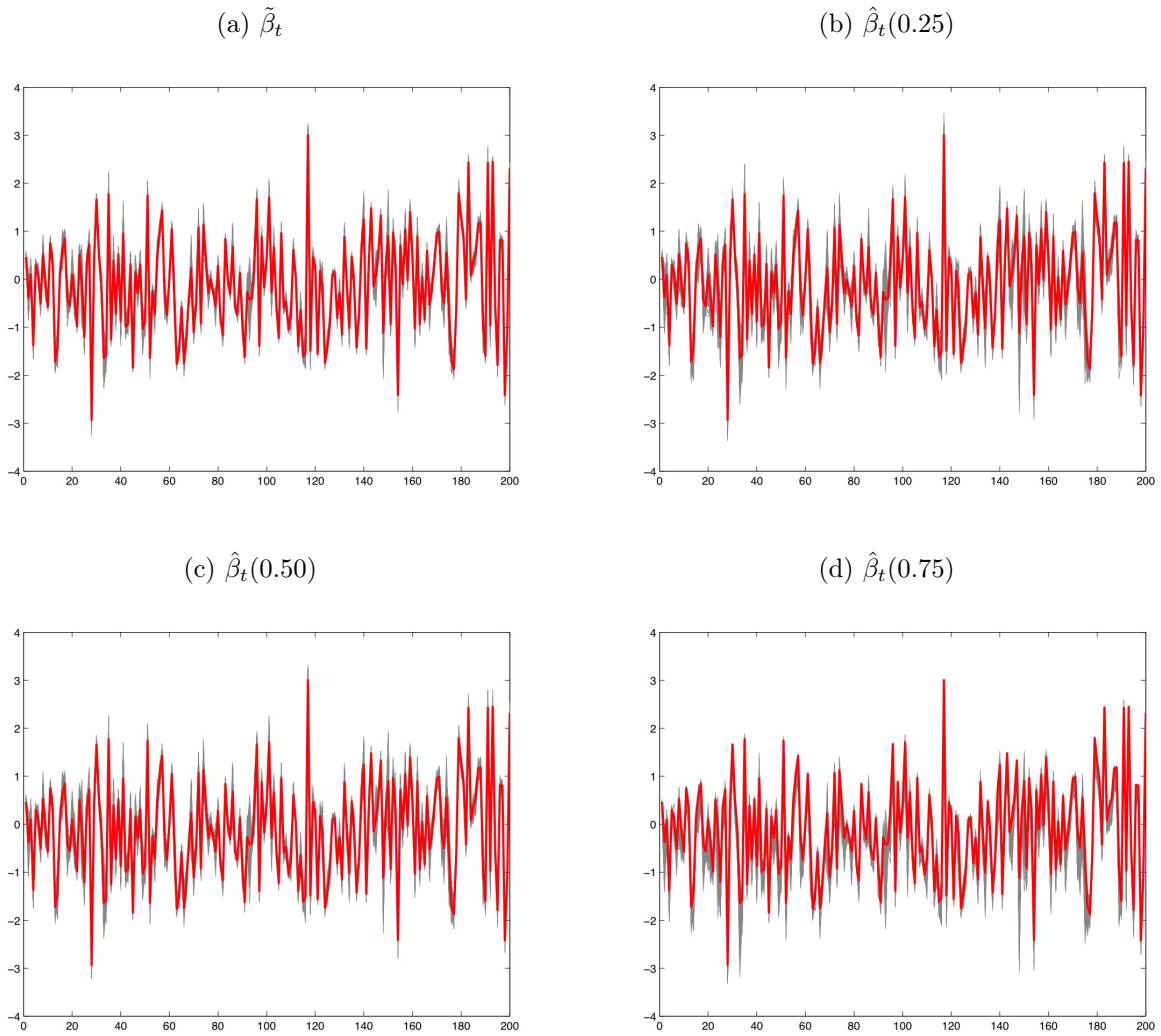
- Siddiqui, M. M. (1960), “Distribution of Quantiles in Samples from a Bivariate Population”, *Journal of Research of the National Bureau of Standards* 64(3): 145–150.
- Skrondal, A. and S. Rabe-Hesketh (2004), *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*, Chapman & Hall/CRC.
- Smith, R. L. (1994), “Nonregular Regression”, *Biometrika* 81(1): 173–183.
- Stock, J. H. and M. W. Watson (1989), “New Indexes of Coincident and Leading Economic Indicators”, in O. J. Blanchard and S. Fischer, editors, “NBER Macroeconomics Annual”, volume 4, pages 351–409, The MIT Press.
- Stock, J. H. and M. W. Watson (1999), “Forecasting Inflation”, *Journal of Monetary Economics* 44(2): 293–335.
- Stock, J. H. and M. W. Watson (2002), “Macroeconomic Forecasting Using Diffusion Indexes”, *Journal of Business & Economic Statistics* 20(2): 147–162.
- Tipping, M. E. and C. M. Bishop (1999), “Probabilistic Principal Components Analysis”, *Journal of the Royal Statistical Society* 61(3): 611–622.
- Tukey, J. W. (1965), “Which Part of the Sample Contains the Information?”, *Proceedings of the National Academy of Sciences of the United States of America* 53(1): 127–134.
- Wansbeek, T. and E. Meijer (2000), *Measurement Error and Latent Variables in Econometrics*, volume 37 of *Advanced Textbooks in Economics*, North-Holland.
- Zhou, H. and K. L. Lange (2010), “On the Bumpy Road to the Dominant Mode”, *Scandinavian Journal of Statistics* 37: 612–631.

Figure 1: Performance of QPC and PC Estimators of β_t^0 - DGP 1
 ($T = 200, N = 10$)



The red line corresponds to the simulated quantile factor β_t^0 . The grey shaded area corresponds to the QPC estimators $\hat{\beta}_t(\tau)$ for $\tau = \{0.25, 0.50, 0.75\}$ and the PC estimator $\tilde{\beta}_t$ that were computed from 1,000 simulations of a standard factor model (DGP 1).

