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

Svetlana Bryzgalova, Jiantao Huang,Christian Julliard.

Working paper, 2021

解读者：屠雪永

2021.11.21

Outline

- Introduction
- Research design
  - Frequentist Estimation of Linear SDFs
  - Bayesian Analysis of Linear SDFs
- Simulation
- Empirical study
- Conclusion

1. Introduction-- Motivation

- Two observations in the the empirical asset pricing literature
  - The factor zoo phenomenon
  - The so-called weak factors
- No general method has been suggested:
  - i) is applicable to both tradable and non-tradable factors,
  - ii) can handle the entire factor zoo,
  - iii) remains valid under misspecification,
  - iv) is robust to the weak inference problem,
  - v) delivers an empirical pricing kernel that outperforms (in- and out-of-sample) popular models (with either observable or latent factors).

1. Introduction-- Related Literature

- **Asset allocation**
  - Shanken (1987) and Harvey and Zhou (1990)-first
  - Pastor and Stambaugh (2000) and Pastor (2000)--prior distribution
- **Model selection、parameter estimation and shrinkage-based one**
  - Garlappi, Uppal, and Wang (2007)、Kozak, Nagel, and Santosh (2020)
- **Weak identification**
  - Kan and Zhang (1999a,b)、Gospodinov, Kan, and Robotti (2014, 2019)
- **Performance evaluation**
  - Harvey, Liu, and Zhu (2016)、Giglio,Feng, and Xiu (2020)

1. Introduction-- Framework

$$M_t = 1 - (\boldsymbol{f}_t - E[\boldsymbol{f}_t])^\top \boldsymbol{\lambda}_{\boldsymbol{f}}$$

$$E[M_t \boldsymbol{R}_t] = \boldsymbol{0}_N$$

Frequentist Estimation of Linear SDFs

GMM

weak identification

Bayesian Analysis of Linear SDFs

Correct SDF specification

$\boldsymbol{\mu_R} = \boldsymbol{C\lambda}$

Average pricing errors

$\boldsymbol{\mu_R} = \boldsymbol{C\lambda} + \boldsymbol{\alpha}$

Flat Priors for Risk Prices

$$\pi(\boldsymbol{\lambda}, \sigma^2) \propto \sigma^{-2}$$

Spike-and-Slab Prior

$$\pi(\boldsymbol{\lambda}, \sigma^2, \boldsymbol{\gamma}) \propto \pi(\boldsymbol{\lambda} | \sigma^2, \boldsymbol{\gamma}) \pi(\sigma^2) \pi(\boldsymbol{\gamma})$$

Continuous Spike

$$\pi(\boldsymbol{\lambda}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega}) = \pi(\boldsymbol{\lambda} \mid \sigma^2, \boldsymbol{\gamma}) \pi(\sigma^2) \pi(\boldsymbol{\gamma} \mid \boldsymbol{\omega}) \pi(\boldsymbol{\omega})$$

1. Introduction-- Contribution

- We propose a **novel framework** for analyzing linear asset pricing models: simple, robust, and applicable to **high dimensional** problems.
- It provides reliable price of risk estimates for **both tradable and non-tradable factors**, and detects those **weakly identified**.
- The method automatically selects **the best specification** – if a dominant one exists – or provides a Bayesian model averaging (BMA-SDF), if there is no clear winner.

2.1. Frequentist Estimation of Linear SDFs

- K factors $\boldsymbol{f_t} = (f_{1t} \dots f_{kt})^\top, t = 1, \dots T$. N test assets $\boldsymbol{R}_t = (R_{1t} \dots R_{Nt})^\top$
- linear stochastic discount factors $M_t = 1 - (\boldsymbol{f}_t - E[\boldsymbol{f}_t])^\top \boldsymbol{\lambda_f}$

$$E[M_t \boldsymbol{R_t}] = \mathbf{0_N}$$

$$\boldsymbol{\mu_R} = \mathbb{E}[\boldsymbol{R}_t] = \boldsymbol{C_f} \boldsymbol{\lambda_f}$$

- $\boldsymbol{C}_f$: the covariance matrix between $\boldsymbol{R}_t$ and $\boldsymbol{f_t}$
- $\boldsymbol{\lambda_f}$: vector of prices of risk of the factors, estimated via

$$\boldsymbol{\mu_R} = \lambda_c \mathbf{1_N} + \boldsymbol{C_f} \boldsymbol{\lambda_f} + \boldsymbol{\alpha} = \boldsymbol{C} \boldsymbol{\lambda} + \boldsymbol{\alpha}$$

- Estimated via GMM

$$\mathbb{E}[\boldsymbol{g_t}(\lambda_c, \boldsymbol{\lambda_f}, \boldsymbol{\mu_f})] = \mathbb{E}\begin{pmatrix} \boldsymbol{R_t} - \lambda_c \mathbf{1_N} - \boldsymbol{R_t}(\boldsymbol{f_t} - \boldsymbol{\mu_f})^\top \boldsymbol{\lambda_f} \\ \boldsymbol{f_t} - \boldsymbol{\mu_f} \end{pmatrix} = \begin{pmatrix} \mathbf{0_N} \\ \mathbf{0_K} \end{pmatrix}$$

$$\boldsymbol{g_T}(\lambda_c, \boldsymbol{\lambda_f}, \boldsymbol{\mu_f}) \equiv \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{g_t}(\lambda_c, \boldsymbol{\lambda_f}, \boldsymbol{\mu_f})$$

$$\{\widehat{\lambda}_c, \widehat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \widehat{\boldsymbol{\mu}}_{\boldsymbol{f}}\} \equiv \underset{\lambda_c, \boldsymbol{\lambda_f}, \boldsymbol{\mu_f}}{\arg\min} \; \boldsymbol{g_T}(\lambda_c, \boldsymbol{\lambda_f}, \boldsymbol{\mu_f})^\top \boldsymbol{W} \boldsymbol{g_T}(\lambda_c, \boldsymbol{\lambda_f}, \boldsymbol{\mu_f}).$$

2.1. Frequentist Estimation of Linear SDFs

$$\{\widehat{\lambda}_c, \widehat{\boldsymbol{\lambda}}_{\boldsymbol{f}}, \widehat{\boldsymbol{\mu}}_{\boldsymbol{f}}\} \equiv \underset{\lambda_c, \boldsymbol{\lambda}_{\boldsymbol{f}}, \boldsymbol{\mu}_{\boldsymbol{f}}}{\arg\min}\ \boldsymbol{g}_{\boldsymbol{T}}(\lambda_c, \boldsymbol{\lambda}_{\boldsymbol{f}}, \boldsymbol{\mu}_{\boldsymbol{f}})^\top \boldsymbol{W} \boldsymbol{g}_{\boldsymbol{T}}(\lambda_c, \boldsymbol{\lambda}_{\boldsymbol{f}}, \boldsymbol{\mu}_{\boldsymbol{f}}).$$

