QFM Exploration Notes

April 2, 2025

1 Location-Scale Model Notes

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 **??**.