Figure 2: Performance of QPC and PC Estimators of β_t^0 - DGP 2
 ($T = 200, N = 100$)

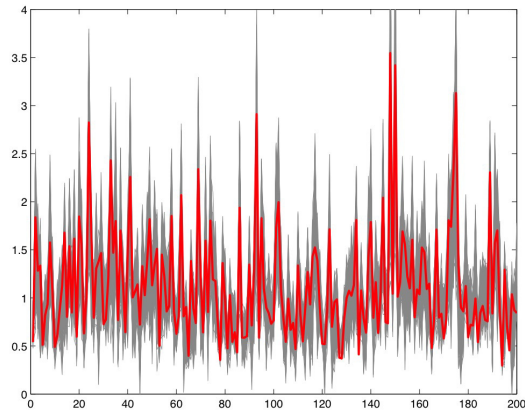


The red line corresponds to the simulated quantile factor β_t^0 . The grey shaded area corresponds to the QPC estimators $\hat{\beta}_t(\tau)$ for $\tau = \{0.25, 0.50, 0.75\}$ and the PC estimator $\tilde{\beta}_t$ that were computed from 1,000 simulations of a location-scale factor model (DGP 2).

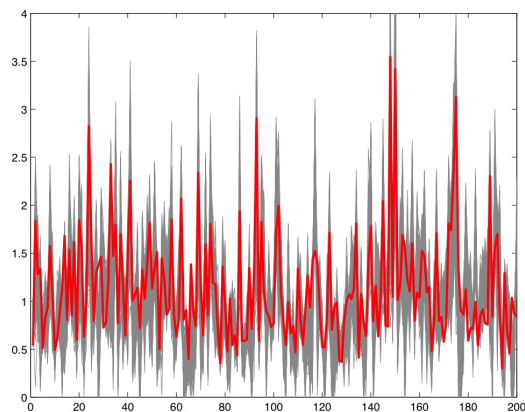
Figure 3: Performance of QPC Estimators of γ_t^0 - DGP 2

$(T = 200, N = 100)$

(a) $\hat{\gamma}_t(0.25)$

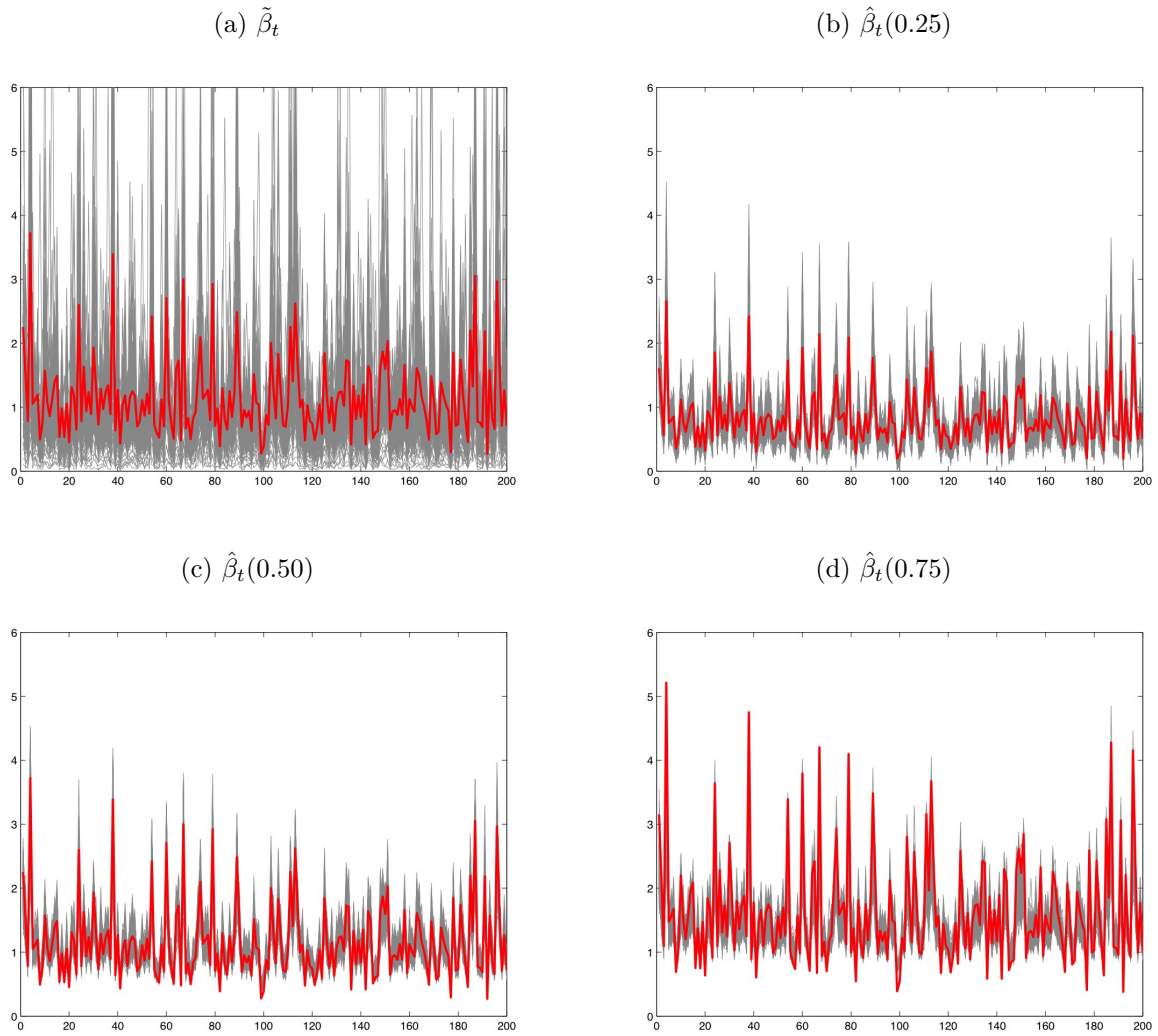


(b) $\hat{\gamma}_t(0.75)$



The red line corresponds to the simulated quantile factor γ_t^0 . The grey shaded area corresponds to the QPC estimators $\hat{\gamma}_t(\tau)$ for $\tau = \{0.25, 0.50, 0.75\}$ that were computed from 1,000 simulations of a location-scale factor model (DGP 2).

