Schrödinger's Sparsity in the Cross Section of Stock Returns

Doron Avramov[1], Guanhao Feng[2], Jingyu He[2] and Shuhua Xiao[2]

July 30, 2025

SoFiE Financial Machine Learning Summer School at Yale

[1]Reichman University [2]City University of Hong Kong

High-dimensional AP has two *different* modeling choices and assumptions.

- **Sparse modeling: L_1 penalty, Lasso regression**
  - Feng, Giglio, and Xiu (JF 2020), Freybergr, Neuhierl, and Weber (RFS 2020), and Bybee, Kelly, and Su (RFS 2023)
- **Dense modeling: L_2 penalty, Ridge regression**
  - Kelly, Pruitt, and Su (JFE 2019), Kozak, Nagel, and Santosh (JFE 2020)
  - Kozak and Nagel (WP 2023) — SDF requires a large number of characteristics.
  - Shen and Xiu (WP 2025) — Ridge outperforms Lasso when signals are weak.

Giannone, Lenza, and Primiceri (ECTA 2021) (GLP2021) develop a Bayesian sparse model that learns sparsity levels in linear regression.

- GLP2021 test six high-dimensional datasets (Macro/Finance/Micro) and find the posterior distribution rarely concentrates on a single sparse model.
  - Two applications more strongly favor dense models with a full set of predictors.
  - Only one application where the posterior density focuses on a sparse model.
  - Those posterior distributions behave quite differently across applications.

$\Rightarrow$ *illusion of sparsity*

- Link L_1 and L_2: **no assumption**, but posterior learning $\Rightarrow$ *Proportion of non-zero coefficients is unknown and can be learned.*

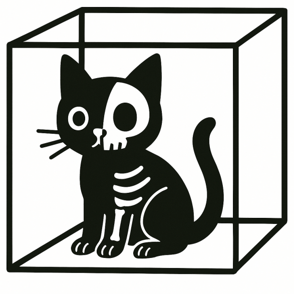
A Tale of Two Philosophies: Sparse vs Dense

- Traditional AP models demand an ex ante decision on sparsity or density.
- Empirical findings frequently mirror prior assumptions instead of revealing the actual structure of expected returns.

A Tale of Two Philosophies: Sparse vs Dense

- Traditional AP models demand an ex ante decision on sparsity or density.
- Empirical findings frequently mirror prior assumptions instead of revealing the actual structure of expected returns.
- Can sparsity be treated not as a fixed assumption, but as an inferred property of the data?

Challenge and Motivation: Schrödinger's Sparsity



Schrödinger's cat

- A cat, entangled with a quantum system, remains in a superposition of alive and dead states until observed.
- The nature of AP models — sparse or dense — are in a state of superposition until empirical data is observed.

We examine the sparsity of AP models within the conditional latent factor framework of IPCA. (It can also be tested in beta-pricing or SDF models.)

$$
\begin{aligned}
r_{i,t} &= \boldsymbol{\alpha}(\mathbf{Z}_{i,t-1}) + \boldsymbol{\beta}(\mathbf{Z}_{i,t-1})\mathbf{f}_t + \epsilon_{i,t} \\
\text{where} \quad \boldsymbol{\alpha}(\mathbf{Z}_{i,t-1}) &= \alpha_0 + \boldsymbol{\alpha}_1 \mathbf{Z}_{i,t-1} \\
\boldsymbol{\beta}(\mathbf{Z}_{i,t-1}) &= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1(\mathbf{I}_K \otimes \mathbf{Z}_{i,t-1}) \\
\epsilon_{i,t} &\sim \mathcal{N}\left(0, \sigma_i^2\right)
\end{aligned}
$$

- $\mathbf{f}_t$: K latent factors (can be extended to both observable and latent factors).
- $\mathbf{Z}_{i,t-1}$: L characteristics.

Research Questions

- Built on IPCA (Kelly, Pruitt, and Su, JFE 2019; Bybee, Kelly, and Su, RFS 2023) and Bayesian unconditional latent factor model (Geweke and Zhou, RFS 1996).

  — A New Perspective: probability of char sparsity

- Our focuses are the char-driven alphas (mispricings) and betas (loadings).

