2027 AEA Annual Meeting

UCONN
UNIVERSITY OF CONNECTICUT

Bootstrap Inference for Factor-Augmented Quantile Regression

Department of Economics, University of Connecticut

Feifan Wang

FAQR Model and Research Question

How can we use information from a high-dimensional predictor panel X_t in time-series quantile regression?

High-Dimensional Panel $X_t = (X_{1t}, \dots, X_{Nt})'$ → **Factor Model** $X_{it} = \lambda_i' F_t + e_{it}$ → **Dimension Reduction** PCA → Estimated factors $\tilde{F}_t$

Factor-augmented quantile regression (FAQR)

$$y_{t+h} = \underbrace{\left(\alpha'(\tau)H^{-1} \quad \beta'(\tau)\right)}_{=\delta'(\tau)} \underbrace{\begin{pmatrix}\tilde{F}_t \\ W_t\end{pmatrix}}_{=\hat{z}_t} + \underbrace{\alpha'(\tau)H^{-1}\left(HF_t - \tilde{F}_t\right)}_{\text{factor-estimation error}} + \varepsilon_{t+h}(\tau),$$

$$q_\tau(\varepsilon_{t+h}(\tau) \mid F_t, W_t) = 0, \qquad t = 1, \dots, T-h.$$

➢ **Goal: Develop valid inference for FAQR in time-series settings by accounting for factor-estimation uncertainty and serial dependence in regression disturbances.**

Why Is Inference for FAQR Challenging?

➢ Challenge 1: Factor-estimation uncertainty

- Existing FAQR theory (Bai and Ng, 2008; Ando and Tsay, 2011) requires

$$\frac{T^{5/8}}{N} \to 0,$$

under which factor-estimation error is asymptotically negligible.

- When the cross-sectional dimension N is not sufficiently large relative to the time-series dimension T, factor-estimation uncertainty can affect the limiting distribution and therefore inference.

➢ Challenge 2: Serial dependence

- Existing HAC and bootstrap methods account for serial dependence in time-series quantile regression (Galvao and Yoon, 2024; Gregory et al., 2018; Cai and Long, 2026).
- However, these methods assume **observed regressors** and do not account for factor-estimation uncertainty in FAQR.

Contribution I | New Asymptotic Theory

We derive the limiting distribution of the smoothed FAQR estimator $\hat{\delta}_{m_{NT}}(\tau)$ under the rate condition

$$\frac{\sqrt{T}}{N} \to c \in [0, \infty).$$

$$\sqrt{T}\left(\hat{\delta}_{m_{NT}}(\tau) - \delta_0(\tau)\right) \xrightarrow{d} N\left(-c\,\kappa(\tau), \Sigma(\tau)\right),$$

- $\delta_0(\tau)$ is the true population parameter.
- m_{NT} denotes the smoothing bandwidth.
- If $c = 0$, factor-estimation error is asymptotically negligible,

$$\sqrt{T}\left(\hat{\delta}_{m_{NT}}(\tau) - \delta_0(\tau)\right) \xrightarrow{d} N\left(0, \Sigma(\tau)\right).$$

- If $c > 0$, factor-estimation error enters at first order through the non-negligible bias $-c\,\kappa(\tau)$.

Contribution II | New Bootstrap Method: The FAQR-DWB Procedure

➢ Bootstrap Setup

- Consider the scalar restriction $\mathcal{H}_0: R\delta(\tau) = r_0$, where R selects the coefficient of interest and r_0 is its hypothesized value.

Under $\mathcal{H}_0$, obtain the restricted smoothed FAQR estimator $\hat{\delta}^r_{m_{NT}}(\tau) = \left(\hat{\alpha}^{r\prime}_{m_{NT}}(\tau), \hat{\beta}^{r\prime}_{m_{NT}}(\tau)\right)'$ and define the restricted smoothed score

$$s_{m_{NT}}\left(\hat{z}_t; \hat{\delta}^r_{m_{NT}}(\tau)\right) = \hat{z}_t\left[K\left(-\frac{y_{t+h} - \hat{z}_t'\hat{\delta}^r_{m_{NT}}(\tau)}{m_{NT}}\right) - \tau\right]$$

- Let $\hat{D}^r_{m_{NT}}(\tau)$ denote the corresponding Jacobian and construct the projected score

$$g^r_{m_{NT},t}(\tau) = -R\left[\hat{D}^r_{m_{NT}}(\tau)\right]^{-1} s_{m_{NT}}\left(\hat{z}_t; \hat{\delta}^r_{m_{NT}}(\tau)\right).$$

Define the numerator $N^r_{m_{NT}}(\tau) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T-h} g^r_{m_{NT},t}(\tau)$.