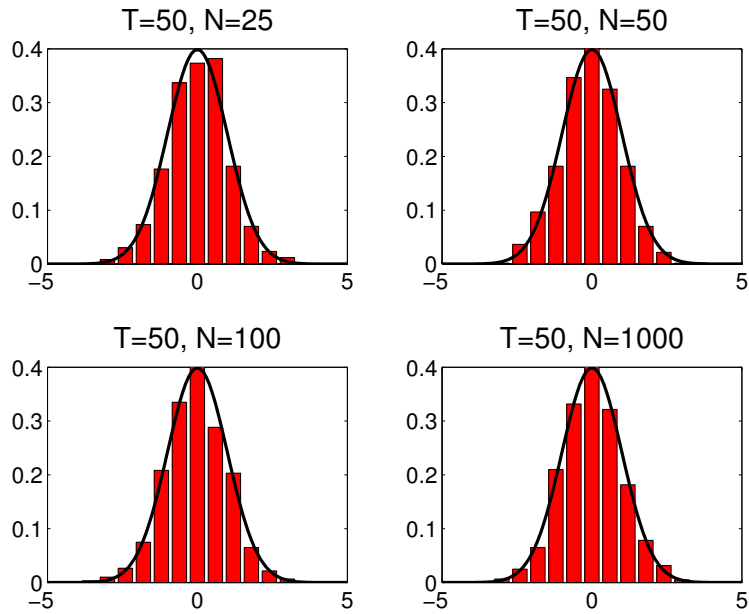
Figure 4: Performance of QPC and PC Estimators of β_t^0 - DGP 3
 ($T = 200, N = 100$)



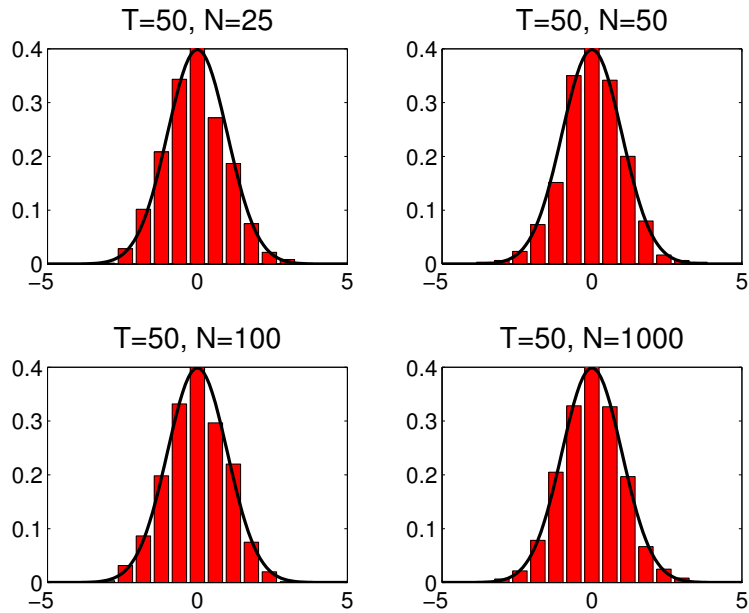
The red line corresponds to the simulated quantile factor β_t^0 . The grey shaded area corresponds to the QPC estimators $\hat{\beta}_t(\tau)$ for $\tau = \{0.25, 0.50, 0.75\}$ and the PC estimator $\tilde{\beta}_t$ that were computed from 1,000 simulations of a nonlinear factor model (DGP 3).

Figure 5: Histogram of Standardized QPC Factors ($T = 50$)