  — Our approach does not aim to benchmark against its frequentist counterpart but rather emphasizes targeted Bayesian model evaluation.

- We allow sparsity prob. to be data-inferred (following GLP2021) or exogenously fixed, enabling model performance evaluation under varying sparsity assumptions.

Empirical Highlights

- Best-performing models are neither extremely sparse nor dense.
- When the imposed sparsity matches the endogenous level inferred by the posterior, model performance peaks.

Empirical Highlights

- Best-performing models are neither extremely sparse nor dense.
- When the imposed sparsity matches the endogenous level inferred by the posterior, model performance peaks.
- Sparsity prob. change across test asset sets and positively correlate with pricing difficulty.
  - 5×5 ME-BM portfolios $\Rightarrow$ Sparse model
  - Alphas is generally sparser than betas.
  - Assets with higher Jensen's alpha $\sim$ denser mispricings.
  - Assets with higher Sharpe ratios $\sim$ denser loadings.
- Sparsity is time-varying. Models grow sparser during recessions.
- Models integrating observable and latent factors outperform.

Model

Spike-and-Slab Prior

Spike-and-slab prior, a Bayesian variable selection prior.

Let $d = 1$ or 0 denote selected or not selected, the spike and slab prior on β is

$$\beta \mid d \sim d\mathcal{N}\left(0, \xi_1^2\sigma^2\right) + (1-d)\mathcal{N}\left(0, \xi_0^2\sigma^2\right)$$

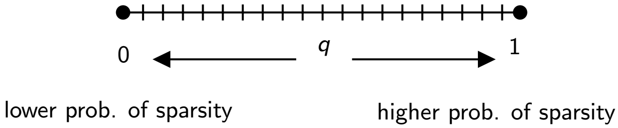
$$P(d=0) = 1 - P(d=1) = q$$

Hence, when ξ_1 is related large and ξ_0 shrinks to zero:

$$\beta = \begin{cases} 0 \text{ with prob. } q & \text{The regressor is not chosen.} \\ \mathcal{N}\left(0, \gamma^2\right) \text{ with prob. } 1-q & \text{The regressor is chosen.} \end{cases}$$

$$\beta = \begin{cases} 0 \text{ with prob. } q & \text{The regressor is not chosen.} \\ \mathcal{N}\left(0, \gamma^2\right) \text{ with prob. } 1-q & \text{The regressor is chosen.} \end{cases}$$

- Standard spike-and-slab prior: q is a specific value.
- GLP2021: q has its prior so that one can sample: $q \sim \text{Beta}(a, b)$
  - These priors probabilistically balance variable selection and shrinkage.



- **Prior settings of $q \neq$ precise control of sparsity levels!**

Priors Over Sparsity Structures

$$r_{i,t} = \alpha_0 + \boldsymbol{\alpha}_1 \mathbf{Z}_{i,t-1} + \boldsymbol{\beta}_0 \mathbf{f}_t + \boldsymbol{\beta}_1 [\mathbf{f}_t \otimes \mathbf{Z}_{i,t-1}] + \epsilon_{i,t}.$$

- Independent spike-and-slab priors on $\boldsymbol{\alpha}_1$ and $\boldsymbol{\beta}_1$
- **Separate priors**: different sparsity levels of alpha and beta.

$$[\boldsymbol{\alpha}_1]_l \sim \begin{cases} \mathcal{N}(0, \gamma_\alpha^2) & \text{if } d_l^\alpha = 1 \\ 0 & \text{if } d_l^\alpha = 0 \end{cases} \qquad [\boldsymbol{\beta}_1]_l \sim \begin{cases} \mathcal{N}(0, \gamma_\beta^2) & \text{if } d_l^\beta = 1 \\ 0 & \text{if } d_l^\beta = 0 \end{cases}$$

$$d_l^\alpha \sim \text{Bernoulli}(1 - q_\alpha) \qquad d_l^\beta \sim \text{Bernoulli}(1 - q_\beta)$$

$$q_\alpha \sim \text{Beta}(a_{q_\alpha}, b_{q_\alpha}) \qquad q_\beta \sim \text{Beta}(a_{q_\beta}, b_{q_\beta})$$

$$\gamma_\alpha^2 \sim \mathcal{IG}\left(A_{\gamma_\alpha}/2, B_{\gamma_\alpha}/2\right) \qquad \gamma_\beta^2 \sim \mathcal{IG}\left(A_{\gamma_\beta}/2, B_{\gamma_\beta}/2\right)$$

