Asset Heterogeneity and Uncommon Factors

Qianshu Zhang
City University of Hong Kong
Joint With Lin William Cong, Guanhao Feng, Jingyu He, and Junye Li

Highlights

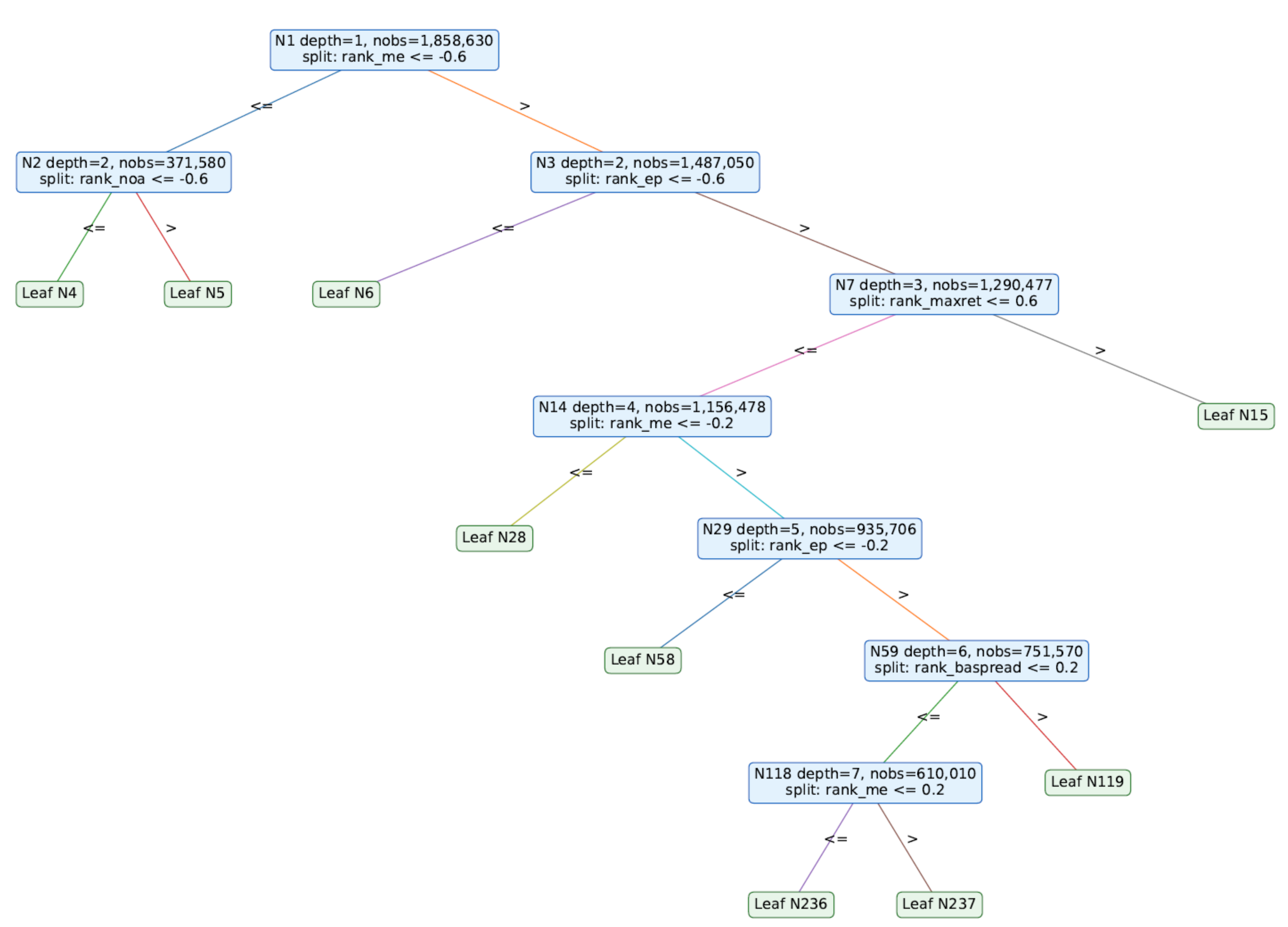
- **Methodology:**
  - Tree from machine learning + Fama Macbeth Regression + Lasso penalty for risk premia.
  - Heterogeneous asset clusters and heterogeneous factor risk premia.
- **Empirical:**
  - Equity show cross-sectional heterogeneity.
  - Factor zoo are locally sparse but globally dense.
  - Cluster-specific investment outperform pool one.