Panel A: $\tau = 0.25$

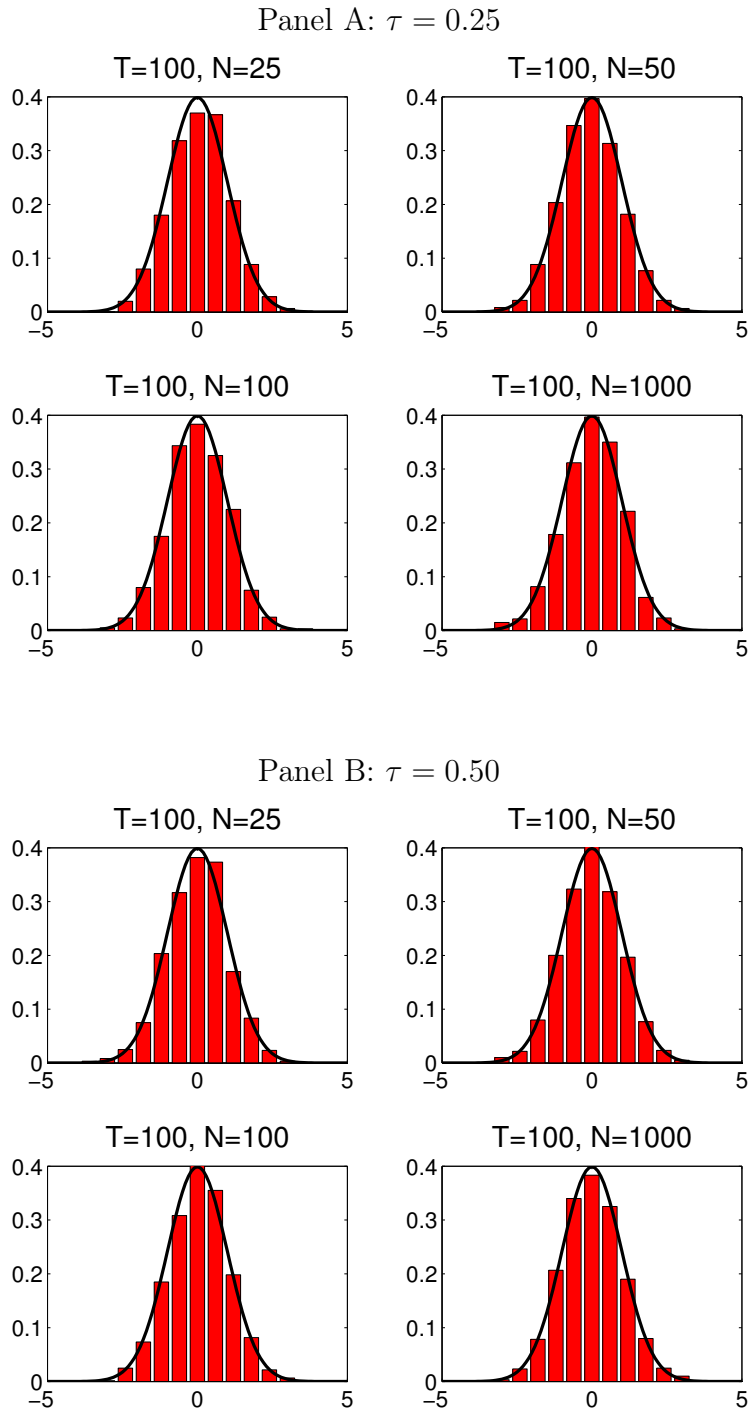


Panel B: $\tau = 0.50$



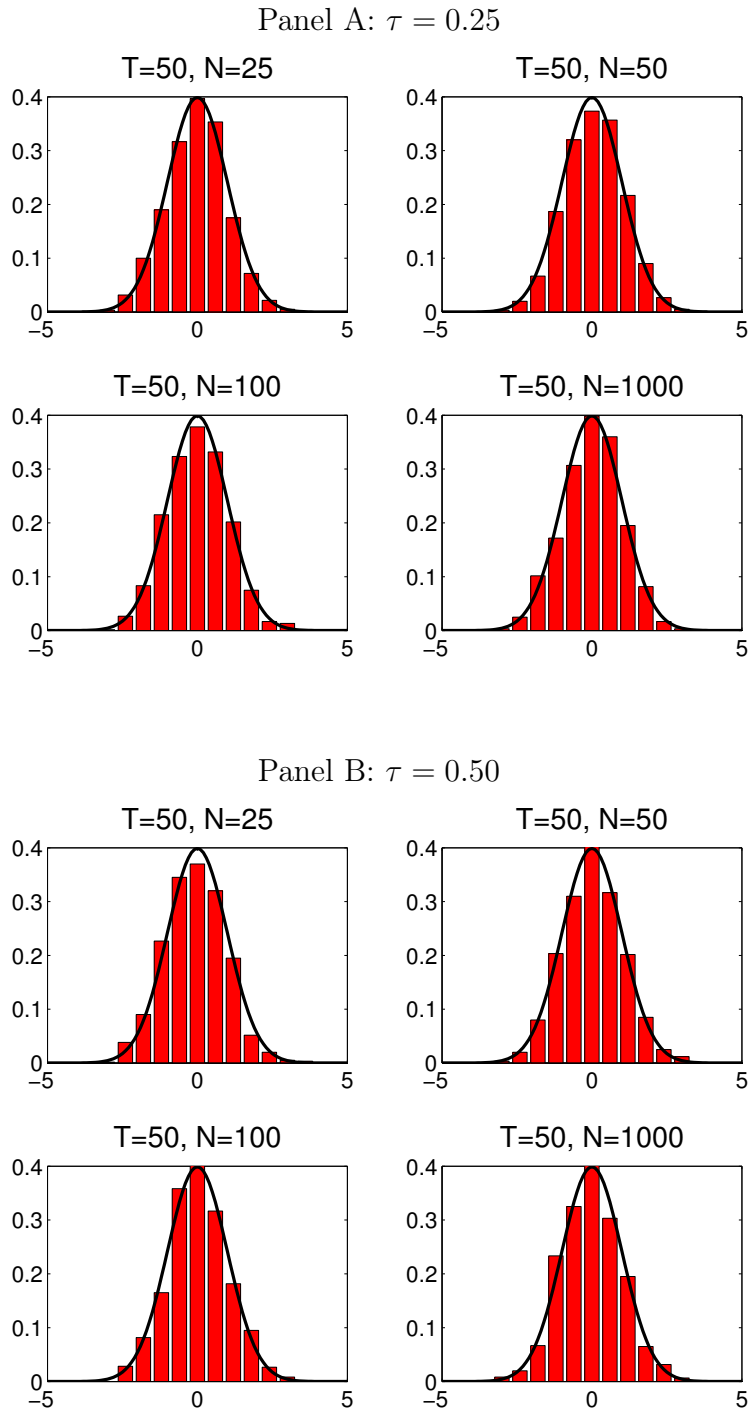
These graphs correspond to the histograms of $S = 1,000$ simulated standardized quantile factors $\tilde{f}_t(\tau)$ for $\tau = \{0.25, 0.50\}$ and $t = \lfloor T/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of x . The solid black line is the density of the standard Normal distribution.

Figure 6: Histogram of Standardized QPC Factors ($T = 100$)



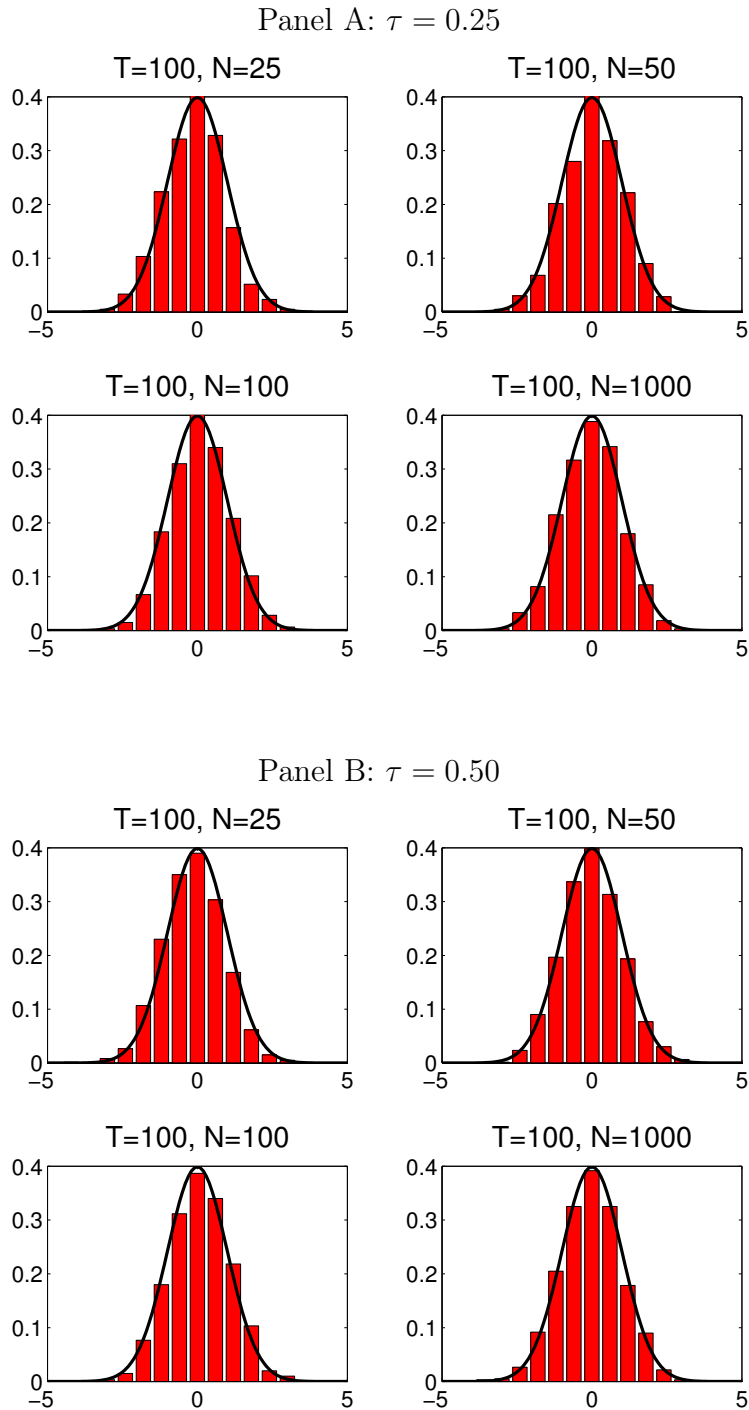
These graphs correspond to the histograms of $S = 1,000$ simulated standardized quantile factors $\tilde{f}_t(\tau)$ for $\tau = \{0.25, 0.50\}$ and $t = \lfloor T/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of x . The solid black line is the density of the standard Normal distribution.