- $$\boldsymbol{W}_{\boldsymbol{ols}} = \begin{pmatrix} \boldsymbol{I}_{\boldsymbol{N}} & \boldsymbol{0}_{\boldsymbol{N}\times\boldsymbol{K}} \\ \boldsymbol{0}_{\boldsymbol{K}\times\boldsymbol{N}} & \kappa \boldsymbol{I}_{\boldsymbol{K}}, \end{pmatrix}, \text{ and } \boldsymbol{W}_{\boldsymbol{gls}} = \begin{pmatrix} \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} & \boldsymbol{0}_{\boldsymbol{N}\times\boldsymbol{K}} \\ \boldsymbol{0}_{\boldsymbol{K}\times\boldsymbol{N}} & \kappa \boldsymbol{I}_{\boldsymbol{K}} \end{pmatrix}$$
- $\boldsymbol{\Sigma}_{\boldsymbol{R}}$ is the covarince matrix of returns, and κ > 0 is a large constant

$$\widehat{\boldsymbol{\lambda}}_{\boldsymbol{ols}} = (\widehat{\boldsymbol{C}}^\top \widehat{\boldsymbol{C}})^{-1} \widehat{\boldsymbol{C}}^\top \bar{\boldsymbol{R}}$$

$$\widehat{\boldsymbol{\lambda}}_{\boldsymbol{gls}} = (\widehat{\boldsymbol{C}}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \widehat{\boldsymbol{C}})^{-1} \widehat{\boldsymbol{C}}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \bar{\boldsymbol{R}}$$

where $\widehat{\boldsymbol{C}} = (\boldsymbol{1}_{\boldsymbol{N}}, \widehat{\boldsymbol{C}}_{\boldsymbol{f}})$ and $\widehat{\boldsymbol{C}}_{\boldsymbol{f}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{R}_{\boldsymbol{t}}(\boldsymbol{f}_{\boldsymbol{t}} - \widehat{\boldsymbol{\mu}}_{\boldsymbol{f}})^\top$

- GMM require factor exposures $\boldsymbol{C}$ to have full rank
- Problem arises: **weak factor**

2.2. Bayesian Analysis of Linear SDFs

- The K_1 tradable factors first ($\boldsymbol{f}_t^{(1)}$), K_2 non−tradable factors($\boldsymbol{f}_t^{(2)}$) ,
- $\boldsymbol{f} \equiv (\boldsymbol{f}_t^{(1),\top}, \boldsymbol{f}_t^{(2),\top})^\top$ and $K_1 + K_2 = K$.
- $\boldsymbol{Y_t} \equiv \boldsymbol{f_t} \cup \boldsymbol{R_t}$ *,Y_t is a $p = N + K$ dimensional vector*
- $\boldsymbol{Y_t} \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{\mu_Y}, \boldsymbol{\Sigma_Y})$, $\mathbf{Y} \equiv \{\boldsymbol{Y_t}\}_{t=1}^T$
- Likelihood function

$$p(\mathbf{Y}|\boldsymbol{\mu_Y}, \boldsymbol{\Sigma_Y}) \propto |\boldsymbol{\Sigma_Y}|^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}tr\left[\boldsymbol{\Sigma_Y^{-1}} \sum_{t=1}^{T} (\boldsymbol{Y_t} - \boldsymbol{\mu_Y})(\boldsymbol{Y_t} - \boldsymbol{\mu_Y})^\top\right]\right\},$$

- *The diffuse prior:* $\pi(\boldsymbol{\mu_Y}, \boldsymbol{\Sigma_Y}) \propto |\boldsymbol{\Sigma_Y}|^{-\frac{p+1}{2}}$

$$P(B/A) = \frac{P(A/B)P(B)}{P(A)}$$

- *Posterior distribution :* $\boldsymbol{\mu_Y}|\boldsymbol{\Sigma_Y}, \mathbf{Y} \sim \mathcal{N}(\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}}, \; \boldsymbol{\Sigma_Y}/T)$

$$\boldsymbol{\Sigma_Y}|\mathbf{Y} \sim \mathcal{W}^{-1}\left(T-1, \; \sum_{t=1}^{T} (\boldsymbol{Y_t} - \widehat{\boldsymbol{\mu}}_{\boldsymbol{Y}})(\boldsymbol{Y_t} - \widehat{\boldsymbol{\mu}}_{\boldsymbol{Y}})^\top\right)$$

where $\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}} \equiv \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{Y_t}$ and $\mathcal{W}^{-1}$ is the inverse-Wishart distribution

2.2. Bayesian Analysis of Linear SDFs

$$\boldsymbol{\mu_R} = \boldsymbol{C\lambda}$$

Definition 1 (Bayesian SDF (B-SDF) Estimates) *Conditional on $\boldsymbol{\mu_Y}$, $\boldsymbol{\Sigma_Y}$ and the data $\mathbf{Y} = \{\boldsymbol{Y_t}\}_{t=1}^{T}$, under the null of unique correct SDF specification*[7] *and any diffuse prior, the posterior distribution of $\boldsymbol{\lambda}$ is a Dirac distribution (that is, a constant) at $(\boldsymbol{C}^\top \boldsymbol{C})^{-1}\boldsymbol{C}^\top \boldsymbol{\mu_R}$. Therefore, conditional on only the data $\mathbf{Y} = \{\boldsymbol{Y_t}\}_{t=1}^{T}$ and the null, the posterior distribution of $\boldsymbol{\lambda}$ can be sampled by drawing $\boldsymbol{\mu}_{Y,(j)}$ and $\boldsymbol{\Sigma}_{Y,(j)}$ from the Normal-inverse-Wishart (6)–(7) and computing the draw* $\boxed{\boldsymbol{\lambda}_{(j)} \equiv \left(\boldsymbol{C}_{(j)}^\top \boldsymbol{C}_{(j)}\right)^{-1} \boldsymbol{C}_{(j)}^\top \boldsymbol{\mu}_{R,(j)}.}$