- Higher posterior mean of q_α (or q_β), higher prob. of sparsity.

$$r_{i,t} = \alpha_0 + \boldsymbol{\alpha}_1 \mathbf{Z}_{i,t-1} + \boldsymbol{\beta}_0 \mathbf{f}_t + \boldsymbol{\beta}_1 [\mathbf{f}_t \otimes \mathbf{Z}_{i,t-1}] + \epsilon_{i,t}.$$

- Directly control the sparsity level (i.e., control # selected char.).

  M_α and M_β restrict the number of char. driving alpha and beta.

- **(Separate) joint priors**:

$$(d_1^\alpha, d_2^\alpha, \cdots, d_L^\alpha) \sim \left[\prod_{l=1}^{L} \text{Bernoulli}(1 - q_\alpha)\right] \times \mathbf{I}\left(\sum_{l=1}^{L} d_l = M_\alpha\right),$$

$$(d_1^\beta, d_2^\beta, \cdots, d_L^\beta) \sim \left[\prod_{l=1}^{L} \text{Bernoulli}(1 - q_\beta)\right] \times \mathbf{I}\left(\sum_{l=1}^{L} d_l = M_\beta\right).$$

- Larger M_α (or M_β), lower **sparsity level**.

We follow Kelly, Pruitt, and Su (JFE 2019).

- $\Gamma_\alpha = [\alpha_0, \boldsymbol{\alpha}_1]$ and $\Gamma_\beta = [\boldsymbol{\beta}_0, \boldsymbol{\beta}_1]$.
- $\Gamma_\beta^\top \Gamma_\beta = \mathbf{I}_K$:
  The unconditional second-moment matrix of $\mathbf{f}_t$ is diagonal with descending diagonal entries, and restricts the mean of $\mathbf{f}_t$ to be non-negative.
  - To preserve the structure of Γ_β: impose these constraints on each factor.
- $\Gamma_\alpha^\top \Gamma_\beta = \mathbf{0}_{1\times K}$
  - Regressing Γ_α on Γ_β and replacing Γ_α with the residual from this regression.

Schrödinger's Sparsity

Main test assets:

- P-Tree (Cong, Feng, He, and He, JFE 2025) test assets (1990-2024)
  - Sequential decreasing alphas by boosted trees
  - Constructed based on the past sample (1980-1989)

Other test assets:

- Portfolios
  - 25 ME/BM portfolios
  - 360 bivariate-sorted portfolios
  - 610 univariate-sorted portfolios
- Individual stocks
  - stocks ranked 1st to 500th by average market equity (500 Big)
  - stocks ranked 501st-1000th by average market equity (500 Small)

(i) Probability of Sparsity

Table 1: Model Performance Under Different Priors

		CSR^2			(q_α, q_β)		
		$K=1$	$K=3$	$K=5$	$K=1$	$K=3$	$K=5$
Panel A: Unrestricted # sel char.							
(q_α prior mean, q_β prior mean)	0.9,0.9	29.2	43.7	58.9	0.66,0.62	0.83,0.57	0.93,0.64
	0.5,0.9	29.4	43.4	57.0	0.51,0.62	0.68,0.60	0.77,0.64
	0.1,0.9	29.4	43.1	56.6	0.37,0.62	0.53,0.60	0.63,0.66
	0.9,0.5	29.3	44.3	59.9	0.66,0.47	0.82,0.47	0.93,0.50
	0.5,0.5	29.5	42.4	58.8	0.52,0.47	0.69,0.43	0.79,0.50
	0.1,0.5	29.5	43.6	58.1	0.37,0.47	0.53,0.46	0.64,0.49
	0.9,0.1	29.5	46.9	58.3	0.66,0.31	0.83,0.31	0.92,0.33
	0.5,0.1	29.6	42.5	57.9	0.52,0.31	0.69,0.30	0.79,0.34
	0.1,0.1	29.7	45.6	53.7	0.37,0.31	0.54,0.31	0.62,0.35
Panel B: Fixed # sel char.		...					
Panel C: No sparsity							
(M_α, M_β)	20,20	29.9	36.9	45.2	/	/	/

CSR^2 Benchmark: CAPM

(i) Probability of Sparsity

- **Prob. of sparsity**:
  - Between the extremes of highly sparse (prob. close to 1) and fully dense (prob. close to 0) specifications.
  - Mispricings show higher sparsity than loadings — theoretically complementarity
  - Sparsity patterns change with the number of latent factors K.