Figure 7: Histogram of Standardized QPC Factor Loadings ($T = 50$)



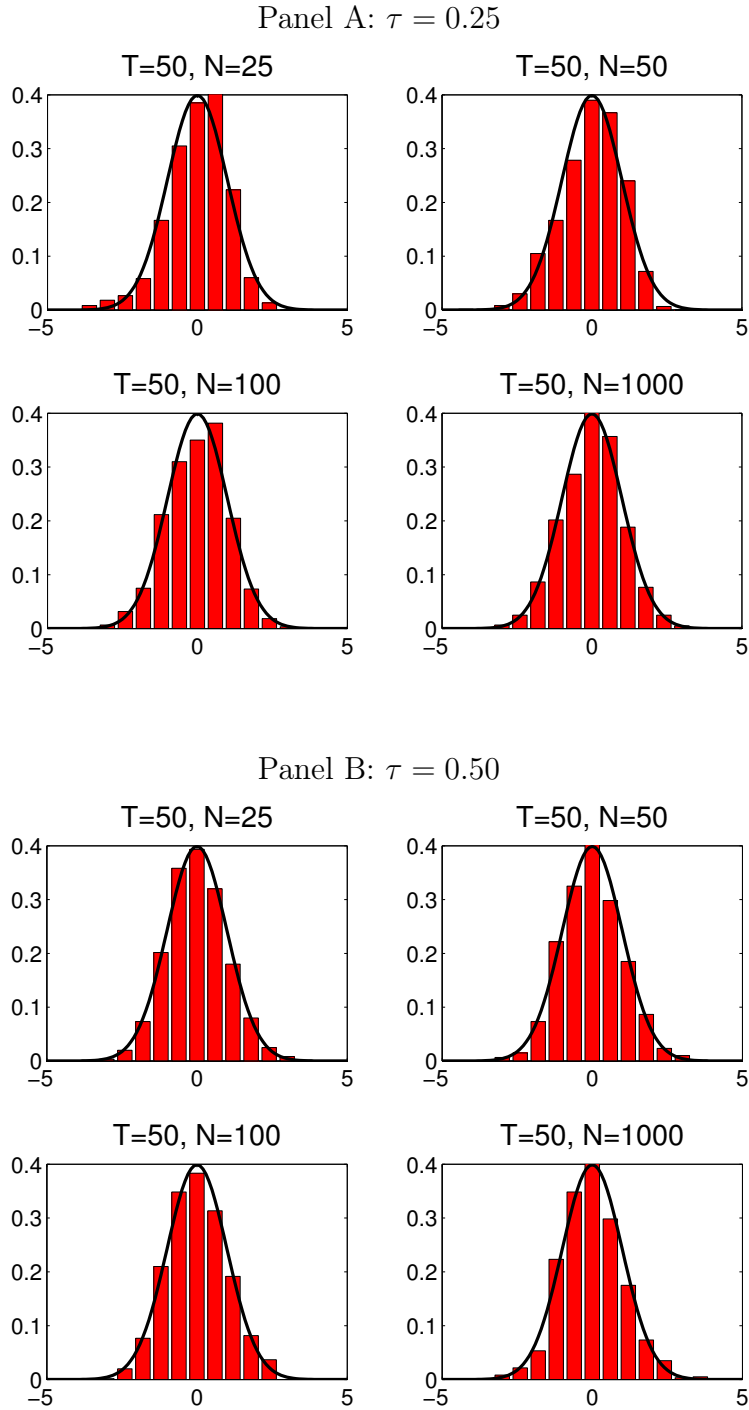
These graphs correspond to the histograms of $S = 1,000$ simulated standardized quantile factor loadings $\tilde{\lambda}_i(\tau)$ for $\tau = \{0.25, 0.50\}$ and $i = \lfloor N/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of x . The solid black line is the density of the standard Normal distribution.