- Estimating the cross−sectional fit of the model

$$R_{ols}^2 = 1 - \frac{(\boldsymbol{\mu_R} - \boldsymbol{C\lambda})^\top(\boldsymbol{\mu_R} - \boldsymbol{C\lambda})}{(\boldsymbol{\mu_R} - \bar{\mu}_R \mathbf{1_N})^\top(\boldsymbol{\mu_R} - \bar{\mu}_R \mathbf{1_N})},$$

where $\bar{\mu}_R = \frac{1}{N}\sum_i^N \mu_{R,i}$

2.2. Bayesian Analysis of Linear SDFs

Definition 2 (Bayesian SDF GLS (B-SDF-GLS)) *Conditional on $\boldsymbol{\mu}_{\mathbf{Y}}$, $\boldsymbol{\Sigma}_{\mathbf{Y}}$ and the data $\mathbf{Y} = \{\mathbf{Y}_t\}_{t=1}^{T}$, under the null of unique correct SDF specification and any diffuse prior, the posterior distribution of $\boldsymbol{\lambda}$ is a Dirac distribution (that is, a constant) at $(\boldsymbol{C}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{C})^{-1} \boldsymbol{C}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{\mu}_{\boldsymbol{R}}$. Therefore, conditional on only the data $\mathbf{Y} = \{\mathbf{Y}_t\}_{t=1}^{T}$ and the null, the posterior distribution of $\boldsymbol{\lambda}$ can be sampled by drawing $\boldsymbol{\mu}_{Y,(j)}$ and $\boldsymbol{\Sigma}_{Y,(j)}$ from the Normal-inverse-Wishart (6)–(7) and computing* $\boxed{\boldsymbol{\lambda}_{(j)} \equiv (\boldsymbol{C}_{(j)}^\top \boldsymbol{\Sigma}_{\boldsymbol{R},(j)}^{-1} \boldsymbol{C}_{(j)})^{-1} \boldsymbol{C}_{(j)}^\top \boldsymbol{\Sigma}_{\boldsymbol{R},(j)}^{-1} \boldsymbol{\mu}_{R,(j)}.}$

- Estimating the cross−sectional fit of the model

$$R^2_{gls} = 1 - \frac{(\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}\boldsymbol{\lambda})^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} (\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}\boldsymbol{\lambda})}{(\boldsymbol{\mu}_{\boldsymbol{R}} - \bar{\mu}_R \mathbf{1}_{\boldsymbol{N}})^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} (\boldsymbol{\mu}_{\boldsymbol{R}} - \bar{\mu}_R \mathbf{1}_{\boldsymbol{N}})}$$

where $\bar{\mu}_R = \frac{1}{N} \sum_i^N \mu_{R,i}$

2.2. Bayesian Analysis of Linear SDFs

- We now allow for mode limplied **average pricing errors**, α, $\alpha_i \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$
- $\boldsymbol{\mu_R} = \boldsymbol{C\lambda} + \boldsymbol{\alpha}$
- cross-sectional likelihood function

$$p(data|\boldsymbol{\lambda}, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{\mu_R} - \boldsymbol{C\lambda})^\top(\boldsymbol{\mu_R} - \boldsymbol{C\lambda})\right\}$$

- the "data" are the expected risk premia, $\boldsymbol{\mu_R}$, and the factor loadings, $\boldsymbol{C}$
- We assume a diffuse prior for $(\boldsymbol{\lambda}, \sigma^2)$: $\pi(\boldsymbol{\lambda}, \sigma^2) \propto \sigma^{-2}$
- Posterior distribution of $(\boldsymbol{\lambda}, \sigma^2)$

$$\boldsymbol{\lambda}|\sigma^2, data \sim \mathcal{N}\left(\underbrace{(\boldsymbol{C}^\top\boldsymbol{C})^{-1}\boldsymbol{C}^\top\boldsymbol{\mu_R}}_{\widehat{\boldsymbol{\lambda}}}, \underbrace{\sigma^2(\boldsymbol{C}^\top\boldsymbol{C})^{-1}}_{\boldsymbol{\Sigma_\lambda}}\right)$$

$$\sigma^2|data \sim \mathcal{IG}\left(\frac{N-K-1}{2}, \frac{(\boldsymbol{\mu_R} - \boldsymbol{C}\hat{\boldsymbol{\lambda}})^\top(\boldsymbol{\mu_R} - \boldsymbol{C}\hat{\boldsymbol{\lambda}})}{2}\right)$$

2.2. Bayesian Analysis of Linear SDFs

- Average pricing errors are cross−sectionally correlated, $\boldsymbol{\alpha}$, $\boldsymbol{\alpha} \sim \mathcal{N}(\mathbf{0}_{\boldsymbol{N}}, \sigma^2 \boldsymbol{\Sigma}_{\boldsymbol{R}})$
  - $\boldsymbol{\mu}_{\boldsymbol{R}} = \boldsymbol{C}\boldsymbol{\lambda} + \boldsymbol{\alpha}$
  - Posterior distribution of $(\boldsymbol{\lambda}, \sigma^2)$

$$\boldsymbol{\lambda} | \sigma^2, data \sim \mathcal{N}\left(\underbrace{(\boldsymbol{C}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-\mathbf{1}} \boldsymbol{C})^{-1} \boldsymbol{C}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-\mathbf{1}} \boldsymbol{\mu}_{\boldsymbol{R}}}_{\widehat{\boldsymbol{\lambda}}}, \ \underbrace{\sigma^2 (\boldsymbol{C}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-\mathbf{1}} \boldsymbol{C})^{-1}}_{\boldsymbol{\Sigma}_{\boldsymbol{\lambda}}} \right)$$

$$\sigma^2 | data \sim \mathcal{IG}\left(\frac{N - K - 1}{2}, \frac{(\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}\hat{\boldsymbol{\lambda}})^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-\mathbf{1}} (\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}\hat{\boldsymbol{\lambda}})}{2} \right)$$

2.3. Model Selection and Aggregation

- **Pitfalls of Flat Priors for Risk Prices** $\pi(\boldsymbol{\lambda}_{\boldsymbol{\gamma}}, \sigma^2) \propto \frac{1}{\sigma^2}$ and $\boldsymbol{\lambda}_{-\boldsymbol{\gamma}} = 0$
- $\boldsymbol{\gamma}^\top = (\gamma_0, \gamma_1, \ldots, \gamma_K)$,where $\gamma_j \in \{0, 1\}$
- $\gamma_j = 1$, factor j should be included
- $\boldsymbol{C}_{\boldsymbol{\gamma}} = [\boldsymbol{C}_j]_{\gamma_j=1}$ represents a p_γ-columns sub-matrix of C.
- The posterior probability of model γ

$$\Pr(\boldsymbol{\gamma}|data) = \frac{p(data|\boldsymbol{\gamma})}{p(data|\boldsymbol{\gamma}) + p(data|\boldsymbol{\gamma}')} \longrightarrow \text{equal prior probability}$$

- $p(data|\gamma)$ denotes the **marginal likelihood** of the model indexed by γ, using a flat prior for λ

$$p(data|\boldsymbol{\gamma}) \propto (2\pi)^{\frac{p_\gamma}{2}} |\boldsymbol{C}_{\boldsymbol{\gamma}}^\top \boldsymbol{C}_{\boldsymbol{\gamma}}|^{-\frac{1}{2}} \frac{\Gamma(\frac{N-p_\gamma}{2})}{(\frac{N\hat{\sigma}_\gamma^2}{2})^{\frac{N-p_\gamma}{2}}}$$

- Under a flat prior for risk prices, the model containing a weak factor will always be selected.