    $K \uparrow$, q_β decreases or remains low, q_α increases substantially.
- **Models' performance**:
  - Robust across prior settings and increasing with more factors.
  - Models incorporating probabilistic sparsity consistently outperform both fully dense specifications (Panel C).

(ii) Effect of Misspecified Assumption of Sparsity

Table 1: Model Performance Under Different Priors

		CSR^2			(q_α, q_β)		
		$K=1$	$K=3$	$K=5$	$K=1$	$K=3$	$K=5$
Panel A: Unrestricted # sel char.							
(q_α prior mean, q_β prior mean)	0.9,0.9	29.2	43.7	58.9	0.66,0.62	0.83,0.57	0.93,0.64
	0.5,0.9	29.4	43.4	57.0	0.51,0.62	0.68,0.60	0.77,0.64
	0.1,0.9	29.4	43.1	56.6	0.37,0.62	0.53,0.60	0.63,0.66
	0.9,0.5	29.3	44.3	59.9	0.66,0.47	0.82,0.47	0.93,0.50
	0.5,0.5	29.5	42.4	58.8	0.52,0.47	0.69,0.43	0.79,0.50
	0.1,0.5	29.5	43.6	58.1	0.37,0.47	0.53,0.46	0.64,0.49
	0.9,0.1	29.5	46.9	58.3	0.66,0.31	0.83,0.31	0.92,0.33
	0.5,0.1	29.6	42.5	57.9	0.52,0.31	0.69,0.30	0.79,0.34
	0.1,0.1	29.7	45.6	53.7	0.37,0.31	0.54,0.31	0.62,0.35
Panel B: Fixed # selected char.							
(M_α, M_β)	2,2	25.4	49.3	48.4	/	/	/
	10,2	28.0	51.1	50.0	/	/	/
	18,2	25.2	46.9	37.8	/	/	/
	2,10	28.8	50.9	59.6	/	/	/
	10,10	29.6	38.3	41.1	/	/	/
	18,10	27.2	40.9	39.5	/	/	/
	2,18	29.8	54.9	56.1	/	/	/
	10,18	29.9	34.5	51.0	/	/	/
	18,18	27.5	39.3	42.1	/	/	/

(ii) Effect of Misspecified Assumption of Sparsity

- Fixing M_α and M_β significantly affect model performance.

- Model performance peaks when fixed inclusion sizes in the constrained model match the sparsity levels of the probabilistic model.
  - $K = 5$, the best model is inferred at a sparsity level near $(M_\alpha, M_\beta) = (1, 10)$, while the best fixed model is at $(2, 10)$.

Learn rather than impose sparsity in conditional asset pricing models.

(iii) Excluding Mispricing

$$r_{i,t} = \boldsymbol{\beta}_0 \mathbf{f}_t + \boldsymbol{\beta}_1 [\mathbf{f}_t \otimes \mathbf{Z}_{i,t-1}] + \epsilon_{i,t}.$$

Table 2: Model Performance under Different Priors (without Mispricing)

		CSR^2			q_β		
		$K=1$	$K=3$	$K=5$	$K=1$	$K=3$	$K=5$
Panel A: Unrestricted # sel char.							
	0.9	20.4	52.8	58.1	0.62	0.59	0.60
q_β prior mean	0.5	20.5	53.2	58.3	0.48	0.43	0.50
	0.1	20.6	53.6	60.1	0.32	0.27	0.26
Panel B: Fixed # sel char.							
	2	15.6	49.5	50.4	/	/	/
M_β	10	20.1	52.1	58.5	/	/	/
	18	19.9	53.8	60.6	/	/	/

- With a prior mean of 0.9 ($K=5$), the posterior q_β is 0.64–0.66 with mispricing and 0.60 without.

$\Rightarrow$ Without a mispricing channel, loadings must capture more return variation, requiring a denser specification.

