QFM Exploration Notes

April 2, 2025

1 Location-Scale Model Notes

I consider the following DGP:

$$y_{it} = \alpha_i \beta_t + \eta_i \gamma_t e_{it}$$

We require that $\eta_t \gamma_t > 0$, so we will draw them from the χ^2 distribution. All together we have:

$$\alpha_i \sim \mathcal{N}(0,1); \quad \beta_t \sim \mathcal{N}(0,1); \quad \eta_i \sim \chi^2(1); \quad \gamma_t \sim \chi^2(1); \quad e_{it} \sim \mathcal{N}(0,1)$$

Suppose y_{it} is a scalar random variable, $\alpha_i, \beta_t, \gamma_t$, and η_i are random vectors, and e_{it} are i.i.d., in time and cross-section, and independent of y, α, β, γ, and η_i. Let F be the cdf of e_{it}, and $Q(\tau)$ its τ-th quantile. Suppose the DGP is

$$y_{it} = \alpha_i' \beta_t + \eta_i' \gamma_t e_{it}.$$

Assume $\eta_i' \gamma_t$ is always positive, then the conditional quantile function is

$$Q_\tau(y_{it} | \alpha_i, \beta_t, \eta_i' \gamma_t) = \alpha_i' \beta_t + \eta_i' \gamma_t F^{-1}(\tau) = \alpha_i' \beta_t + \eta_i' \gamma_t Q(\tau)$$

This is what the QPC algorithm estimates, and so with a two-factor model the factors estimated should*** span the same space as β_t and γ_t. Then a regression the true γ_t on $\hat{\beta}_t + \hat{\gamma}_t$ should have a high R^2. This is importantly only true if the estimated $\hat{\gamma}_t$ is actually converging to the true one. The estimated loadings should be an estimate of the true loading on the first factor $\hat{\alpha}_i$ and the τ-th quantile of F.

For simplicity, we will consider $e_{it} \sim \mathcal{N}(0,1)$. Then for estimates at the median, we should expect that there is only one factor. For estimates at $\tau = 0.25$, we should expect that the mean value of the second estimated factor is $\Phi^{-1}(0.25) = -0.67$.

We consider the DGP as above with $\alpha_i, \beta_t, w_i, x_t$ all i.i.d. $\mathcal{N}(0,1)$. To satisfy $\eta_i' \gamma_t > 0$, we define

$$\eta_i = e^{w_i}$$
$$\gamma_t = e^{x_t}$$

We then generate data according to this DGP for $T = 1000, N = 100$. The average R^2 is reported in table **??**.

Table 1.2: Correlation Between Estimated and True Parameters - DGP 2

	PC			QPC ($\tau = 0.25$)			QPC ($\tau = 0.50$)			QPC ($\tau = 0.75$)		
$T \setminus N$	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

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

Figure 1: Caption