2.3. Model Selection and Aggregation

$$\pi(\boldsymbol{\lambda}_{\gamma}, \sigma^2) \propto \frac{1}{\sigma^2} \text{ and } \boldsymbol{\lambda}_{-\gamma} = 0$$

- **Spike-and-Slab Prior for Risk Prices**
  - $\pi(\boldsymbol{\lambda}, \sigma^2, \boldsymbol{\gamma}) \propto \pi(\boldsymbol{\lambda}|\sigma^2, \boldsymbol{\gamma})\pi(\sigma^2)\pi(\boldsymbol{\gamma})$
  - γ_j= 1, the factor included in the model, the prior (the "slab") follows a normal distribution $\lambda_j|\sigma^2, \gamma_j = 1 \sim \mathcal{N}(0, \sigma^2\psi_j)$
  - γ_j = 0, the factor not included in the model, the prior (the "spike") is a Dirac distribution at zero. $\delta(x) = \begin{cases} +\infty, x = 0 \\ 0, x = otherwise \end{cases}$
  - Cross-sectional variance prior $\pi(\sigma^2) \propto \sigma^{-2}$
  - Let $\boldsymbol{D}$ denote a diagonal matrix with elements $c, \psi_1^{-1}, \cdots \psi_K^{-1}$
  - $\boldsymbol{D}_{\gamma}$ the sub-matrix of $\boldsymbol{D}$ corresponding to model $\boldsymbol{\gamma}$

$$\boldsymbol{\lambda}_{\gamma}|\sigma^2, \boldsymbol{\gamma} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{D}_{\gamma}^{-1})$$

2.3. Model Selection and Aggregation

- **Spike-and-Slab Prior for Risk Prices**

Proposition 2 (B-SDF OLS Posterior with Dirac Spike-and-Slab) *The posterior distribution of* $(\lambda_\gamma, \sigma^2, \boldsymbol{\gamma})$ *under the assumption of Dirac spike-and-slab prior and spherical* $\boldsymbol{\alpha}$ *(OLS), conditional on the draws of* $\boldsymbol{\mu_Y}$ *and* $\boldsymbol{\Sigma_Y}$ *from equations (6)–(7), is given by the following conditional distributions:*

$$\boxed{\boldsymbol{\lambda}_\gamma | data, \sigma^2, \boldsymbol{\gamma} \sim \mathcal{N}\left(\hat{\boldsymbol{\lambda}}_\gamma, \hat{\sigma}^2(\hat{\boldsymbol{\lambda}}_\gamma)\right)} \tag{16}$$

$$\sigma^2 | data, \boldsymbol{\gamma} \sim \mathcal{IG}\left(\frac{N}{2}, \frac{SSR_\gamma}{2}\right), \text{ and} \tag{17}$$

$$p(\boldsymbol{\gamma} \mid data) \propto \frac{|\boldsymbol{D}_\gamma|^{\frac{1}{2}}}{|\boldsymbol{C}_\gamma^\top \boldsymbol{C}_\gamma + \boldsymbol{D}_\gamma|^{\frac{1}{2}}} \frac{1}{(SSR_\gamma/2)^{\frac{N}{2}}}, \tag{18}$$

where $\hat{\boldsymbol{\lambda}}_\gamma = (\boldsymbol{C}_\gamma^\top \boldsymbol{C}_\gamma + \boxed{\boldsymbol{D}_\gamma})^{-1} \boldsymbol{C}_\gamma^\top \boldsymbol{\mu_R}$, $\hat{\sigma}^2(\hat{\boldsymbol{\lambda}}_\gamma) = \sigma^2(\boldsymbol{C}_\gamma^\top \boldsymbol{C}_\gamma + \boxed{\boldsymbol{D}_\gamma})^{-1}$, *and* $SSR_\gamma = \boldsymbol{\mu_R}^\top \boldsymbol{\mu_R} - \boldsymbol{\mu_R}^\top \boldsymbol{C}_\gamma (\boldsymbol{C}_\gamma^\top \boldsymbol{C}_\gamma + \boldsymbol{D}_\gamma)^{-1} \boldsymbol{C}_\gamma^\top \boldsymbol{\mu_R} = \min_{\boldsymbol{\lambda}_\gamma} \{(\boldsymbol{\mu_R} - \boldsymbol{C}_\gamma \boldsymbol{\lambda}_\gamma)^\top (\boldsymbol{\mu_R} - \boldsymbol{C}_\gamma \boldsymbol{\lambda}_\gamma) + \boldsymbol{\lambda}_\gamma^\top \boldsymbol{D}_\gamma \boldsymbol{\lambda}_\gamma\}$ *and* $\mathcal{IG}$ *denotes the inverse-Gamma distribution.*

2.3. Model Selection and Aggregation

- **Spike-and-Slab Prior for Risk Prices**

Proposition 3 (B-SDF GLS Posterior with Dirac Spike-and-Slab) *The posterior distribution of* $(\lambda_\gamma, \sigma^2, \boldsymbol{\gamma})$ *under the assumption of Dirac spike-and-slab prior and and non-spherical* $\boldsymbol{\alpha}$ *(GLS), conditional on the draws of* $\boldsymbol{\mu_Y}$ *and* $\boldsymbol{\Sigma_Y}$ *from equations (6)–(7), is given by the following conditional distributions:*

$$\boxed{\boldsymbol{\lambda}_\gamma | data, \sigma^2, \boldsymbol{\gamma} \sim \mathcal{N}\left(\hat{\boldsymbol{\lambda}}_\gamma, \hat{\sigma}^2(\hat{\boldsymbol{\lambda}}_\gamma)\right),} \tag{19}$$

$$\sigma^2 | data, \boldsymbol{\gamma} \sim \mathcal{IG}\left(\frac{N}{2}, \frac{SSR_\gamma}{2}\right), \text{ and} \tag{20}$$

$$p(\boldsymbol{\gamma} \mid data) \propto \frac{|\boldsymbol{D}_\gamma|^{\frac{1}{2}}}{|\boldsymbol{C}_\gamma^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{C}_\gamma + \boldsymbol{D}_\gamma|^{\frac{1}{2}}} \frac{1}{(SSR_\gamma/2)^{\frac{N}{2}}}, \tag{21}$$