- **The observed studentized test statistic** is

$$T^r_{\text{obs}}(\tau) = \frac{N^r_{m_{NT}}(\tau)}{\left[\mathcal{L}^B_{T-h}\left(\{g^r_{m_{NT},t}(\tau)\}; L_T\right)\right]^{1/2}}$$

where $\mathcal{L}^B_{T-h}(\cdot; L_T)$is the Bartlett long-run variance estimator.

➢ Bootstrap Procedure

1 Reproduce Factor-Estimation Uncertainty

1. Generate a bootstrap panel using a **wild bootstrap**:

$$X^*_t = \tilde{\Lambda}\tilde{F}_t + e^*_t, \qquad e^*_{it} = \hat{e}_{it}\eta^*_{it},$$
$$\hat{e}_{it} = X_{it} - \tilde{\lambda}_i'\tilde{F}_t, \qquad \eta^*_{it} \overset{i.i.d.}{\sim} N(0,1).$$

2. Extract bootstrap factors $\{\tilde{F}^*_t\}_{t=1}^T$ by PCA from X^*. Define $\hat{z}^*_t = \left(\tilde{F}^{*\prime}_t, W_t'\right)'$.
3. Construct **the factor-estimation component**:

$$N^*_{F,m_{NT}}(\tau) = -\frac{1}{\sqrt{T}}\sum_{t=1}^{T-h} R\left[\hat{D}^r_{m_{NT}}(\tau)\right]^{-1} \frac{1}{m_{NT}} k\left(-\frac{y_{t+h} - \hat{z}_t'\hat{\delta}^r_{m_{NT}}(\tau)}{m_{NT}}\right)\hat{z}^*_t \times\left(\tilde{F}^*_t - H^*\tilde{F}_t\right)'(H^{*-1})'\hat{\alpha}^r_{m_{NT}}(\tau).$$

2 Reproduce Serial Dependence in the Quantile Score (Following Cai and Long, 2026)

4. Independently of the first-stage multipliers η^*_{it}, generate $\zeta^*_t \overset{i.i.d.}{\sim} N(0,1), t = 1, \dots, T-h+L_T-1$, and construct

$$\xi^*_t = \frac{1}{\sqrt{L_T}}\sum_{j=0}^{L_T-1}\zeta^*_{t+j} - \frac{1}{T}\sum_{s=1}^{T-h}\left(\frac{1}{\sqrt{L_T}}\sum_{j=0}^{L_T-1}\zeta^*_{s+j}\right), \quad t = 1, \dots, T-h.$$

5. Apply ξ^*_t to the recentered restricted smoothed score:

$$g^*_{\text{DWB},m_{NT},t}(\tau) = -\xi^*_t R\left[\hat{D}^r_{m_{NT}}(\tau)\right]^{-1}\tilde{s}_{m_{NT}}\left(\hat{z}_t; \hat{\delta}^r_{m_{NT}}(\tau)\right).$$

6. Construct **the DWB score component**:

$$N^*_{\text{DWB},m_{NT}}(\tau) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T-h} g^*_{\text{DWB},m_{NT},t}(\tau).$$

3 Combine the Two Sources of Uncertainty

7. Combine the two bootstrap components:

$$N^*_{m_{NT}}(\tau) = N^*_{\text{DWB},m_{NT}}(\tau) + N^*_{F,m_{NT}}(\tau)$$

8. Construct **the studentized bootstrap statistic**:

$$T^*(\tau) = \frac{N^*_{m_{NT}}(\tau)}{\left[\mathcal{L}^B_{T-h}\left(\{g^*_{\text{DWB},m_{NT},t}(\tau)\}; L_T\right)\right]^{1/2}}.$$

where $\mathcal{L}^B_{T-h}(\cdot; L_T)$ denotes the Bartlett long-run variance estimator.

