Empirical Pricing Factors in Theoretical Economies

Kerry Back Alexander Ober Seth Pruitt

Overview

- Large literature documents relations between characteristics and returns in rational theoretical models.
- Pricing is by unobservable (latent) SDF. SDF betas are correlated with characteristics.
- What is the best way to uncover the observable (tradable) SDF?

Overview

- Large literature documents relations between characteristics and returns in rational theoretical models.
- Pricing is by unobservable (latent) SDF. SDF betas are correlated with characteristics.
- What is the best way to uncover the observable (tradable) SDF?

Theoretical models: Berk-Green-Naik (1999), Kogan-Papanikolaou (2014), Gomes-Schmid (2021)

Overview

- Large literature documents relations between characteristics and returns in rational theoretical models.
- Pricing is by unobservable (latent) SDF. SDF betas are correlated with characteristics.
- What is the best way to uncover the observable (tradable) SDF?

Theoretical models: Berk-Green-Naik (1999), Kogan-Papanikolaou (2014), Gomes-Schmid (2021)

Empirical methods: Fama-French, Fama-MacBeth (& Rosenberg), Kelly-Pruitt-Su (IPCA), Didisheim-Ke-Kelly-Malamud (Complexity)

Empirical Methods

Two Empirical Approaches

1. Define factor weights by some recipe as functions of cross-section of characteristics. Examples: Fama-French, Fama-MacBeth-Rosenberg, DKKM
2. Or assume latent factors. Estimate betas. Then extract factors. Examples: PCA, IPCA.

Kelly-Pruitt-Su, 2019: IPCA = Instrumented PCA

- Latent factors f_t
- Characteristics z_{it}
- Returns $r_{i,t+1} = z_{it}'\Gamma f_{t+1} + \varepsilon_{i,t+1}$
- Choose Γ and f_t to minimize sum over i and t of squared residuals $r_{i,t+1} - z_{it}'\Gamma f_{t+1}$

Kelly-Pruitt-Su Empirics

- 36 firm characteristics
- Rank-standardize characteristics to $[-0.5, +0.5]$ interval
- Test assets: individual stocks
- 5 factors do very well
- Out-of-sample Sharpe ratios:
  - IPCA tangency portfolio: 2.5 annually
  - FF5 tangency portfolio: 1.3 annually

Didisheim-Ke-Kelly-Malamud, 2023: Random Fourier Features

- Factors = returns of portfolios whose weights are rank-standardized characteristic values in $[-0.5, 0.5]$
  - Example: book-to-market factor would be return of portfolio that is long value stocks (above median bm) and short growth stocks (below median bm).
  - "More value" $\Rightarrow$ higher weight. "More growth" $\Rightarrow$ more negative weight.

Didisheim-Ke-Kelly-Malamud, 2023: Random Fourier Features

- Factors = returns of portfolios whose weights are rank-standardized characteristic values in $[-0.5, 0.5]$
  - Example: book-to-market factor would be return of portfolio that is long value stocks (above median bm) and short growth stocks (below median bm).
  - "More value" $\Rightarrow$ higher weight. "More growth" $\Rightarrow$ more negative weight.
- Start with M characteristics c_{ik}. Generate M random numbers h_k. Compute two new composite characteristics

$$\cos\left(\sum_{k=1}^{M} h_k c_{ik}\right) \quad \text{and} \quad \sin\left(\sum_{k=1}^{M} h_k c_{ik}\right)$$

Didisheim-Ke-Kelly-Malamud, 2023: Random Fourier Features

- Factors = returns of portfolios whose weights are rank-standardized characteristic values in $[-0.5, 0.5]$
  - Example: book-to-market factor would be return of portfolio that is long value stocks (above median bm) and short growth stocks (below median bm).
  - "More value" $\Rightarrow$ higher weight. "More growth" $\Rightarrow$ more negative weight.
- Start with M characteristics c_{ik}. Generate M random numbers h_k. Compute two new composite characteristics

$$\cos\left(\sum_{k=1}^{M} h_k c_{ik}\right) \quad \text{and} \quad \sin\left(\sum_{k=1}^{M} h_k c_{ik}\right)$$

- Repeat many times to get thousands of new composite characteristics.
- Define a factor as described before for each of the new composite characteristics

DKKM Empirical Results

- 130 stock characteristics
- Out-of-sample Sharpe ratios reach 4.0 with high-complexity models (10,000+ factors)
- Fama-French-Carhart 6-factor Sharpe ratio ≈ 1.1

Data-Generating Models

Theoretical Models: Overview

- In all models, risk premia depend on covariances with a stochastic discount factor.
- In all models, the covariances are correlated in the cross-section with firm characteristics.
- Hence, characteristics ‘explain’ risk premia.
- In all models, we can compute size, book-to-market, profitability, asset growth, and momentum (the FFC6 characteristics).
- We can also compute the true traded stochastic discount factor and the true tangency portfolio.