where $\hat{\boldsymbol{\lambda}}_\gamma = (\boldsymbol{C}_\gamma^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{C}_\gamma + \boxed{\boldsymbol{D}_\gamma})^{-1} \boldsymbol{C}_\gamma^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{\mu}_{\boldsymbol{R}}$, $\hat{\sigma}^2(\hat{\boldsymbol{\lambda}}_\gamma) = \sigma^2 (\boldsymbol{C}_\gamma^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{C}_\gamma + \boxed{\boldsymbol{D}_\gamma})^{-1}$, *and* $SSR_\gamma = \boldsymbol{\mu}_{\boldsymbol{R}}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{\mu}_{\boldsymbol{R}}^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{C}_\gamma (\boldsymbol{C}_\gamma^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{C}_\gamma + \boldsymbol{D}_\gamma)^{-1} \boldsymbol{C}_\gamma^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} \boldsymbol{\mu}_{\boldsymbol{R}} = \min_{\boldsymbol{\lambda}_\gamma} \{(\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}_\gamma \boldsymbol{\lambda}_\gamma)^\top \boldsymbol{\Sigma}_{\boldsymbol{R}}^{-1} (\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}_\gamma \boldsymbol{\lambda}_\gamma) + \boldsymbol{\lambda}_\gamma^\top \boldsymbol{D}_\gamma \boldsymbol{\lambda}_\gamma\}$ *and* $\mathcal{IG}$ *denotes the inverse-Gamma distribution.*

2.3. Model Selection and Aggregation

- **Spike-and-Slab Prior for Risk Prices**

Let $\boldsymbol{D}$ denote a diagonal matrix with elements $c, \psi_1^{-1}, \cdots \psi_K^{-1}$

$$\psi_j = \psi \times \boldsymbol{\rho}_j^\top \boldsymbol{\rho}_j$$

$\boldsymbol{\rho}_j$ is an $N \times 1$ vector of **correlation coefficients** between factor j and the test assets

$\psi \in \mathbb{R}_+$ is a **tuning parameter** that controls the degree of shrinkage over all factors

$$\hat{\boldsymbol{\lambda}}_\gamma = (\boldsymbol{C}_\gamma^\top \boldsymbol{C}_\gamma + \boldsymbol{D}_\gamma)^{-1} \boldsymbol{C}_\gamma^\top \boldsymbol{\mu}_{\boldsymbol{R}}$$

This Bayesian formulation is robust to weak factors

2.3. Model Selection and Aggregation

- **Continuous Spike**
  - Following Bayesian variable selection (see, e.g., George and McCulloch (1993, 1997) and Ishwaran, Rao, et al. (2005))

$$\pi(\boldsymbol{\lambda}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega}) = \pi(\boldsymbol{\lambda} \mid \sigma^2, \boldsymbol{\gamma})\pi(\sigma^2)\pi(\boldsymbol{\gamma} \mid \boldsymbol{\omega})\pi(\boldsymbol{\omega})$$

$$\lambda_j \mid \gamma_j, \sigma^2 \sim \mathcal{N}\left(0, r(\gamma_j)\psi_j\sigma^2\right)$$

  - $r(\gamma_j = 1) = 1$, hence we have the same "slab" as before
  - $r(\gamma_j = 0) = r \ll 1$ Dirac "spike" is replaced by a Gaussian spike

  - $\pi(\omega)$ encodes our ex ante beliefs about **the sparsity** of the true model

  - $\boldsymbol{D}$ as a diagonal matrix with elements $c, \ (r(\gamma_1)\psi_1)^{-1}, \ldots, (r(\gamma_K)\psi_K)^{-1}$

$$\boldsymbol{\lambda}|\sigma^2, \boldsymbol{\gamma} \sim \mathcal{N}(0, \sigma^2\boldsymbol{D}^{-1})$$

$$\pi(\gamma_j = 1|\omega_j) = \omega_j, \quad \omega_j \sim Beta\left(a_\omega, b_\omega\right) \qquad \frac{a_\omega}{a_\omega + b_\omega}$$

$$\mathbb{E}_\pi[SR_{\boldsymbol{f}}^2 \mid \sigma^2] = \boxed{\frac{a_\omega}{a_\omega + b_\omega}}\psi\sigma^2 \sum_{k=1}^K \tilde{\boldsymbol{\rho}}_k^\top \tilde{\boldsymbol{\rho}}_k$$

2.3. Model Selection and Aggregation

- **Continuous Spike**

Proposition 5 (B-SDF OLS Posterior with Continuous Spike-and-Slab) *The posterior distribution of* $(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{\omega}, \sigma^2)$ *under the assumption of continuous spike-and-slab prior and spherical* $\boldsymbol{\alpha}$ *(OLS), conditional on the draws of* $\boldsymbol{\mu}_{\boldsymbol{Y}}$ *and* $\boldsymbol{\Sigma}_{\boldsymbol{Y}}$ *from equations (6)–(7), is given by the following conditional distributions:*

$$\boxed{\boldsymbol{\lambda} | data, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega} \sim \mathcal{N}\left(\hat{\boldsymbol{\lambda}}, \hat{\sigma}^2(\hat{\boldsymbol{\lambda}})\right),} \tag{28}$$

$$\frac{p(\gamma_j = 1 | data, \boldsymbol{\lambda}, \boldsymbol{\omega}, \sigma^2, \boldsymbol{\gamma}_{-j})}{p(\gamma_j = 0 | data, \boldsymbol{\lambda}, \boldsymbol{\omega}, \sigma^2, \boldsymbol{\gamma}_{-j})} = \frac{\omega_j}{1-\omega_j} \frac{p(\lambda_j | \gamma_j = 1, \sigma^2)}{p(\lambda_j | \gamma_j = 0, \sigma^2)}, \tag{29}$$

$$\omega_j | data, \boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma^2 \sim Beta\left(\gamma_j + a_\omega, 1 - \gamma_j + b_\omega\right), \text{ and} \tag{30}$$

$$\sigma^2 | data, \boldsymbol{\omega}, \boldsymbol{\lambda}, \boldsymbol{\gamma} \sim \mathcal{IG}\left(\frac{N+K+1}{2}, \frac{(\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}\boldsymbol{\lambda})^\top (\boldsymbol{\mu}_{\boldsymbol{R}} - \boldsymbol{C}\boldsymbol{\lambda}) + \boldsymbol{\lambda}^\top \boldsymbol{D} \boldsymbol{\lambda}}{2}\right), \tag{31}$$

where $\hat{\boldsymbol{\lambda}} = (\boldsymbol{C}^\top \boldsymbol{C} + \boldsymbol{D})^{-1} \boldsymbol{C}^\top \boldsymbol{\mu}_{\boldsymbol{R}}$ *and* $\hat{\sigma}^2(\hat{\boldsymbol{\lambda}}) = \sigma^2 (\boldsymbol{C}^\top \boldsymbol{C} + \boldsymbol{D})^{-1}$.