Figure 8: Histogram of Standardized QPC Factor Loadings ($T = 100$)



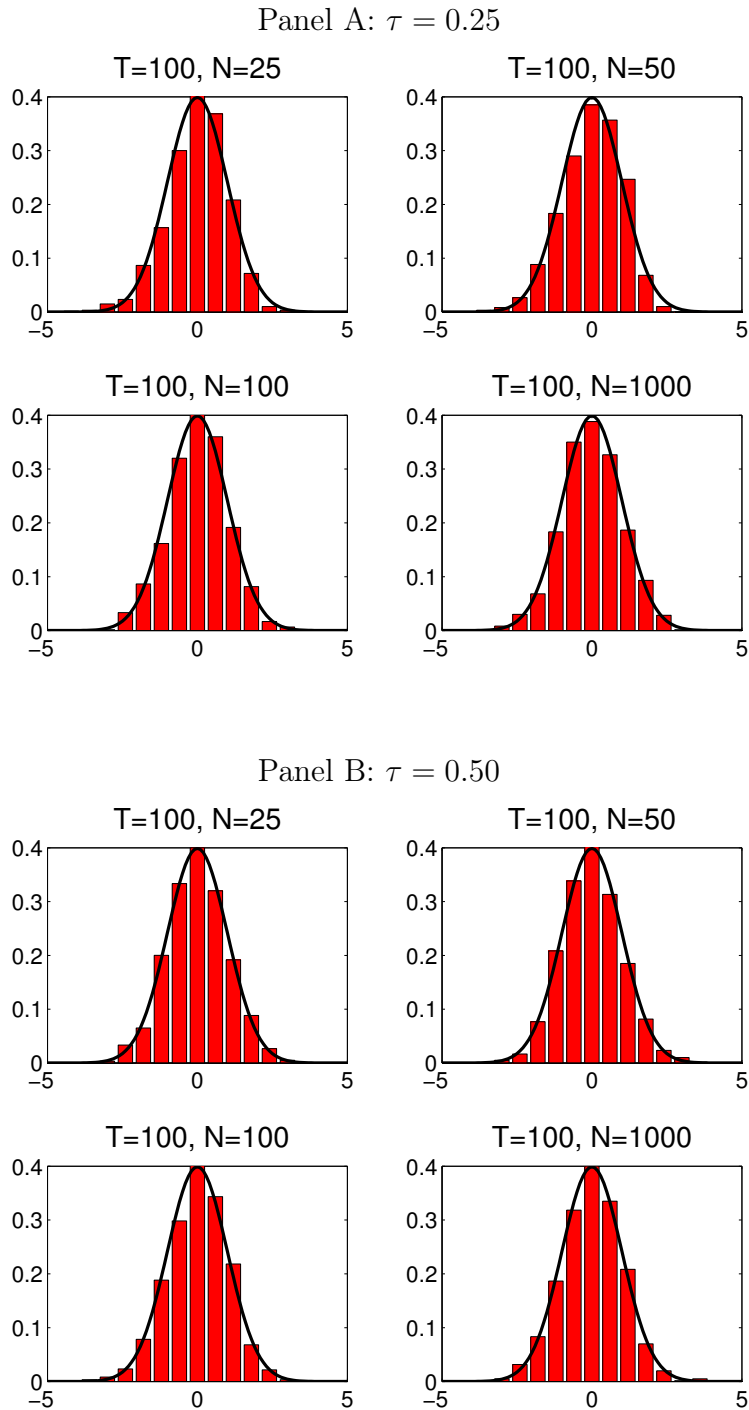
These graphs correspond to the histograms of $S = 1,000$ simulated standardized quantile factor loadings $\tilde{\lambda}_i(\tau)$ for $\tau = \{0.25, 0.50\}$ and $i = \lfloor N/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of x . The solid black line is the density of the standard Normal distribution.

Figure 9: Histogram of Standardized QPC Common Components ($T = 50$)



These graphs correspond to the histograms of $S = 1,000$ simulated standardized quantile common components $\tilde{c}_{it}(\tau)$ for $\tau = \{0.25, 0.50\}$ and $(i, t) = (\lfloor N/2 \rfloor, \lfloor T/2 \rfloor)$, where $\lfloor x \rfloor$ is the integer part of x . The solid black line is the density of the standard Normal distribution.

Figure 10: Histogram of Standardized QPC Common Components ($T = 100$)



These graphs correspond to the histograms of $S = 1,000$ simulated standardized quantile common components $\tilde{c}_{it}(\tau)$ for $\tau = \{0.25, 0.50\}$ and $(i, t) = (\lfloor N/2 \rfloor, \lfloor T/2 \rfloor)$, where $\lfloor x \rfloor$ is the integer part of x . The solid black line is the density of the standard Normal distribution.

Table 1: Correlation Between Estimated and True Factors and Factor Loadings - DGP 1

$T \setminus N$	PC			QPC ($\tau = 0.25$)			QPC ($\tau = 0.50$)			QPC ($\tau = 0.75$)		
	10	50	100	10	50	100	10	50	100	10	50	100
	Panel A: Factor											
50	0.9367	0.9901	0.9947	0.8794	0.9716	0.9896	0.8951	0.9844	0.9917	0.8768	0.9640	0.9824
100	0.9413	0.9908	0.9951	0.8795	0.9743	0.9844	0.9061	0.9856	0.9924	0.8894	0.9688	0.9884
200	0.9397	0.9904	0.9950	0.8682	0.9535	0.9897	0.9008	0.9851	0.9922	0.8746	0.9512	0.9800
1000	0.9438	0.9911	0.9953	0.8798	0.9654	0.9745	0.8983	0.9862	0.9927	0.8783	0.9666	0.9793
	Panel B: Factor Loading											
50	0.9846	0.9901	0.9895	0.9078	0.9762	0.9049	0.9608	0.9846	0.9837	0.9223	0.9734	0.9190
100	0.9930	0.9953	0.9951	0.9147	0.9897	0.9536	0.9778	0.9927	0.9925	0.9466	0.9847	0.9395
200	0.9964	0.9976	0.9975	0.9149	0.9938	0.9748	0.9807	0.9962	0.9961	0.9543	0.9892	0.9575
1000	0.9993	0.9996	0.9995	0.9539	0.9998	0.9845	0.9987	0.9992	0.9995	0.9896	0.9957	0.9825

Average correlation between the QPC estimators $\{\hat{\alpha}_i^{(s)}(\tau)\}_{i=1}^N$ and $\{\hat{\beta}_t^{(s)}(\tau)\}_{t=1}^T$, the PC estimators $\{\hat{\alpha}_i^{(s)}(\tau)\}_{i=1}^N$ and $\{\hat{\beta}_t^{(s)}(\tau)\}_{t=1}^T$, and their actual counterparts $\{\alpha_i^0\}_{i=1}^N$ and $\{\beta_t^0\}_{t=1}^T$, respectively, for $N = \{10, 50, 100\}$, $T = \{50, 100, 200, 1000\}$, $\tau = \{0.25, 0.50, 0.75\}$, and $s = 1, \dots, S = 1000$ simulations. The average correlation was computed as $\bar{\rho}_X = S^{-1} \sum_{s=1}^S \rho(\hat{X}^{(s)}, X^0)$, where $\hat{X}^{(s)}$ is an estimator (QPC or PC), and X^0 is its actual counterpart.