Schrödinger's Sparsity Everywhere

(i) Test Assets and Sparsity

Table 3: Sparsity for Different Test Assets

	CSR^2	TP. Sp	(q_α, q_β)
Panel A: P-Tree			
100	42.4	1.00	0.69,0.43
200	51.0	1.09	0.60,0.37
400	45.2	0.49	0.54,0.32
Panel B: Ind. Stock			
500 big	31.4	0.80	0.61,0.29
500 small	3.9	3.64	0.49,0.38
Panel C: Others			
ME/BM25	33.6	0.25	0.80,0.50
Bi360	7.8	1.03	0.50,0.20
Uni610	48.0	0.61	0.44,0.20

- Sparsity levels change across different types of test assets.

  E.g., ME/BM 25 sparser.

Table 3: Sparsity for Different Test Assets

	CSR^2	TP. Sp	(q_α, q_β)
Panel A: P-Tree			
100	42.4	1.00	0.69,0.43
200	51.0	1.09	0.60,0.37
400	45.2	0.49	0.54,0.32
Panel B: Ind. Stock			
500 big	31.4	0.80	0.61,0.29
500 small	3.9	3.64	0.49,0.38
Panel C: Others			
ME/BM25	33.6	0.25	0.80,0.50
Bi360	7.8	1.03	0.50,0.20
Uni610	48.0	0.61	0.44,0.20

- Within the same category of test assets, a larger number of assets generally requires more char.

(i) Test Assets and Sparsity

Table 3: Sparsity for Different Test Assets

	CSR^2	TP. Sp	(q_α, q_β)
Panel A: P-Tree			
100	42.4	1.00	0.69,0.43
200	51.0	1.09	0.60,0.37
400	45.2	0.49	0.54,0.32
Panel B: Ind. Stock			
500 big	31.4	0.80	0.61,0.29
500 small	3.9	3.64	0.49,0.38
Panel C: Others			
ME/BM25	33.6	0.25	0.80,0.50
Bi360	7.8	1.03	0.50,0.20
Uni610	48.0	0.61	0.44,0.20

- Those assets that are more difficult to price tend to require more char. to capture alpha.

(i) Test Assets and Sparsity

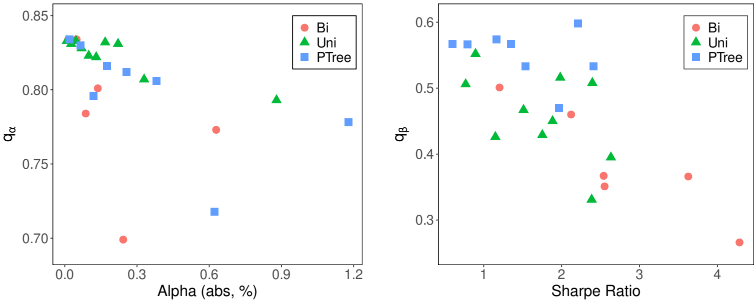
Table 3: Sparsity for Different Test Assets

	CSR^2	TP. Sp	(q_α, q_β)
Panel A: P-Tree			
100	42.4	1.00	0.69,0.43
200	51.0	1.09	0.60,0.37
400	45.2	0.49	0.54,0.32
Panel B: Ind. Stock			
500 big	31.4	0.80	0.61,0.29
500 small	3.9	3.64	0.49,0.38
Panel C: Others			
ME/BM25	33.6	0.25	0.80,0.50
Bi360	7.8	1.03	0.50,0.20
Uni610	48.0	0.61	0.44,0.20

- Substantial heterogeneity in sparsity probabilities across standard test assets.

Alpha, Sharpe ratio and sparsity

Figure 1: Sparsity Levels and Pricing Difficulty of Test Assets



(a) q_α and Alpha **(b)** q_β and Sharpe Ratio

- Sparsity levels vary across test assets, reflecting pricing difficulty differences.