Berk, Green, and Naik (1999)

- Firms invest optimally given an exogenous pricing kernel
- Fixed number of firms, each receives take-it-or-leave-it investment opportunities each period
- Projects generate operating cash flows until they randomly die
- Investment depends on project NPV (which varies with beta and interest rates)
- Model generates: book value, market value, net income, stock returns
- Characteristics: size, book-to-market, ROE, asset growth, momentum

Kogan and Papanikolaou (2014)

- Firms invest optimally given an exogenous pricing kernel
- Two aggregate state variables:
  - Disembodied productivity affecting all capital
  - Productivity of newly installed capital
- Firms acquire projects stochastically at firm-specific rates
- Optimal capital investment choice for each project
- Firm-specific and project-specific productivity processes
- Projects produce cash flows until they randomly expire

Gomes and Schmid (2021)

- General equilibrium model with heterogeneous firms
- Representative household with Epstein-Zin preferences (endogenous SDF)
- Firms make optimal investment and financing decisions
- Single aggregate productivity process (AR(1)) + firm-specific shocks
- Stochastic investment opportunities with random costs
- Lumpy investment (discrete project adoption)
- Debt: consol bonds paying coupons until random expiration
- Tax benefits of debt; costly equity issuance (pecking order)
- Strategic default when equity value ≤ 0

Our Simulations

Motivation for Simulations

- In data-generating models, true betas with respect to the SDF depend on entire history of firm-specific and macro shocks
- Characteristics-based factor models use observable firm characteristics to construct traded factors
- Betas with respect to these factors partially explain risk premia
- Our questions: For a given set of characteristics, what is the best way to construct traded factors? Is the answer robust across models?

Simulation Design

- 1,000 firms and 920 months in each panel (discard first 200 as burn-in, leaving 60 years)
- 10 independent panels for each data-generating model (results very consistent across panels for each model)
- In each panel for each model in each month, compute true conditional SDF and true conditional max Sharpe ratio
- Use calibrations from original papers, except we substitute exogenous lognormal SDF in GS, calibrated to match market risk premium

Evaluations of Empirical Methods

Don't use any specific set of test assets. Instead:

- Compare conditional max Sharpe ratios of estimated MVE portfolios of factors

  Barillas-Shanken, 2017: in horse race between factor models, assuming test assets include competing factors, model with highest Sharpe ratio wins

- Estimate conditional SDFs implied by the models and compute Hansen-Jagannathan distance to true conditional SDF

  HJ distance is the maximum pricing error over all test assets with unit uncentered second moment

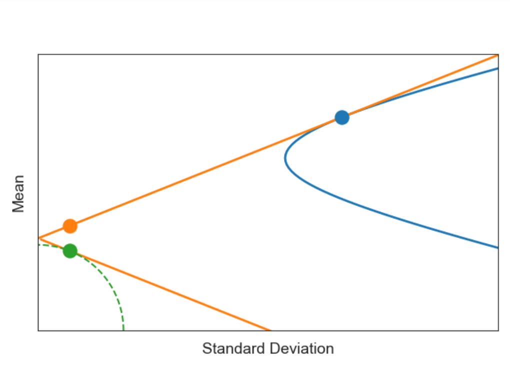
Britten-Jones Regression

- Can find empirical mean-variance frontier by linear regression of constant 1 on asset excess returns
- Write fitted regression as

$$
\begin{aligned}
1 &= \sum_{i=1}^{N} \hat{\beta}_i (r_i - r_f) + \hat{\varepsilon} \\
&= \hat{z} + \hat{\varepsilon}
\end{aligned}
$$

- Max Sharpe ratio is $\bar{z}/\sigma_z$
- Result originally due to Hansen and Richard for population: projection in population rather than regression in sample

Britten-Jones Regression



- Green dot $= (1 + r_f)\hat{\varepsilon} - 1$
- Orange dot $= r_f + (1 + r_f)\hat{z}$, the Sharpe ratio of which equals $\bar{z}/\sigma_z$
- Efficient part of frontier is $\{r_f + b\hat{z} \mid b \geq 0\}$

Our Implementation

- Use known true distributions to calculate $\hat{\varepsilon}$. Use as SDF (orthogonal to excess returns).
- In empirical factor models, use Britten-Jones regression on rolling 360 month windows (following DKKM)
- With many factors (maybe more than 360), use ridge penalization in Britten-Jones regression (following DKKM)

Performance Measures for Each Factor Model

- Mean theoretical conditional max Sharpe ratio in each panel
- Realized HJ distance: Square root of mean value of $(\hat{\varepsilon}_{\text{factors}} - \hat{\varepsilon}_{\text{all-returns}})^2$ in each panel
- Both averaged across panels

