QFM Exploration Notes

April 9, 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:

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

In general, we let F be the cdf of e_{it}. Then the conditional quantile function is:

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

We call the estimates of this conditional quantile function from the QPC algorithm $\hat{\alpha}, \hat{\beta}, \hat{\eta}, \hat{\gamma}$. If the algorithm converges to the true values (as captured by the common component), we would expect that the estimated factors span the same space as the true ones. Because the rotation of the factors is not identified, we need to measure this in a way that is invariant to rotation. To do so, we consider the R^2 of the following regressions:

$$\beta = a + b_1 \hat{\beta} + b_2 \hat{\gamma}; \quad \gamma = c + d_1 \hat{\gamma} + d_2 \hat{\gamma};$$

If the algorithm is converging properly, we would expect high values of R^2 in each of these regressions.

Note that under this DGP, beacuse $e_i t$ is $\mathcal{N}(0,1)$ we have that at the median $(\tau = 0.50)$ there will only be one factor.

For now, I will consider the errors-in-variables rotation, where with k_τ factors, the factor loadings are restricted such that $\Lambda = [I_{k_\tau} \Lambda_2]'$. This is the most simple computationally, though there are others I could consider from Bai and Ng (2013).

For now, I want to see how this DGP behaves and wether we can do any estimation consistently. My target is to replicate table 2 from Sagner (2019), shown in Figure 1.

I will not report PC estimates. In addition to the mean R^2 values for both the first and second factor, I will include the proportion of simulations where the process took a long time to converge (> 100 iterations), where it didn't converge at all (not converged at 1000

Figure 1

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 $\overline{\rho}_X = S^{-1} \sum_{s=1}^{S} \rho(\hat{X}^{(s)}, X)$, where $\hat{X}^{(s)}$ is an estimator (QPC or PC), and X is its true counterpart.

iterations), the proportion of simulations where the R^2 for the second factor is > 0.5, and the proportion of simulations where the R^2 of the first factor is < 0.9. For small N and T this last value may be sizable, but for larger N and T a large value of this proportion is concerning as it implies that the first factor is not being estimated consistently.

I want to first check that my estimation procedure is working correctly. To do so, I will consider 2 simple DGPs:

$$y_{it} = \alpha_i \beta_t + e_{it} \tag{1}$$

$$y_{it} = \alpha_i \beta_t + \eta_i \gamma_t + e_{it} \tag{2}$$

These DGPs are simple location models, and so we should expect a good fit for both the one factor case (Equation 1) and the two factor case (Equation 2). Results of a small simulation study are reported in.

Now that the simulation is behaving as anticipated, I will explore the question of initial values. For this study, I consider the following DGP:

$$y_{it} = \alpha_i \beta_t + \eta_i \gamma_t e_{it} \tag{3}$$

Where $\eta_i, \gamma_t \sim \chi_1^2$ to satisfy $\eta_i \gamma_t > 0$. I fit a QFM using the same procedure as before, but with a change in the initialization step. Instead of beginning with PCA to guess the initial values, I add noise from a $\mathcal{N}(0,1)$ to the initial PCA estimate. This is then rotated according to the errors-in-variables rotation I have used previously. The model is fit from these new initial values, and some interesting patterns appear. Consider the examples in tables 1, 2, and 3.

Table 1: QFM Fits With Different Noise - Bad Second Factor Fit

Noise Seed	First Factor Fit	Second Factor Fit	Iterations	Objective Function Value
No Noise	0.9921	0.1600	14	3909.9828
134935	0.9912	0.7791	33	3595.2488
363439	0.9924	0.7245	18	3645.0155
880628	0.9927	0.1601	24	4005.6856
252318	0.9954	0.7844	84	3677.4222
344982	0.9965	0.7391	16	3591.6591
702677	0.9920	0.8051	54	3555.9745
116075	0.9910	0.1600	19	3915.6519
100298	0.9934	0.7755	25	3667.7547
178700	0.9943	0.1599	28	3940.8150
310893	0.9903	0.7956	28	3564.4116
True	0.9950	0.8008	47	3525.9113

Notes: This table reports the R^2 value for a regression each of the true factors on the estimated factors, as well as the value of the objective function evaluated at the estimate. Each observation is a different seed for random noise added to initial values generated by PCA.

In table 1, the fit based on the PCA starting values is quite poor for the second factor. Using different starting values, I am able to get a better fit for the second factor which is mirrored with a large decrease in the value of the objective function. Adding noise can also make the fit worse in terms of objective function value.

In table 2, the fit based on the PCA starting values is quite poor for both factors. Using different starting values, the fit is improved dramatically with the first factor being fit well and the second factor being fit decently.

In table 3, the fit based on the PCA starting values is very good. However, different starting values can produce a lover value of the objective function.

Table 2: QFM Fits With Different Noise - Poor PCA Fit

Noise Seed	First Factor Fit	Second Factor Fit	Iterations	Objective Function Value
No Noise	0.0119	0.0639	3	4548.2392
32851	0.9983	0.4868	91	2921.2252
305155	0.9985	0.1999	28	3004.2602
530302	0.9979	0.0911	26	3056.5396
53515	0.9965	0.4573	21	2967.1183
965115	0.9975	0.0176	100	3185.0630
201177	0.0703	0.3880	95	4157.0256
380261	0.9975	0.0422	11	3015.5472
406115	0.9982	0.0828	100	3128.7703
445781	0.9974	0.3452	15	2956.7660
912360	0.9981	0.3322	72	2969.2764
True	0.9971	0.5003	7	2901.3943

Notes: This table reports the R^2 value for a regression each of the true factors on the estimated factors, as well as the value of the objective function evaluated at the estimate. Each observation is a different seed for random noise added to initial values generated by PCA.

Table 3: QFM Fits With Different Noise - Good PCA Fit

Noise Seed	First Factor Fit	Second Factor Fit	Iterations	Objective Function Value
No Noise	0.9969	0.8077	68	3119.1333
688017	0.9968	0.8088	48	3112.0576
7268	0.9971	0.7998	20	3155.9562
934064	0.9969	0.8006	25	3164.4899
343860	0.9974	0.7982	18	3141.5936
798375	0.9965	0.7849	19	3050.5442
631366	0.9967	0.7847	71	3257.3209
124501	0.9957	0.7886	38	3184.5473
163844	0.9959	0.7902	51	3116.7062
673964	0.9975	0.7865	45	3170.8194
272234	0.9973	0.7840	19	3144.6454
True	0.9992	0.8096	10	2989.2456

Notes: This table reports the R^2 value for a regression each of the true factors on the estimated factors, as well as the value of the objective function evaluated at the estimate. Each observation is a different seed for random noise added to initial values generated by PCA.

Next, I consider starting estimation with the true factor levels and loadings. In each of the 3 examples above, using the true levels and loadings improves the fit and lowers the objective function value. I think this can be thought of as the ideal scenario, at least for the angle from which I am approaching the problem. The trouble we have with PCA is that it is based on the mean, and so is uninformative of the true value of the second factor because at the median, the second factor doesn't contribute. We want the initial value to be close to the true value so that we are close to the global optimum in the objective function. This is then a benchmark against which methods can be compared.

References

Bai, J., and S. Ng (2013) "Principal components estimation and identification of static factors," *Journal of Econometrics*, 176(1), 18–29.

Sagner, A. G. (2019) "Three Essays on Quantile Factor Analysis," .