(ii) Marcro Regimes and Sparsity Probability

Table 4: Time Variation Analysis: Sparsity in Structural Breaks / Business Cycles

	CSR^2	TP. Sp	(q_α, q_β)
Panel A: Sequential segmentation			
Regime1	48.5	1.92	0.72,0.56
Regime2	24.1	0.82	0.71,0.53
Regime3	59.7	0.72	0.77,0.46
Panel B: Macro-driven segmentation			
Normal	53.8	1.18	0.67,0.46
Recession	14.2	0.77	0.76,0.50
Panel C: Full period			
Whole	42.4	1.00	0.69,0.43

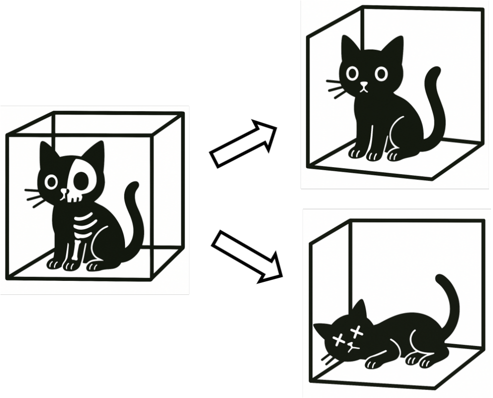
- Settings of time periods:
  - Breakpoints in Smith and Timmermann (RFS 2021): July 1998 and June 2010.
  - Define recession periods based on the Sahm Rule (88 months).
- AP models tend to be sparser during recessions.

- **Recession-induced sparsity**
  - Heightened market uncertainty:
    Investors tend to focus on macroeconomic and systematic risks
  - Large macroeconomic shocks:
    Disrupt the relationship between firm char. and return variation
- **Regime-based char. importance**
  - During recessionary periods, both the bid-ask spread and 12-month momentum exhibit substantially lower posterior inclusion prob. for mispricing.
  - Characteristics such as asset growth , net stock issues , operating profitability and R&D-to-market equity lose their ability to explain systematic risk.

Schrödinger's Sparsity Everywhere!

Sparsity Prob. change across both **cross-sectional** and **time-series** dimensions.

$\Rightarrow$ i) Test assets / Pricing difficulty; ii) Time periods / Macro conditions



Assuming AP model to be either sparse or dense ex ante may be wrong.

Conditional Model for Observable Factors and Sparsity

Model with Observable and Latent Factors

- In the conditional observable factor model, alpha and beta are sparse functions of high-dimensional char.
- Augmenting latent factors aids in recovering unspanned risk factors within observable factor models.

$$r_{i,t} = \boldsymbol{\alpha}(\mathbf{Z}_{i,t-1}) + \boldsymbol{\beta}(\mathbf{Z}_{i,t-1}) \underbrace{\left[\mathbf{f}_t^O, \mathbf{f}_t^L\right]}_{\mathbf{F}_t} + \epsilon_{i,t}$$

$$= \underbrace{\alpha_0 + \boldsymbol{\alpha}_1 \mathbf{Z}_{i,t-1}}_{\text{mispricing}} + \underbrace{\boldsymbol{\beta}_0 \mathbf{f}_t^O + \boldsymbol{\beta}_1 [\mathbf{f}_t^O \otimes \mathbf{Z}_{i,t-1}]}_{\text{obs. factors, conditional beta}} + \underbrace{\boldsymbol{\beta}_0 \mathbf{f}_t^L + \boldsymbol{\beta}_1 [\mathbf{f}_t^L \otimes \mathbf{z}_{i,t-1}]}_{\text{latent factors, dynamic loadings}} + \epsilon_{i,t}.$$

Table 5: Augmented Observable Factor Models

	CSR^2	TP.Sp	(q_α, q_β)	$\alpha RMSE$
Panel A: only obs				
MKT	14.9	0.57	0.55,0.37	0.0032
FF3	27.3	0.60	0.65,0.26	0.0026
FF5	50.4	1.13	0.74,0.39	0.0014
Panel B: only latent				
LF1	29.5	0.35	0.52,0.47	0.0036
LF3	45.0	1.00	0.68,0.58	0.0021
LF5	56.8	1.02	0.77,0.66	0.0011
Panel C: obs + latent				
MKT+LF1	53.9	0.87	0.69,0.35	0.0015
MKT+LF5	56.5	1.06	0.79,0.48	0.0005
FF3+LF1	41.6	1.07	0.67,0.27	0.0014
FF3+LF5	57.4	1.26	0.80,0.56	0.0003
FF5+LF1	50.6	1.21	0.67,0.35	0.0012
FF5+LF5	55.8	1.25	0.79,0.58	0.0005
Panel D: uncond. model				
MKT	/	0.57	/	0.0060
FF3	11.5	0.60	/	0.0056
FF5	49.3	1.13	/	0.0042

- A vs D: The informational value of conditioning: Our framework uncovers pricing structures via latent factor models and adaptive sparsity.

