Complexity and Shrinkage in Simple Economic Models

Kerry Back[1,2]   Alexander Ober[1]   Seth Pruitt[3]

[1] Jones Graduate School of Business, Rice University

[2] School of Social Sciences, Rice University

[3] W.P. Carey School of Business, Arizona State University

Overview

Motivation:

- Investigate recent factor construction methods 'out of sample'
  – in simulated economies

- Compare performance of Fama-French to
  Fama-MacBeth-Rosenberg (cf., e.g., Hoberg-Welch) and to
  recent methods

Motivation:

- Investigate recent factor construction methods 'out of sample'
  – in simulated economies
- Compare performance of Fama-French to
  Fama-MacBeth-Rosenberg (cf., e.g., Hoberg-Welch) and to
  recent methods

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

Overview

Motivation:

- Investigate recent factor construction methods 'out of sample'
  – in simulated economies
- Compare performance of Fama-French to
  Fama-MacBeth-Rosenberg (cf., e.g., Hoberg-Welch) and to
  recent methods

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

Empirical methods: Kelly-Pruitt-Su (2019),
Didisheim-Ke-Kelly-Malamud (2023)

Empirical Methods

Empirical Methods: Overview

Kelly-Pruitt-Su (KPS):

- A version of principal components in which the factor loadings are set as linear combinations of firm characteristics
- This shrinks the number of factors.

Didisheim et al. (DKKM):

- Form factors as returns of portfolios in which the weights are random nonlinear combinations of firm characteristics.
- They generate many thousands of such factors. This is complexity. They then estimate the stochastic discount factor using penalized regression on the factors. This is shrinkage.

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).

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).
- Start with M characteristics c_{im}. Generate M random numbers h_m. Compute two new composite characteristics

$$\cos\left(\sum_{m=1}^{M} h_m c_{im}\right) \quad \text{and} \quad \sin\left(\sum_{m=1}^{M} h_m c_{im}\right)$$

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).
- Start with M characteristics c_{im}. Generate M random numbers h_m. Compute two new composite characteristics

$$\cos\left(\sum_{m=1}^{M} h_m c_{im}\right) \quad \text{and} \quad \sin\left(\sum_{m=1}^{M} h_m c_{im}\right)$$

- 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

BGN's Simulation Results

- Replicates Fama-French (1992) value and size results in sign and magnitude
- Beta becomes insignificant when size included (as in data)
- Contrarian strategies profitable at short horizons (1-6 months)
- Momentum strategies profitable at longer horizons (1-5 years)

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

KP's Simulation Results

- Firm value = Assets in place (existing projects) + PVGO (future projects)

- Assets in place: affected only by disembodied productivity shock

- PVGO: affected by both embodied and disembodied productivity shocks

- Model replicates value premium in data

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

GS's Simulation Results

- Endogenous leverage drives risk premia
- Leverage is countercyclical: rises in recessions when default risk increases
- Credit spreads predict stock returns and business cycles
- Value firms have higher leverage $\Rightarrow$ value premium

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 mean-variance efficient portfolio

- Choose quadratic-optimal MVE portfolio: max $E[(r-1)^2]$

Evaluations of Empirical Methods

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

- Compare realized 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
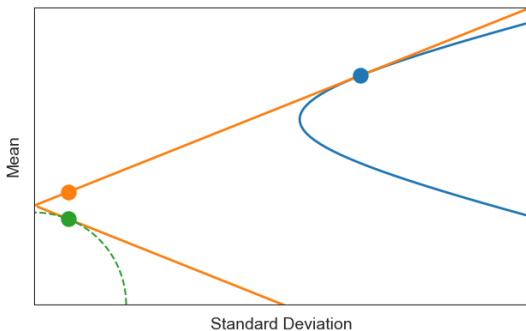
Britten-Jones Regression

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

- Write fitted regression as

$$1 = \sum_{i=1}^{n} \hat{\beta}_i (r_i - r_f) + \hat{\varepsilon}$$
$$= \hat{z} + \hat{\varepsilon}$$

- 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

- Green dot $= (1 + r_f)\hat{\varepsilon}$
- Orange dot $= r_f + (1 + r_f)\hat{z}$, the Sharpe ratio of which equals $\bar{z}/\sigma_z$. Solves max $E[(r - 1)^2]$
- 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).