Motivation

- A zoo of factors are proposed, aiming to price the cross section. (e.g., Cochrane, 2011, JF)
- One direction seeks a parsimonious set of factors (e.g., Feng, Giglio and Xiu, 2020, JF, Bryzgalova, Huang, and Julliard, 2023, JF)
- Another direction emphasize that the "true" SDF can not be spanned by a small number of factors. (Kozak, Nagel and Santosh, 2020, JFE)
- We argue, however, that both sides overlook heterogenous factor risk premia.
- We relax the universal pricing assumption by identifying cluster-specific cross-sectional factors.

Our solution for the cross-sectional heterogeneity in factor pricing: adapting and modifying the panel tree framework, which enables endogenous asset cluster detection through a data-driven methodology.

Figure1: Asset Heterogeneity in the cross section



Methodology

Cluster-specific cross-sectional factor model (Fama and French 2020)

$$r_{i,t} = A(\mathbf{z}_{i,t-1}) + \mathbf{z}_{i,t-1}^{\intercal} F(\mathbf{z}_{i,t-1}) + \epsilon_{i,t},$$

where

$$A(\mathbf{z}_{i,t-1}) = \sum_{j=1}^{J} \mathbb{1}_{\{\mathcal{T}(\mathbf{z}_{i,t-1})=j\}} r_{j,z,t},$$

$$F(\mathbf{z}_{i,t-1}) = \sum_{j=1}^{J} \mathbb{1}_{\{\mathcal{T}(\mathbf{z}_{i,t-1})=j\}} \mathbf{f}_{j,t},$$