Table 2: Correlation Between Estimated and True Factors and Factor Loadings - DGP 2

$T \setminus N$	PC			QPC ($\tau = 0.25$)			QPC ($\tau = 0.50$)			QPC ($\tau = 0.75$)		
	10	50	100	10	50	100	10	50	100	10	50	100
Panel A: First Factor												
50	0.9418	0.9908	0.9947	0.9078	0.9725	0.9900	0.9147	0.9868	0.9928	0.9129	0.9751	0.9910
100	0.9411	0.9910	0.9951	0.8441	0.8937	0.9566	0.9175	0.9862	0.9928	0.8175	0.9280	0.9748
200	0.9421	0.9909	0.9952	0.8736	0.9677	0.9874	0.9164	0.9861	0.9927	0.8824	0.9462	0.9812
1000	0.9426	0.9909	0.9950	0.8640	0.9509	0.9861	0.9118	0.9859	0.9926	0.8612	0.9478	0.9845
Panel B: Second Factor												
50				0.4640	0.6617	0.8567				0.5641	0.8209	0.9088
100				0.5194	0.7112	0.8465				0.5479	0.7142	0.8739
200				0.5561	0.8114	0.8594				0.5513	0.7484	0.8321
1000				0.5666	0.7660	0.8999				0.5151	0.7827	0.8494

Average correlation between the QPC estimators $\{\hat{\beta}_t^{(s)}(\tau)\}_{t=1}^T$ and $\{\hat{\gamma}_t^{(s)}(\tau)\}_{t=1}^T$, the PC estimators $\{\hat{\beta}_t^{(s)}(\tau)\}_{t=1}^T$, and their actual counterparts $\{\beta_t^0\}_{t=1}^T$ and $\{\gamma_t^0\}_{t=1}^T$, respectively, for $N = \{10, 50, 100\}$, $T = \{50, 100, 200, 1000\}$, $\tau = \{0.25, 0.50, 0.75\}$, and $s = 1, \dots, S = 1000$ simulations. The average correlation was computed as $\bar{\rho}_X = S^{-1} \sum_{s=1}^S \rho(\hat{X}^{(s)}, X^0)$, where $\hat{X}^{(s)}$ is an estimator (QPC or PC), and X^0 is its actual counterpart.

Table 3: Correlation Between Estimated and True Factors and Factor Loadings - DGP 3

$T \setminus N$	PC			QPC ($\tau = 0.25$)			QPC ($\tau = 0.50$)			QPC ($\tau = 0.75$)		
	10	50	100	10	50	100	10	50	100	10	50	100
	Panel A: Factor											
50	0.6644	0.8034	0.8550	0.6847	0.9109	0.9535	0.7467	0.9262	0.9618	0.7016	0.9133	0.9556
100	0.6614	0.7992	0.8467	0.6745	0.9072	0.9514	0.7377	0.9209	0.9592	0.6949	0.9051	0.9519
200	0.6464	0.7890	0.8368	0.6573	0.8963	0.9452	0.7174	0.9112	0.9532	0.6721	0.8954	0.9455
1000	0.6675	0.8068	0.8520	0.6680	0.8997	0.9469	0.7221	0.9131	0.9545	0.6720	0.8973	0.9462
	Panel B: Factor Loading											
50	0.6733	0.8075	0.8203	0.7261	0.9142	0.9156	0.8337	0.9291	0.9286	0.8277	0.9180	0.9162
100	0.7647	0.8762	0.8936	0.7805	0.9560	0.9577	0.8939	0.9641	0.9642	0.8892	0.9589	0.9575
200	0.8327	0.9294	0.9381	0.8202	0.9786	0.9791	0.9197	0.9829	0.9823	0.9196	0.9799	0.9794
1000	0.9360	0.9771	0.9801	0.8507	0.9949	0.9956	0.9452	0.9960	0.9964	0.9448	0.9951	0.9956

Average correlation between the QPC estimators $\{\hat{\alpha}_i^{(s)}(\tau)\}_{i=1}^N$ and $\{\hat{\beta}_t^{(s)}(\tau)\}_{t=1}^T$, the PC estimators $\{\hat{\alpha}_i^{(s)}(\tau)\}_{i=1}^N$ and $\{\hat{\beta}_t^{(s)}(\tau)\}_{t=1}^T$, and their actual counterparts $\{\alpha_i^0\}_{i=1}^N$ and $\{\beta_t^0\}_{t=1}^T$, respectively, for $N = \{10, 50, 100\}$, $T = \{50, 100, 200, 1000\}$, $\tau = \{0.25, 0.50, 0.75\}$, and $s = 1, \dots, S = 1000$ simulations. The average correlation was computed as $\bar{\rho}_X = S^{-1} \sum_{s=1}^S \rho(\hat{X}^{(s)}, X^0)$, where $\hat{X}^{(s)}$ is an estimator (QPC or PC), and X^0 is its actual counterpart.