2.3. Model Selection and Aggregation

- **Continuous Spike**

Proposition 6 (B-SDF GLS Posterior with Continuous Spike-and-Slab) *The posterior distribution of* $(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{\omega}, \sigma^2)$ *under the assumption of continuous spike-and-slab prior and non-spherical* $\boldsymbol{\alpha}$ *(GLS), conditional on the draws of* $\boldsymbol{\mu_Y}$ *and* $\boldsymbol{\Sigma_Y}$ *from equations (6)–(7), differs from ones in Proposition 5 only for the posterior distributions of* $(\boldsymbol{\lambda}, \sigma^2)$*:*

$$\boxed{\boldsymbol{\lambda}|data, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega} \sim \mathcal{N}\left(\hat{\boldsymbol{\lambda}}, \hat{\sigma}^2(\hat{\boldsymbol{\lambda}})\right)}, \text{ and} \tag{32}$$

$$\sigma^2|data, \boldsymbol{\omega}, \boldsymbol{\lambda}, \boldsymbol{\gamma} \sim \mathcal{IG}\left(\frac{N+K+1}{2}, \frac{(\boldsymbol{\mu_R} - \boldsymbol{C}\boldsymbol{\lambda})^\top \boldsymbol{\Sigma_R}^{-1}(\boldsymbol{\mu_R} - \boldsymbol{C}\boldsymbol{\lambda}) + \boldsymbol{\lambda}^\top \boldsymbol{D}\boldsymbol{\lambda}}{2}\right), \tag{33}$$

where $\hat{\boldsymbol{\lambda}} = (\boldsymbol{C}^\top \boldsymbol{\Sigma_R}^{-1} \boldsymbol{C} + \boldsymbol{D})^{-1} \boldsymbol{C}^\top \boldsymbol{\Sigma_R}^{-1} \boldsymbol{\mu_R}$ *and* $\hat{\sigma}^2(\hat{\boldsymbol{\lambda}}) = \sigma^2 (\boldsymbol{C}^\top \boldsymbol{\Sigma_R}^{-1} \boldsymbol{C} + \boldsymbol{D})^{-1}$.

2.3. Model Selection and Aggregation

- **Selection vs. Aggregation**

  - If we are interested in some quantity Δ that is well-defined for every model $m = 1, \dots, \bar{m}$ (e.g., price of risk, risk premia, and maximum Sharpe ratio)

$$\mathbb{E}\left[\Delta|\mathrm{data}\right] = \sum_{m=0}^{\bar{m}} \mathbb{E}\left[\Delta|\mathrm{data}, \mathrm{model} = m\right] \Pr\left(\mathrm{model} = m|\mathrm{data}\right)$$

  - The BMA expectation of Δ, conditional on only the data is simply the weighted average of the expectation in every model, with weights equal to the models' posterior probabilities

3.2. Empirical Analysis

- **Sampling Two Quadrillion Models**

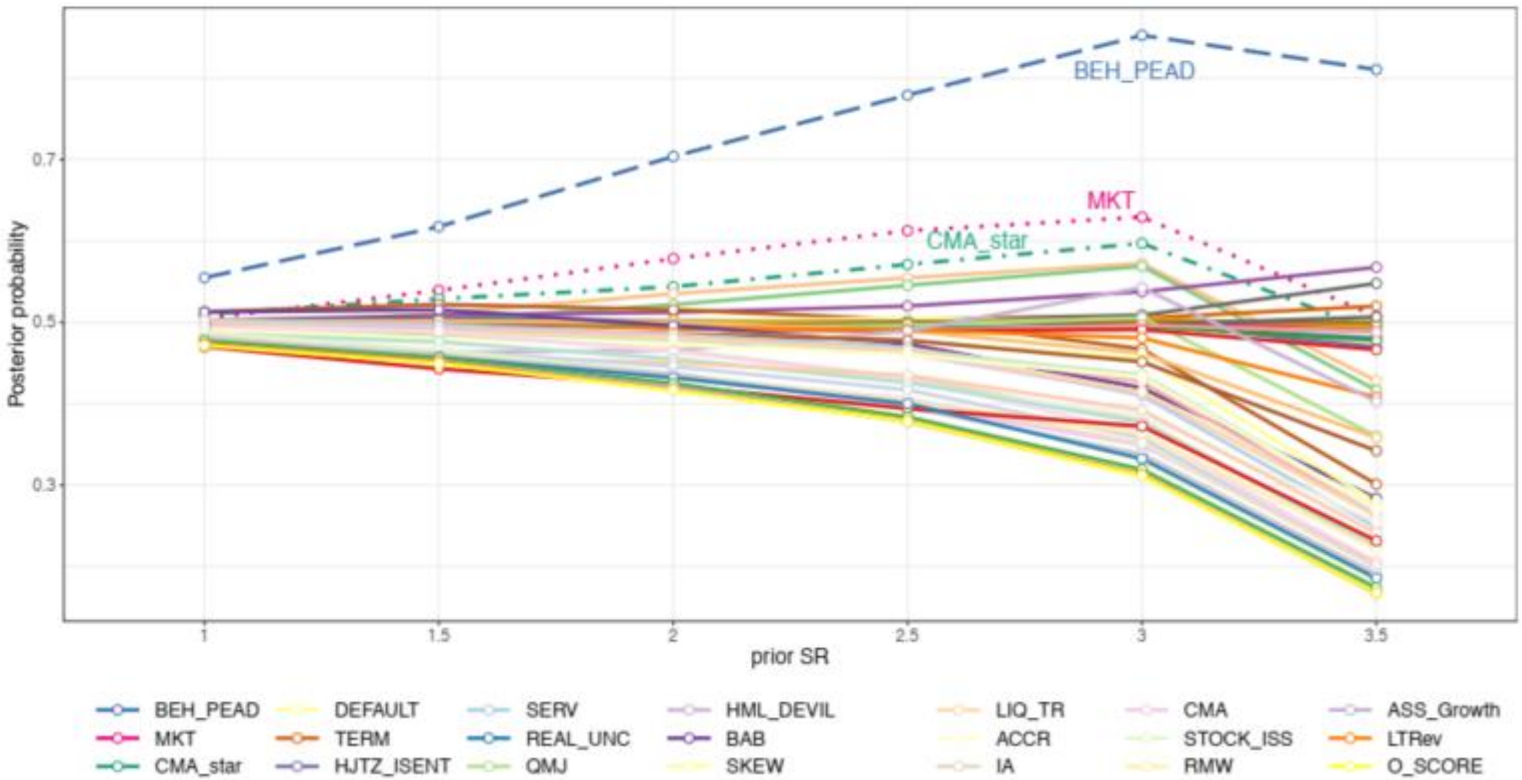
- We focus on 51 (both tradable and non-tradable) monthly factors available from October 1973 to December 2016

- Models: $2^{51} = 2.25\ quadrillion$

- cross-section of 60 asset returns(long-short portfolio)

