Uncommon Factors for Bayesian Asset Clusters

Lin William Cong, Guanhao Feng, Jingyu He, & Junye Li

Discussion at the 2023 MFA by
ROHIT ALLENA
C.T. Bauer College of Business
University of Houston

March 17, 2023

Summary of the Paper

Unified factor model does not explain all stock returns sufficiently well

- Different factor models fit different set of stock returns

The paper's contribution is to develop a novel Bayesian Clustering Method that

1. classifies stocks into different clusters
2. and fits separate factor models for different stock clusters

The main highlights of the paper are that it

1. demonstrates how to cluster stocks into different buckets
2. and shows how to choose the best factor model among many others that fits a given cluster of stocks
3. validates empirically that fitting cluster-specific factor models yield better return forecasts

Comments on the Contribution

The paper makes a significant contribution to the asset pricing literature.

1. Factor model heterogeneity has been acknowledged by many studies
2. But limited prior work that demonstrates how this heterogeneity can be modeled
3. Patton and Weller (RFS 2022) model risk-price heterogeneity, but they assume a known factor model
4. This paper takes the first (and the important) step in mathematically formulating the factor-model heterogeneity
5. And I believe that many papers will adapt/generalize this paper's method in the future, as modeling factor heterogeneity is important

Methodology

Summary of the method:

1. The paper classifies stocks into clusters based on stock characteristics
2. For each asset cluster, given K factors, the paper considers all 2^K factor models and picks the best model among them

How to classify stocks into clusters based on stock characteristics?

1. The paper uses a stylized tree model
2. Example: Cluster 1: Size$<$0.5; Cluster 2: Size$>$0.5, BM$<$0.2; Cluster 3: Size$>$0.5, BM$>$0.2

How to choose split points?

1. Consider all split points, choose the one that maximizes the overall fit

Comments related to the methodology

Comment #1: The paper's method is not entirely Bayesian

- The paper first classifies stocks into different clusters by considering only those splits that have the maximum posterior probabilities (MAP) in respective clusters.
- By doing so, the method ignores all other split combinations that may be useful
- For instance, what are the factor models on other splits that do not have the MAP but that still have significant posterior probabilities
- Ignoring split uncertainty not Bayesian may be detrimental for forecasting returns

Suggestion:

- Since considering all split combinations may be computationally infeasible, at least take the ones with significant posterior model probabilities

Comments on the Methodology

Comment #2 The (frequentist) properties of the methodology is not clear

1. Under what conditions does the method recover the correct clusters and the true factor models?
2. For instance, the paper describes excess returns using a linear model but implements a stylized tree model for clustering. Why?
3. Under what assumptions can the stylized tree model identify the correct clusters?

The main advantage of a unified (rather than clustered) factor model is that

- It exploits the entire cross-section to deliver precise risk premium estimates

But clustered factor models do not have the "blessings of dimensionlity"

- While they can deliver accurate forecasts, the forecast precision deteriorates
- Thus, analyzing bias-variance trade-off becomes important

More comments on the methodology and the empirics

1. The paper assumes that model residuals are iid. Is it justifiable?

2. While comparing factor models, it is important to impose $\alpha = 0$ restriction. But the paper allows for mispricing. Philosophically, are models comparable under mispricing?

3. Would it be possible to test the null of "no clusters" based on Bayes Factors? That would be a great contribution

4. Empirically, BCM delivers improved return forecasts and efficient portfolios.
   - But how much of that improvement is due to clustering?
   - How much of it is due to selecting the appropriate factor model?
   - Would it be possible to decompose the improvements?

Conclusions

- Great asset pricing paper!
- Perhaps the first paper to mathematically formulate the factor-model heterogeneity
- Look forward to reading the next version