Discussion of:

"Comparing Asset Pricing Models with Traded and Non-Traded Factors"

by Rohit Allena

Christian Julliard

London School of Economics

The problem: Comparing Models

"All models are wrong, but some are useful."
Box (1976)

Suppose we have:

- **Data** Z
- **Model 0** given by the likelihood $f(Z|\theta)$ with $\theta \in \Theta$
- **Model 1** given by the likelihood $g(Z|\psi)$ with $\psi \in \Psi$

Questions: *How do we decide which model is more likely of being the data generating process of* Z*?*

The Bayesian Answer: Posterior Model Probabilities

- define $m = 0$ if model 0 is true and $= 1$ otherwise
- we would like to know $\Pr(m = 0|Z)$.
- From Bayes' thm:

$$\begin{aligned}\Pr(m=0|Z) &= \frac{\Pr(Z|m=0)\Pr(m=0)}{\Pr(Z)} \\ &= \frac{\Pr(Z|m=0)\Pr(m=0)}{\Pr(Z|m=0)\Pr(m=0)+\Pr(Z|m=1)(1-\Pr(m=0))}\end{aligned}$$

Note: the distribution of $Z|m,\theta,\psi$ is simply

$$f(Z|\theta)^{1-m} g(Z|\psi)^m$$

⇒ if we have *prior probability of* $m = 0$ $(\mu(\Theta))$, and priors over θ and ψ $(p(\theta)$ and $q(\psi))$, from Bayes Th. we can compute

posterior probability of $m = 0|Z$

$$\frac{\mu(\Theta)\int_\Theta f(Z|\theta)\,p(\theta)\,d\theta}{\mu(\Theta)\int_\Theta f(Z|\theta)\,p(\theta)\,d\theta+(1-\mu(\Theta))\int_\Psi g(Z|\psi)\,q(\psi)\,d\psi} \quad (1)$$

Bayes Factors and Posterior Odds

In the two models case the **posterior odds** of Model 0 (odds = prob./(1 -prob.)) is

$$\frac{P(m=0|Z)}{P(m=1|Z)} = \underbrace{\frac{\int_{\Theta} f(Z|\theta)\, p(\theta)\, d\theta}{\int_{\Psi} g(Z|\psi)\, q(\psi)\, d\psi}}_{\text{Bayes Factor } BF \equiv ML_0/ML_1} \times \underbrace{\frac{\mu(\Theta)}{1-\mu(\Theta)}}_{\text{Prior odds}}$$

Recall: in a time series regression of an asset excess return on traded factors the intercept, α, should be zero.

$\Rightarrow$ **Barillas & Shanken (2018)**:

1. compare restricted ($\alpha = 0$) and unrestricted time series regression via BF (prior odds set to 1/2).

   note: under a suitable diffuse prior, the BF has a simple analytical expression: just proportional to a $F-$statistic
2. Fixing a "model 1" (denominator of the BF), use the BFs of various other factor models to construct the models' posterior probs.
3. If factors are not traded (e.g. GDP), hence the α restriction does not hold, use their mimicking portfolios (i.e. their linear projection on the space of returns) and proceed as above (since in this case the α should be zero).

This paper: uncertainty about the mimicking portfolios

- Correctly point out that BS disregard the uncertainty about the parameters, say ω, of the projection of factors on returns.

Idea: since the ω come from a linear regression, we know their posterior distribution $\pi(\omega)$ (a normal-inverse-Wishart under standard assumptions).

$\Rightarrow$ Integrate out ω (via MCMC) to construct for each model j:

$$ML_j = \int ML_j(\omega)\pi(\omega_j)d\omega_j$$

and then proceed as in BS(2018).

- Similarly, when factors are principal components, incorporate the uncertainty about estimating the covariance matrix of return (Σ)

$$ML_j = \int ML_j(\Sigma)\pi(\Sigma)d\Sigma$$

(where $\pi(\Sigma)$ is an inverse-Wishart under standard assumptions)

Simple and clever! And leads to some surprising results: now non-traded factor models do much better than in BS... and that's very odd... unless something goes wrong...

A similar problem: Sim's thinking about IV

The mimicking portfolio cum *alpha* regression is analogous to the classical IV model:

$$\underset{T\times 1}{y} = x\beta + \epsilon = \underset{T\times k}{Z}\gamma\beta + v\beta + \epsilon, \quad (2)$$

$$\underset{T\times 1}{x} = Z\gamma + v, \qquad Var([v\beta + \epsilon \;\; v]) = \Sigma \quad (3)$$

The problem:

- Under the (2)-(3) parametrization the likelihood does not go to zero as $\beta \to \infty$. This is because for very large β, the best fit can be achieved with very small $||\gamma||$ so that $\beta\gamma$ gives a good fit. This will make the fit of (3) poor, but this is bounded from below, hence it does not send the likelihood to zero.

$\Rightarrow$ Baseline: MCMC under a flat/diffuse prior will not converge, and integrating parameters is likely to fail, leading to improper marginals.

But: a non-flat prior can be used to deliver a proper posterior e.g. a prior proportional to $||\gamma||(1+\beta^2)^{1/2}$ (but not a silver bullet)

Note: this is what frequentists call the "weak instrument" problem (it's a general problem, not a Bayesian one).

Weak/spurious factors as great factors

Recall: a spurious/weak factor is one whose covariance with returns goes too zero i.e. when constructing the mimicking portfolio

$$f_t = c + \omega^\top R_t + \eta_t \quad \to f_t^m = \omega^\top R_t$$

we have that $\omega^\top \to 0$.

But: when assessing the factor we estimate alphas from

$$\underset{N\times 1}{R_t} = \underset{N\times 1}{\alpha} + \underset{N\times 1}{\beta} f_t^m + \underset{N\times 1}{\epsilon_t} = \underset{N\times 2}{B}\underset{2\times 1}{X_t} + \epsilon_t$$

$$\Rightarrow R = BX + \varepsilon \quad \Rightarrow \quad \hat{B} = RX^\top (XX^\top)^{-1}$$

- but if $\omega^\top \to 0$ the projection diverges! (same as IV problem)
- Furthermore, integrating the likelihood with a flat/Jeffrey's prior delivers a quantity $\propto |XX^\top|^{-N/2} \to \infty$ as $\omega^\top \to 0$.

$\Rightarrow$ weak/spurious factors will have posterior probability $\to 1$!

- is this driving the result for non-traded factors? It would all make sense...

In summary

- (+) After important question
- (+) Natural approach...
  - (-) ... but one cannot do naive MCMC under a flat/Jeffrey's prior in this setting
  - (-) ... and flat/Jeffrey's prior will lead to selection of weak/spurious non-traded factors.
- (+) the critique of PCs as factors is spot on, as well as the MCMC integration idea in that case...
  - (-) ... but Barillas & Shanken (2017) not ideal in this case (pricing PCs with other PCs, i.e. pivotal factor becomes key) $\Rightarrow$ instead, just do straight post prob. from pricing the cross-section.

Baseline: a very exciting start for a paper project... but it's not a low hanging fruit (and that's probably why BS haven't done it).