3.2. Empirical Analysis

- **Sampling Two Quadrillion Models**



- BEH PEAD、MKT 、CMA

3.2. Empirical Analysis

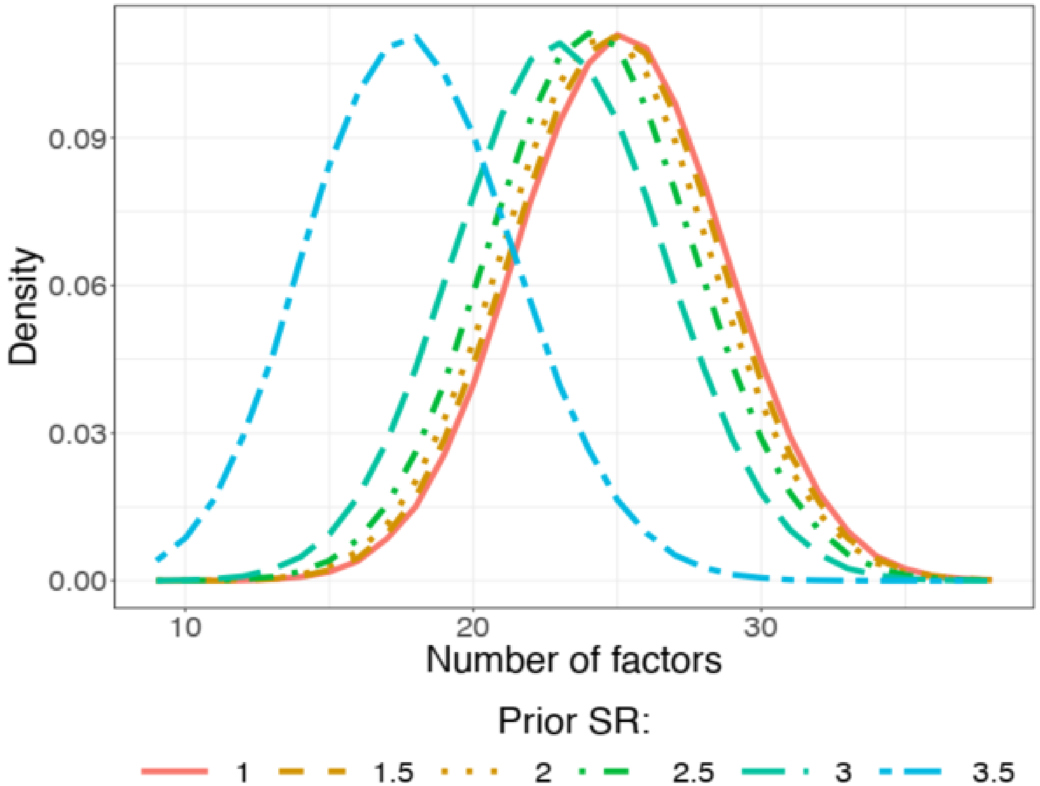
- **Cross-Sectional Performance**

Table 4: Cross-sectional asset pricing

Model		RMSE	MAPE	R^2_{ols}	R^2_{gls}	Model	RMSE	MAPE	R^2_{ols}	R^2_{gls}
Panel A: In-sample pricing, test assets: 60 anomalies										
BMA-SDF:	$SR_{pr}=1$	0.287	0.227	39.2%	24.2%	51 factors	0.041	0.022	98.1%	97.7%
	$SR_{pr}=1.5$	0.253	0.197	49.8%	30.3%	CAPM	0.418	0.338	-29.4%	16.8%
	$SR_{pr}=2$	0.223	0.170	59.1%	37.4%	FF5	0.301	0.223	24.5%	23.2%
	$SR_{pr}=2.5$	0.193	0.148	68.2%	45.5%	Carhart	0.317	0.244	21.5%	21.2%
	$SR_{pr}=3$	0.162	0.128	76.6%	54.7%	q4	0.267	0.189	37.5%	28.1%
	$SR_{pr}=3.5$	0.157	0.128	78.4%	58.8%	KNS_{CV_3}	0.296	0.237	53.7%	19.6%
Panel B: Out-of-sample pricing, test assets: 25 size-value portfolios										
BMA-SDF:	$SR_{pr}=1$	0.108	0.082	42.1%	17.5%	51 factors	0.200	0.163	-98.5%	-1653%
	$SR_{pr}=1.5$	0.094	0.070	55.7%	24.5%	CAPM	0.145	0.112	-4.6%	5.2%
	$SR_{pr}=2$	0.085	0.063	64.5%	30.2%	FF5	0.079	0.059	69.2%	28.0%
	$SR_{pr}=2.5$	0.077	0.058	70.5%	34.9%	Carhart	0.086	0.063	63.2%	27.1%
	$SR_{pr}=3$	0.073	0.054	73.9%	38.4%	q4	0.083	0.065	66.1%	28.2%
	$SR_{pr}=3.5$	0.075	0.055	72.3%	36.8%	KNS_{CV_3}	0.096	0.074	54.4%	28.0%

- Bayesian SDF tends to outperform conventional models

3.2. Empirical Analysis



(a) Model dimensionality

- The posterior mean of the number of factors in the true model is in the 23–25 range.

3.2. Empirical Analysis

model: SR_{pr}:	Panel A: 3 most likely factors					Panel B: 6 most likely factors				
	1	1.5	2	2.5	3	1	1.5	2	2.5	3
Most likely factors	17.5%	24.9%	36.0%	48.8%	59.1%	17.8%	27.0%	44.0%	66.5%	83.7%
CAPM	12.7%	12.5%	11.8%	11.3%	13.1%	12.7%	12.1%	10.3%	7.3%	5.2%
FF3	10.3%	7.9%	5.3%	3.2%	1.7%	10.3%	7.7%	4.7%	2.1%	0.7%
FF5	9.9%	7.0%	4.2%	2.1%	0.7%	9.8%	6.8%	3.7%	1.3%	0.3%
Carhart	10.2%	7.8%	5.2%	2.9%	1.3%	10.2%	7.6%	4.6%	1.9%	0.5%
q4	15.7%	17.8%	17.9%	14.9%	9.6%	15.6%	17.3%	15.7%	9.9%	3.9%
Liq-CAPM	12.5%	12.0%	10.9%	9.6%	9.0%	12.5%	11.7%	9.5%	6.2%	3.6%
FF3-QMJ	11.2%	10.1%	8.8%	7.4%	5.5%	11.1%	9.8%	7.7%	4.8%	2.1%

- The three to six robust factors provide a model that substantially outperforms notable benchmarks

4.Conlusion

- Only a handful of factors seem to be robust explanators of the cross-sections of asset returns

- The three to six robust factors provide a model that substantially outperforms notable benchmarks

- A BMA over the universe of possible models delivers an SDF that presents a novel benchmark for in- and out-of-sample empirical asset pricing

Extensions

- Time-varying expected returns and SDF factor loadings could be accommodated by adopting the **time-varying parameter** approach.