9. Compute the bootstrap p-value:

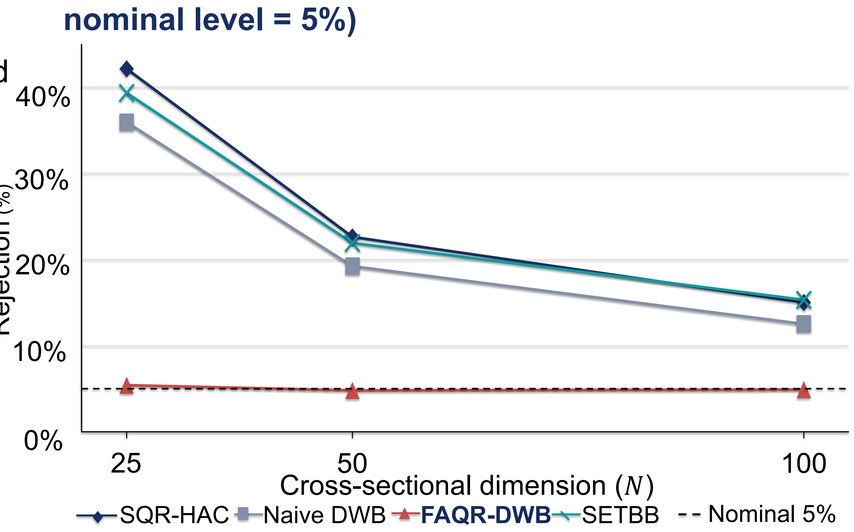
$$\hat{p}_{\text{stud}}(\tau) = \frac{1}{B}\sum_{b=1}^{B}\mathbf{1}\left\{|T^*_b(\tau)| \geq |T^r_{\text{obs}}(\tau)|\right\}.$$

Reject $\mathcal{H}_0$ at level α if

$$\hat{p}_{\text{stud}}(\tau) < \alpha.$$

Monte Carlo Evidence

Figure: Empirical Rejection Rates (τ = 0.10, T = 600; nominal level = 5%)



- ➢ Ignoring factor-estimation uncertainty leads to substantial over-rejection, especially when N is small relative to T.
- ➢ The **studentized FAQR-DWB** rejection rates remain **close to the nominal 5% level** across different cross-sectional dimensions.

Empirical Illustration: U.S. GDP Growth

- ➢ Data: FRED-QD quarterly dataset with 211 U.S. macroeconomic variables from 1960Q1-2019Q2 ($N = 211,\ T = 238$)
- ➢ Forecast target: one-quarter-ahead U.S. real GDP growth.
- ➢ Model: Extract two latent factors from the macroeconomic panel using PCA and estimate

$$q_\tau(y_{t+1} \mid z_t) = \beta_0(\tau) + \beta_1(\tau)y_t + \alpha_1(\tau)F_{1t} + \alpha_2(\tau)F_{2t}.$$

- ➢ Rolling estimation: Use a rolling window of 160 quarters.

Table: Average Length of 95% Confidence Intervals (τ = 0.10)

Method	F_{1t}	F_{2t}
SQR-HAC	3.113	1.568
FAQR-DWB	**3.445**	**2.006**
SETBB	2.187	1.751

- ➢ **FAQR-DWB** produces **wider confidence intervals** than other methods, reflecting additional uncertainty from factor estimation that is not captured when inference conditions on the estimated factors.

Conclusions

- **New Asymptotic Theory:** Show that factor estimation introduces **a first-order, quantile-specific bias** when $c > 0$.
- **New Bootstrap Method:** Develop **the FAQR-DWB** procedure to account for factor-estimation uncertainty and serial dependence and establish its validity.
- **Monte Carlo Simulations:** Show **improved size control**, especially when N is small relative to T.
- **Empirical Application:** Show that factor-estimation uncertainty affects inference on downside risks in GDP growth.
- **Future Research:** Extend to conditional quantile forecasting.

Contact Information

Academic Website: https://sites.google.com/view/feifanwang

Email: feifan.wang@uconn.edu