Table 4: Mean of Standardized QPC Estimators

$T \setminus N$	$\tau = 0.25$			$\tau = 0.50$			$\tau = 0.75$					
	25	50	100	25	50	100	25	50	100			
	Panel A: Quantile Factor											
50	0.158	0.031	0.028	0.016	-0.028	0.025	-0.018	0.008	-0.143	0.031	-0.020	-0.015
100	0.088	-0.011	0.009	0.004	-0.011	0.006	-0.002	-0.002	-0.114	-0.019	0.010	-0.005
	Panel B: Quantile Factor Loading											
50	0.079	0.023	-0.037	0.031	-0.043	0.031	0.034	0.036	-0.083	-0.038	0.056	0.028
100	0.047	0.012	-0.013	0.016	0.016	0.009	0.012	-0.009	0.039	-0.024	-0.028	0.013
	Panel C: Quantile Common Component											
50	0.111	0.029	-0.021	0.024	0.039	-0.026	0.018	-0.014	0.116	-0.035	-0.024	0.024
100	0.073	-0.012	0.011	0.020	-0.013	0.008	0.005	0.006	-0.072	0.019	-0.012	-0.009

Sample mean of the standardized QPC estimator of the quantile factor $\tilde{f}_t(\tau)$, quantile factor loading $\tilde{\lambda}_i(\tau)$ and quantile common component $\tilde{c}_{it}(\tau)$ for $\tau \in \{0.25, 0.50, 0.75\}$, and $i = \lfloor N/2 \rfloor$ and $t = \lfloor T/2 \rfloor$, where $[x]$ is the integer part x . The sample mean was computed across simulations as $\bar{X}_{it}(\tau) = S^{-1} \sum_{s=1}^S X_{it}^{(s)}(\tau)$, where $X_{it}^{(s)}(\tau) = \{\tilde{\lambda}_i^{(s)}(\tau), \tilde{f}_t^{(s)}(\tau), \tilde{c}_{it}^{(s)}(\tau)\}$ and $S = 1,000$.

Table 5: Standard Deviation of Standardized QPC Estimators

$T \setminus N$	$\tau = 0.25$			$\tau = 0.50$			$\tau = 0.75$					
	25	50	100	25	50	100	25	50	100			
	Panel A: Quantile Factor											
50	1.380	1.221	1.258	1.131	1.194	1.140	1.113	1.085	1.321	1.226	1.174	1.098
100	1.207	1.177	1.159	1.078	1.150	1.110	1.079	1.029	1.196	1.176	1.130	1.072
	Panel B: Quantile Factor Loading											
50	1.490	1.344	1.224	1.142	1.259	1.206	1.156	1.093	1.513	1.325	1.203	1.159
100	1.232	1.205	1.139	1.090	1.185	1.146	1.076	1.023	1.807	1.188	1.099	1.091
	Panel C: Quantile Common Component											
50	1.433	1.231	1.217	1.136	1.199	1.193	1.145	1.083	1.454	1.277	1.189	1.110
100	1.221	1.198	1.150	1.087	1.158	1.137	1.078	1.029	1.202	1.178	1.097	1.080

Standard deviation of the standardized QPC estimator of the quantile factor $\tilde{f}_t(\tau)$, quantile factor loading $\tilde{\lambda}_i(\tau)$ and the quantile common component $\tilde{c}_{it}(\tau)$ for $\tau \in \{0.25, 0.50, 0.75\}$, and $i = \lfloor N/2 \rfloor$ and $t = \lfloor T/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of x . The sample standard deviation was computed across simulations as $\sigma_{it}^{\tilde{X}}(\tau) = \sqrt{S^{-1} \sum_{s=1}^S \left(X_{it}^{(s)}(\tau) - \bar{X}_{it}(\tau) \right)^2}$, where $X_{it}^{(s)}(\tau) = \{ \tilde{\lambda}_i^{(s)}(\tau), \tilde{f}_t^{(s)}(\tau), \tilde{c}_{it}^{(s)}(\tau) \}$, $\bar{X}_{it}(\tau)$ is the sample mean, and $S = 1,000$.

<p align="center">Documentos de Trabajo Banco Central de Chile</p>	<p align="center">Working Papers Central Bank of Chile</p>
<p align="center">NÚMEROS ANTERIORES</p>	<p align="center">PAST ISSUES</p>
<p>La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica: www.bcentral.cl/esp/estpub/estudios/dtbc.</p>	<p>Working Papers in PDF format can be downloaded free of charge from: www.bcentral.cl/eng/stdpub/studies/workingpaper.</p>
<p>Existe la posibilidad de solicitar una copia impresa con un costo de Ch\$500 si es dentro de Chile y US\$12 si es fuera de Chile. Las solicitudes se pueden hacer por fax: +56 2 26702231 o a través del correo electrónico: bcch@bcentral.cl.</p>	<p>Printed versions can be ordered individually for US\$12 per copy (for order inside Chile the charge is Ch\$500.) Orders can be placed by fax: +56 2 26702231 or by email: bcch@bcentral.cl.</p>

DTBC – 885

Heterogeneous Paths of Industrialization

Federico Huneus, Richard Rogerson

DTBC – 884

Does the Commodity Super Cycle Matter?

Andrés Fernández, Stephanie Schmitt-Grohé, Martín Uribe

DTBC – 883

Twitter-Based Economic Policy Uncertainty Index for Chile

Andrés Sagner, Juan Sebastián Becerra

DTBC – 882

Corporate-Sector Functional Currency: An International Comparison

Jorge Fernández, Fernando Pino, Francisco Vásquez

DTBC – 881

Back testing fan charts of activity and inflation: the Chilean case

Jorge Fornero, Andrés Gatty

DTBC – 880

Financing Firms in Hibernation during the COVID-19 Pandemic

Tatiana Didier, Federico Huneus, Mauricio Larrain, Sergio L. Schmukler

DTBC – 879

Choice Aversion in Directed Networks

Jorge Lorca, Emerson Melo

DTBC – 878

Big G

Lydia Cox, Gernot Muller, Ernesto Pasten, Raphael Schoenle, Michael Weber

DTBC – 877

Sticky Capital Controls

Miguel Acosta-Henao, Laura Alfaro, Andrés Fernández

DTBC – 876

Measuring the perceived value of an MBA degree

Carlos Madeira

DTBC – 875

The Real Effects of Monetary Shocks: Evidence from Micro Pricing Moments

Gee Hee, Matthew Klepacz, Ernesto Pasten, Raphael Schoenle

DTBC – 874

Measuring Systemic Risk: A Quantile Factor Analysis

Andrés Sagner

DTBC – 873

The impact of information laws on consumer credit access: evidence from Chile

Carlos Madeira

DTBC – 872

On the Response of Inflation and Monetary Policy to an Immigration Shock

Benjamín García, Juan Guerra-Salas

DTBC – 871

Proyecciones de corto plazo para el PIB trimestral: Desempeño reciente de una serie de modelos estándar

Marcus Cobb, Jennifer Peña



BANCO CENTRAL
DE CHILE

DOCUMENTOS DE TRABAJO • Agosto 2020