- Use known true distributions to calculate orange dot $r_f + (1 + r_f)\hat{z}$ and its realized return.

- 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

- Realized Sharpe ratio of $r_f + (1 + r_f)\hat{z}$ in each panel
- Mean conditional HJ distance in each panel

Ridge Regression

- Minimize: $\frac{1}{T} \sum_{t=1}^{T} (1 - \beta' F_t)^2 + \alpha \sum_{k=1}^{M} \beta_k^2$
- 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 $W'y$
- 5 independent variables (characteristics) implies W has 6 columns. $X'W = I$ implies columns of X and W are orthonormal (corresponding columns have unit inner products, others have zero inner products)
- Many solutions W of $X'W = I$, but projection $X(X'X)^{-1}$ 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

| Berk, Green, and Naik (1999)			
(a) Sharpe Ratio (vs FMR)			
κ	6	36	360	3600
---	---	---	---	---
0.000	-0.0094	-0.0631	-0.2140	-0.1922
0.001	-0.0082	-0.0168	-0.0179	-0.0152
0.010	-0.0040	0.0054	0.0086	0.0092
0.050	-0.0025	0.0137	0.0154	**0.0157**
0.100	-0.0049	0.0131	0.0147	0.0152
1.000	-0.0398	-0.0265	-0.0257	-0.0245

(b) Hansen-Jagannathan Distance (vs FMR)			
κ	6	36	360	3600
---	---	---	---	---
0.000	0.0078	0.1930	46.0435	1.0757
0.001	0.0048	0.0357	0.0531	0.0534
0.010	-0.0041	-0.0163	-0.0142	-0.0140
0.050	-0.0026	-0.0228	-0.0246	**-0.0251**
0.100	0.0036	-0.0161	-0.0181	-0.0192

DKKM in BGN Model
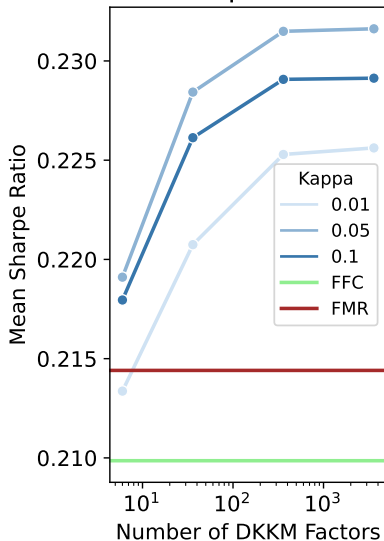
DKKM in KP Model

Kogan and Papanikolaou (2014)			
(a) Sharpe Ratio (vs FMR)			
κ	6	36	360	3600
0.000	-0.0083	-0.0673	-0.1928	-0.1766
0.001	-0.0060	-0.0235	-0.0301	-0.0295
0.010	0.0019	0.0085	0.0087	0.0094
0.050	0.0068	0.0177	0.0185	**0.0190**
0.100	0.0067	0.0174	0.0180	0.0184
1.000	-0.0112	0.0012	0.0036	0.0041
(b) Hansen-Jagannathan Distance (vs FMR)			
κ	6	36	360	3600
0.000	0.0121	0.1684	33.0227	0.9500
0.001	0.0075	0.0491	0.0586	0.0580
0.010	-0.0067	-0.0120	-0.0136	-0.0148
0.050	-0.0152	-0.0281	-0.0296	**-0.0304**
0.100	-0.0145	-0.0274	-0.0286	-0.0293

