Discussion of Bryzgalova, Huang, Julliard

"Bayesian Solutions for the Factor Zoo: We Just Ran Two Quadrillion Models"

Bryan Kelly

Yale

- Excellent paper
- Tour de force in Bayesian AP, with great clarity in exposition
- Next hour is a great investment of your time!

Why Do We Study Factor Models?

Two (Caricature) Perspectives

The Asset Manager: We know the risks, observe returns, ask "were can I get a good deal?"

- Factor model is risk model: Critical to get risk right (high time series R^2)
- Alpha vs. whole model, little interest in nature of individual factors
- Nature of alpha is central interest (arbitrage limits)
- Few subtle utility considerations, focus on variance or drawdown risk

The Behavioral Scientist: Assume returns compensate for risk, observe returns, ask "which risks do people respond to?"

- Particular interest in alignment between expected returns and factors (high cross section R^2)
- Nature of factor is of central interest, less focus on nature of alpha
- Less focus on time series R^2, openness to subtle utility considerations

Why Do We Study Factor Models?

The Empirical Asset Pricer: A bit schizophrenic? "Can I find *any* model that shoehorns into observed return patterns"

- "Tests" asset pricing models and focuses on risk premium estimates (like behavioral scientist)
- But often those models are purely statistical and factors lack economic identification (like asset manager)
- How did we get here? Both the statistics and behavioral science of financial markets are hard

This Paper

- Takes state of AP literature as given
  - Inference centered on risk premia
  - But purely statistical in nature
- Objective: Address a few pitfalls that standard approaches are prone to
- Joins recent "meta-analysis" of factor models (Harvey and Liu; Barillas and Shanken; Kelly, Pruitt, and Su; Feng, Giglio, and Xiu; Kozak, Nagel, and Santosh; etc.)

Two focal points

1. Inference that is robust to "useless" factors
   - Builds on lineage of Kan and Zhang (1999)
2. Model comparison/selection
   - Builds on lineage of Barillas and Shanken (2018)

"Useless" Factors Versus Correlated Factors

Paper motivation

- Useless factors have zero exposures, blow up risk price inference
- "factor proliferation and spurious factors are tightly connected problems"
- Proposed model very well suited to deal with this

How worrisome are "useless" factors?

- Filter out before doing full blown inference?
- If we let these consistently cloud our conclusions, we are worse off then I thought in AP

Correlated factors a bigger problem?

- Factors proliferation and *correlated factors* are tightly connected problems
- Less false discovery in finance than *re*-discovery
- All factors are noisy, some of this noise offsets
- Factor selection likely inferior to factor combination
- Difficult to adjudicate risk premia
- Reduced informativeness of model comparison
- Paper can speak to this

Other Comments

- Non-tradable factors in a statistical factor model?
  - Barillas and Shanken (2018) and Chib, Zeng, and Zhao (2020) cover Bayesian model comparison for tradable factors
- Two quadrillion minus one too many models? Mkt+SMB+HML
- Method might bypass interesting models in the set of two quadrillion
- Many equally probable models, probable models have many factors
- SDF Sharpe ratios suggest high time series R^2

An excellent paper that is a pleasure to study. Thank you!