Ridge Regression

- Minimize: $\frac{1}{T}\sum_{t=1}^{T}(1-\beta' F_t)^2 + \alpha\beta'\beta$
- Penalty parameter α controls shrinkage toward zero
- Essential when number of factors M is large relative to sample size T
- We set $\alpha = \kappa M$ and tune κ to optimize performance
- Tried ridge but performance of Fama-French-Carhart, Fama-MacBeth-Rosenberg, and Kelly-Pruitt-Su declines when regression is penalized

Empirical Factors

- Form factors from size, book-to-market, operating profitability, asset growth, and momentum in all models
- Fama-French-Carhart (FFC): usual 2×3 sorts, use size/book-to-market sort to form SMB
- Fama-MacBeth-Rosenberg (FMR): Fama-MacBeth regressions on characteristics
- Kelly-Pruitt-Su (KPS): latent factors with loadings linearly related to the five characteristics plus an intercept
- Didisheim-Ke-Kelly-Malamud (DKKM): random Fourier features built from the five characteristics plus market return (not penalized in ridge)

Fama-MacBeth-Rosenberg

- Fama-MacBeth (1973), Rosenberg (1976), Fama (1976)
- Regression coefficients $(X'X)^{-1}X'y$ are linear combinations of returns y
- Set $W = X(X'X)^{-1}$ so regression coefficients are $W'y$
- $\#$ columns $W = \#$ characteristics $+1$
- $X'W = I$ implies columns of X and W are orthonormal.
- Many solutions W of $X'W = I$, but projection $W = X(X'X)^{-1}$ solves, for each column, min $w'w$ subject to orthonormal constraint
- Being orthogonal to column of 1's implies long-minus-short portfolio. We rescale so long and short sides each sum to 1.

Results

FFC and FMR Perform about the Same

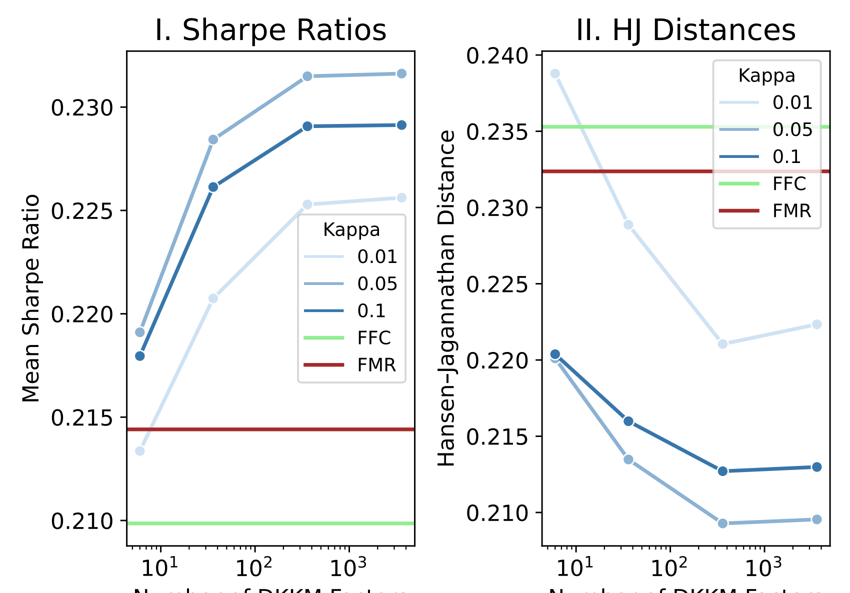
	Berk-Green-Naik		Kogan-Papanikolaou	
	Sh Ratio	HJ Dist	Sh Ratio	HJ Dist
FMR	0.217	0.230	0.194	0.180
FFC	0.209	0.232	0.203	0.168