Lasso penalty on risk premia

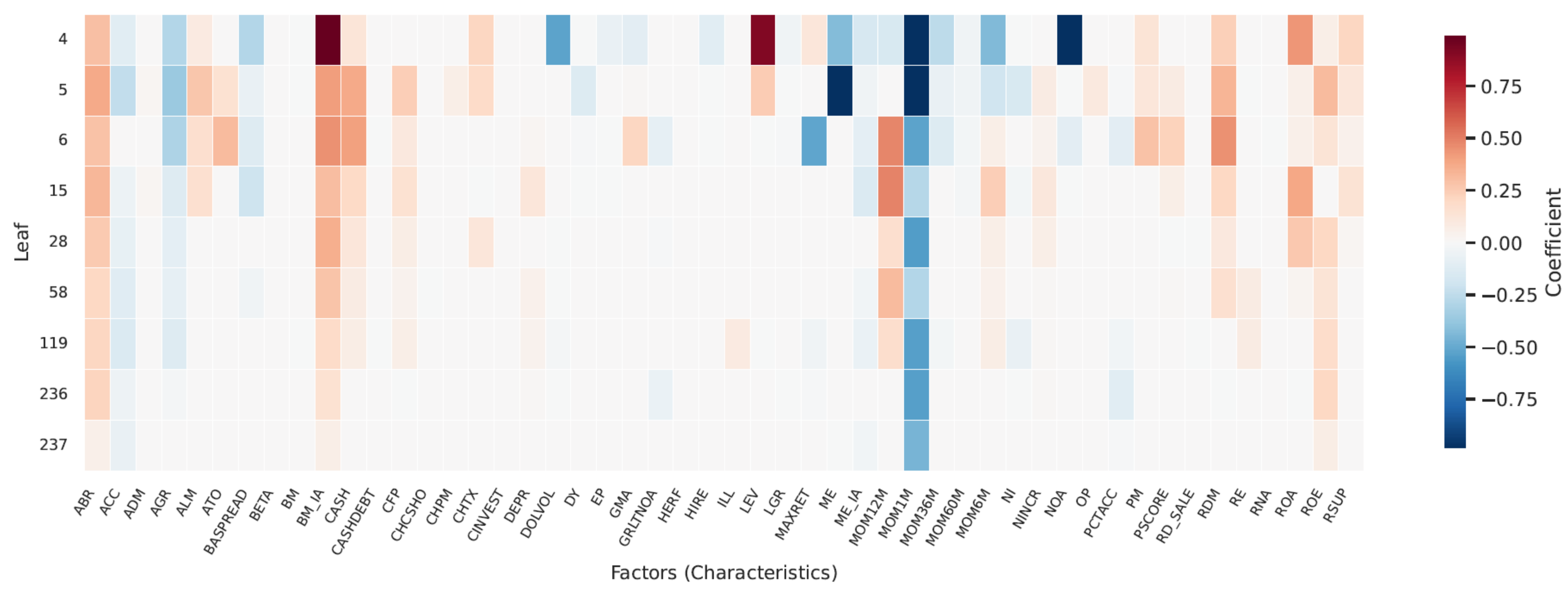
$$\min_{\{r_{j,z,t},\mathbf{f}_{j,t}\}_{t\in\mathcal{T}_j}} \sum_{t\in\mathcal{T}_j} \frac{1}{2n_{j,t}} \|\mathbf{R}_{j,t} - r_{j,z,t}\mathbf{1} - \mathbf{Z}_{j,t-1}\mathbf{f}_{j,t}\|_2^2 + \lambda \sum_{k=1}^{K} \left| \frac{1}{|\mathcal{T}_j|} \sum_{t\in\mathcal{T}_j} f_{j,t,k} \right|.$$

Tree splitting criterion: Minimizing pricing error

$$\text{Score}(p) = \frac{1}{|\mathcal{I}_p|} \sum_{i\in\mathcal{I}_p} \frac{1}{T_{i,p}} \sum_{(i,t)\in\mathcal{O}_p} \left(r_{i,t} - \hat{r}_{i,t}^{(p)}\right)^2$$

Empirical Result

Figure2: Heterogeneous Factor Risk Premia



- Cluster-specific Long-short portfolio achieves better performance, especially value-weighted

Table1: Out-of-sample Investment Performance

	Cluster EW	Cluster VW	Overall EW	Overall VW
Mean Return	3.01	2.11	1.42	0.41
t-stat	(8.26)	(7.73)	(7.90)	(1.60)
Median Return	2.86	1.84	0.93	0.38
SD	4.27	3.49	2.93	3.52
Skewness	0.20	0.46	1.12	0.31
Kurtosis	2.74	3.25	6.67	5.16
AC1	-0.04	-0.12	-0.16	-0.04
Sharpe Ratio	2.44	2.09	1.67	0.41
Max Drawdown	0.17	0.15	0.09	0.32

Table2: Cluster VW spanning test

	(1)	(2)	(3)
Alpha	1.870*** (6.84)	1.856*** (6.71)	1.885*** (6.99)
MktRf	0.217*** (3.24)	0.224*** (3.22)	0.213*** (3.42)
SMB	-0.087 (-0.72)	-0.078 (-0.63)	
HML	0.200* (1.74)	0.212* (1.77)	
RMW	0.053 (0.36)	0.060 (0.40)	
CMA	-0.131 (-0.76)	-0.143 (-0.82)	
UMD		0.031 (0.37)	
LIQ			0.021 (0.26)

- Alpha remain significant after controlling for common factor models.
- Factors have little explanatory power.

Department of Decision Analytics and Operations
香港城市大學 City University of Hong Kong