Table 5: Augmented Observable Factor Models

	CSR^2	TP.Sp	(q_α, q_β)	$\alpha RMSE$
Panel A: only obs				
MKT	14.9	0.57	0.55,0.37	0.0032
FF3	27.3	0.60	0.65,0.26	0.0026
FF5	50.4	1.13	0.74,0.39	0.0014
Panel B: only latent				
LF1	29.5	0.35	0.52,0.47	0.0036
LF3	45.0	1.00	0.68,0.58	0.0021
LF5	56.8	1.02	0.77,0.66	0.0011
Panel C: obs + latent				
MKT+LF1	53.9	0.87	0.69,0.35	0.0015
MKT+LF5	56.5	1.06	0.79,0.48	0.0005
FF3+LF1	41.6	1.07	0.67,0.27	0.0014
FF3+LF5	57.4	1.26	0.80,0.56	0.0003
FF5+LF1	50.6	1.21	0.67,0.35	0.0012
FF5+LF5	55.8	1.25	0.79,0.58	0.0005
Panel D: uncond. model				
MKT	/	0.57	/	0.0060
FF3	11.5	0.60	/	0.0056
FF5	49.3	1.13	/	0.0042

- A vs B: more variation in returns is captured by latent factors ⇒ Reducing # characteristics required to explain these components.

Table 5: Augmented Observable Factor Models

	CSR^2	TP.Sp	(q_α, q_β)	$\alpha RMSE$
Panel A: only obs				
MKT	14.9	0.57	0.55,0.37	0.0032
FF3	27.3	0.60	0.65,0.26	0.0026
FF5	50.4	1.13	0.74,0.39	0.0014
Panel B: only latent				
LF1	29.5	0.35	0.52,0.47	0.0036
LF3	45.0	1.00	0.68,0.58	0.0021
LF5	56.8	1.02	0.77,0.66	0.0011
Panel C: obs + latent				
MKT+LF1	53.9	0.87	0.69,0.35	0.0015
MKT+LF5	56.5	1.06	0.79,0.48	0.0005
FF3+LF1	41.6	1.07	0.67,0.27	0.0014
FF3+LF5	57.4	1.26	0.80,0.56	0.0003
FF5+LF1	50.6	1.21	0.67,0.35	0.0012
FF5+LF5	55.8	1.25	0.79,0.58	0.0005
Panel D: uncond. model				
MKT	/	0.57	/	0.0060
FF3	11.5	0.60	/	0.0056
FF5	49.3	1.13	/	0.0042

- A vs C: Adding latent factors helps mitigate model misspecification.
  - α RMSE: decreases after introducing latent factors.

Summary

- An important problem: **How can researchers determine the appropriate model assumption without first examining the data?**

  $\Rightarrow$ Schrödinger's Sparsity: the true state remains unknowable until observed

  - treating sparsity as a probabilistic property rather than a binary assumption

- A new approach, a flexible Bayesian framework for IPCA
  - Utilizing the independent/joint **spike-and-slab priors**
  - **Endogenously** determine whether the model is sparse or dense, without imposing prior assumptions on sparsity or density.
  - **Exogenously** control the sparsity level of the model.

Summary

- An important problem: **How can researchers determine the appropriate model assumption without first examining the data?**

  ⇒ Schrödinger's Sparsity: the true state remains unknowable until observed

  - treating sparsity as a probabilistic property rather than a binary assumption

- Empirical findings:
  - Well-performing models often lie between the extremes of full sparsity and full density.
  - Sparsity prob. change with the nature of the test assets
    - Sparsity is linked to pricing difficulty.
  - Sparsity increases during recessions, as fewer firm char. remain relevant, and macroeconomic risk becomes more dominant.

⇒ How, when, and why firm characteristics matter in the cross section of returns.