DKKM in KP Model
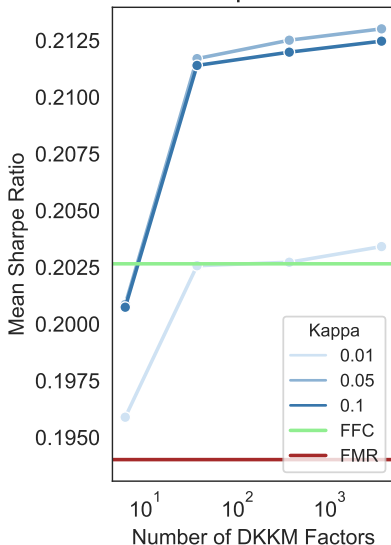
DKKM in GS Model

Gomes and Schmid (2021)			
(a) Sharpe Ratio (vs FMR)			
κ	6	36	360	3600
---	---	---	---	---
0.000	-0.1482	-0.0025	-1.5344	-1.3632
1e-6	-0.1482	0.0018	-0.3309	-0.3049
1e-5	-0.1481	0.0136	-0.0890	-0.0801
1e-4	-0.1474	0.0243	0.0479	**0.0657**
0.001	-0.1445	0.0043	0.0489	0.0576
0.010	-0.1488	-0.0290	-0.0155	-0.0123
(b) Hansen-Jagannathan Distance (vs FMR)			
κ	6	36	360	3600
0.000	0.0422	-0.0574	24.8450	0.7238
1e-6	0.0422	-0.0584	0.1318	0.1021
1e-5	0.0421	-0.0589	0.0032	-0.0162
1e-4	0.0418	-0.0529	-0.0643	**-0.0770**
0.001	0.0409	-0.0321	-0.0535	-0.0596

DKKM in GS Model

I. Sharpe Ratios
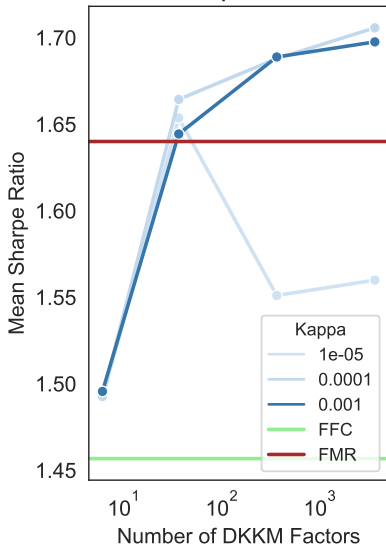
II. HJ Distances

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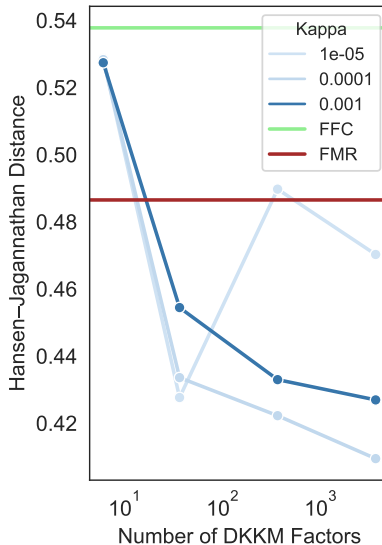
- Optimal with just 2–3 factors
- BGN and GS: 2 factors optimal
- KP: 3 factors optimal

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

- Usual goal in asset pricing is to find the tradeable SDF

- Usually compute factor portfolio weights from firm characteristics - sorts, regression, etc.

- FFC, FMR, and DKKM follow that process. DKKM is clearly better than FFC and FMR in the models we studied.

- KPS (Instrumented PCA) is a bit different in that it assumes latent factors with betas that are linear in characteristics. It is also clearly superior to FMR and FMC.

- Logical next step, which we're working on, is to do IPCA with a more flexible modeling of betas – linear in an exploded characteristic set as in DKKM.