	Gomes-Schmid	
	Sh Ratio	HJ Dist
FMR	1.640	0.487
FFC	1.457	0.538

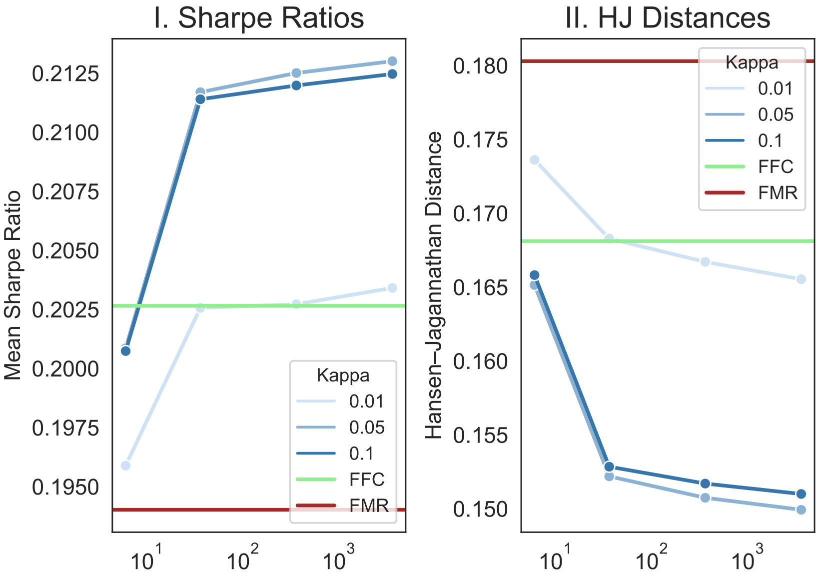
DKKM Results

- Performance increases with number of factors (given sufficient penalization)
- Optimal configurations:
  - BGN and KP: $\kappa = 0.1$, 3,600 factors
  - GS: $\kappa = 10^{-4}$, 3,600 factors

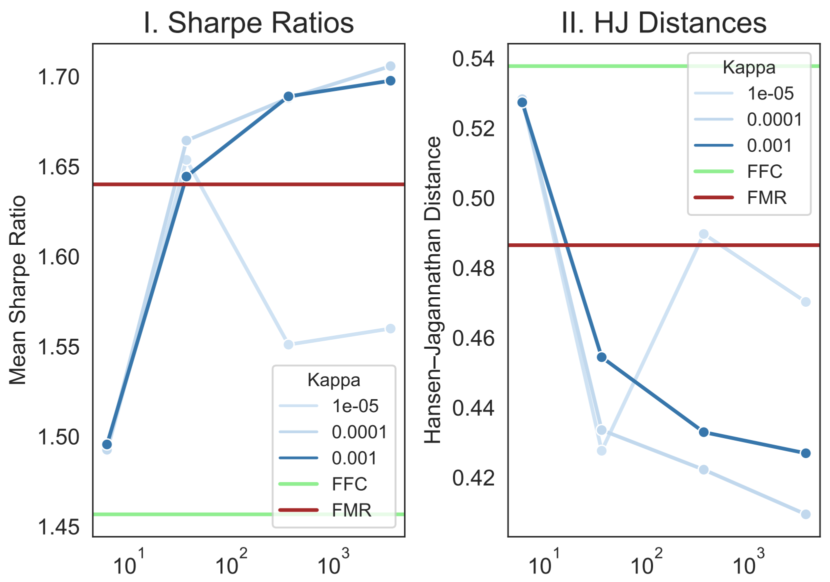
DKKM in BGN Model



DKKM in KP Model



DKKM in GS Model



KPS results

- Optimal with just 2–3 factors
- BGN and GS: 2 factors optimal
- KP: 3 factors optimal

KPS in BGN Model

Berk, Green, and Naik (1999)			
(a) Sharpe Ratio Improvement			
	1 factor	2 factors	3 factors
vs FMR	-0.094	**0.018**	0.014
vs DKKM	-0.110	**0.002**	-0.002
(b) Hansen-Jagannathan Distance Improvement			
	1 factor	2 factors	3 factors
vs FMR	0.059	**-0.038**	-0.028
vs DKKM	0.084	**-0.013**	-0.003

Note: Positive values favor KPS for Sharpe Ratio.
Negative values favor KPS for HJ Distance.

KPS in KP Model

Kogan and Papanikolaou (2014)			
(a) Sharpe Ratio Improvement			
	1 factor	2 factors	3 factors
vs FMR	-0.090	0.013	**0.014**
vs DKKM	-0.109	-0.006	**-0.005**
(b) Hansen-Jagannathan Distance Improvement			
	1 factor	2 factors	3 factors
vs FMR	0.056	-0.023	**-0.024**
vs DKKM	0.086	0.008	**0.006**

Note: Positive values favor KPS for Sharpe Ratio.
Negative values favor KPS for HJ Distance.

KPS in GS Model

Gomes and Schmid (2021)			
(a) Sharpe Ratio Improvement			
	1 factor	2 factors	3 factors
vs FMR	-1.593	**0.042**	0.020
vs DKKM	-1.658	**-0.024**	-0.046
(b) Hansen-Jagannathan Distance Improvement			
	1 factor	2 factors	3 factors
vs FMR	0.408	**-0.009**	0.000
vs DKKM	0.486	**0.069**	0.078

Note: Positive values favor KPS for Sharpe Ratio.
Negative values favor KPS for HJ Distance.

Conclusion

- In theoretical economies, we know the true latent SDF, so can calculate the true errors of pricing models.
- Provides a laboratory for evaluating models
- Current results: DKKM $\sim$ KPS $>$ FFC $\sim$ FMR in three models
- Next steps: more empirical methods, maybe more theoretical models
- Maybe explore complexity version of IPCA: blow up number of characteristics with random Fourier features, then apply IPCA.