- extended to the estimation of **beta** representations of the fundamental pricing equation used in the two-pass procedure, such as **Fama-MacBeth** regressions

3.1. Simulation

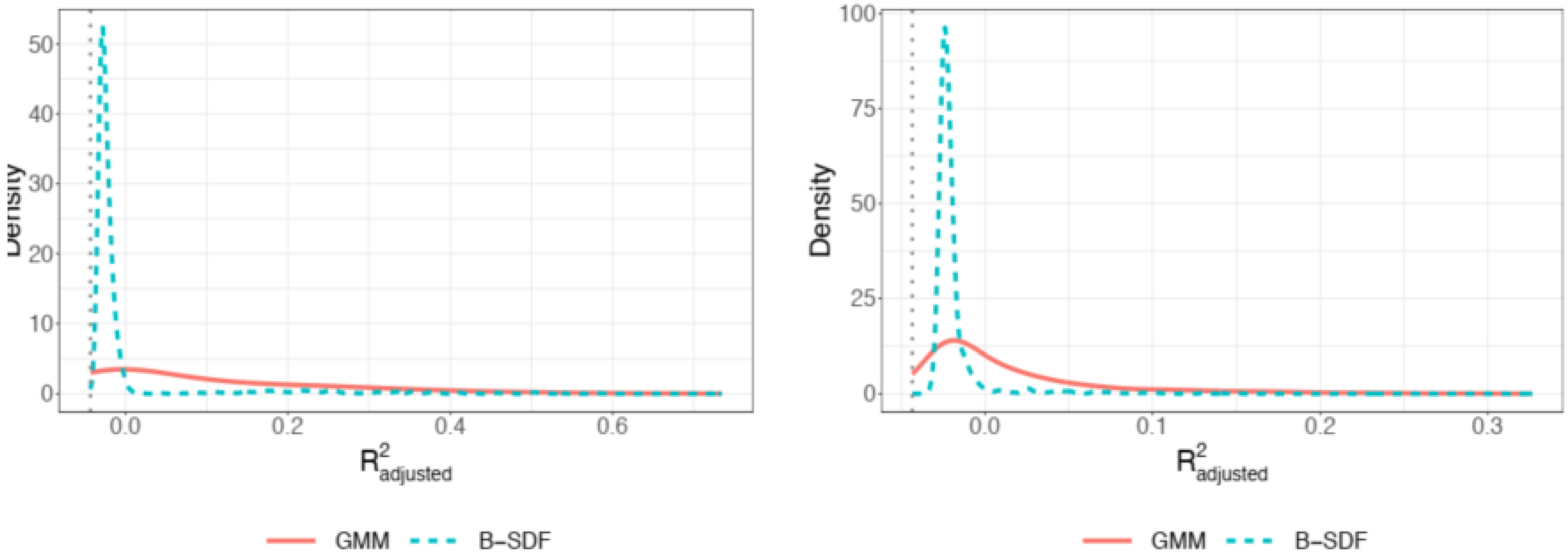
- A linear factor model that includes both **strong and weak** factors
- The cross-section of asset returns mimics the empirical properties of **25 Fama-French portfolios** sorted by size and value
- **HML** is the only useful factor

$$f_{t,useless} \overset{\text{iid}}{\sim} \mathcal{N}(0,(1\%)^2), \qquad \begin{pmatrix} \boldsymbol{R_t} \\ f_{t,hml} \end{pmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(\begin{bmatrix} \bar{\boldsymbol{\mu}}_{\boldsymbol{R}} \\ \bar{f}_{hml} \end{bmatrix}, \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}} & \widehat{\boldsymbol{C}}_{\boldsymbol{hml}} \\ \widehat{\boldsymbol{C}}_{\boldsymbol{hml}}^{\top} & \widehat{\sigma}_{hml}^2 \end{bmatrix} \right), \text{ and}$$

$$\boldsymbol{\mu_R} = \begin{cases} \widehat{\lambda}_c \mathbf{1}_{\boldsymbol{N}} + \widehat{\boldsymbol{C}}_{\boldsymbol{f}} \widehat{\lambda}_{HML}, & \text{if the model is correct, and} \\ \bar{\boldsymbol{R}}, & \text{if the model is misspecified,} \end{cases}$$

- Sample sizes: T = 100, 200, 600, 1,000, and 20,000

3.1. Simulation



(a) OLS-type estimation

(b) GLS-type estimation

Figure 2: Cross-sectional distribution of OLS R^2_{adj} in a model with a useless factor.

- The mode of the posterior distribution in B-SDF is tightly concentrated (across simulations) in the proximity of the true R2 value

3.1. Simulation

Table 2: The probability of retaining risk factors using Bayes factors

T		55%	57%	59%	61%	63%	65%		55%	57%	59%	61%	63%	65%
Panel A: Flat prior														
200	f_{strong}:	0.636	0.602	0.570	0.538	0.509	0.470	$f_{useless}$:	0.980	0.950	0.856	0.724	0.581	0.437
600		0.821	0.802	0.784	0.764	0.733	0.710		0.996	0.983	0.970	0.932	0.878	0.791
1,000		0.880	0.850	0.840	0.840	0.800	0.800		1.000	1.000	0.990	0.980	0.940	0.910
Panel B: Spike-and-Slab, prior of $\sqrt{\mathbb{E}_\pi[SR_f^2 \mid \sigma^2]} = 0.295$														
200	f_{strong}:	0.815	0.761	0.721	0.675	0.630	0.581	$f_{useless}$:	0.004	0.000	0.000	0.000	0.000	0.000
600		0.974	0.961	0.954	0.943	0.926	0.899		0.000	0.000	0.000	0.000	0.000	0.000
1,000		0.980	0.970	0.970	0.960	0.960	0.940		0.000	0.000	0.000	0.000	0.000	0.000
Panel C: Spike-and-Slab, prior of $\sqrt{\mathbb{E}_\pi[SR_f^2 \mid \sigma^2]} = 0.807$														
200	f_{strong}:	0.527	0.489	0.449	0.412	0.381	0.349	$f_{useless}$:	0.041	0.007	0.004	0.000	0.000	0.000
600		0.859	0.832	0.811	0.774	0.734	0.690		0.001	0.000	0.000	0.000	0.000	0.000
1,000		0.910	0.910	0.870	0.850	0.830	0.820		0 .000	0.000	0.000	0.000	0.000	0.000

- The spike-and-slab prior very encouraging for variable and model selection: It successfully eliminates the impact of the useless factors

3.2. Empirical Analysis

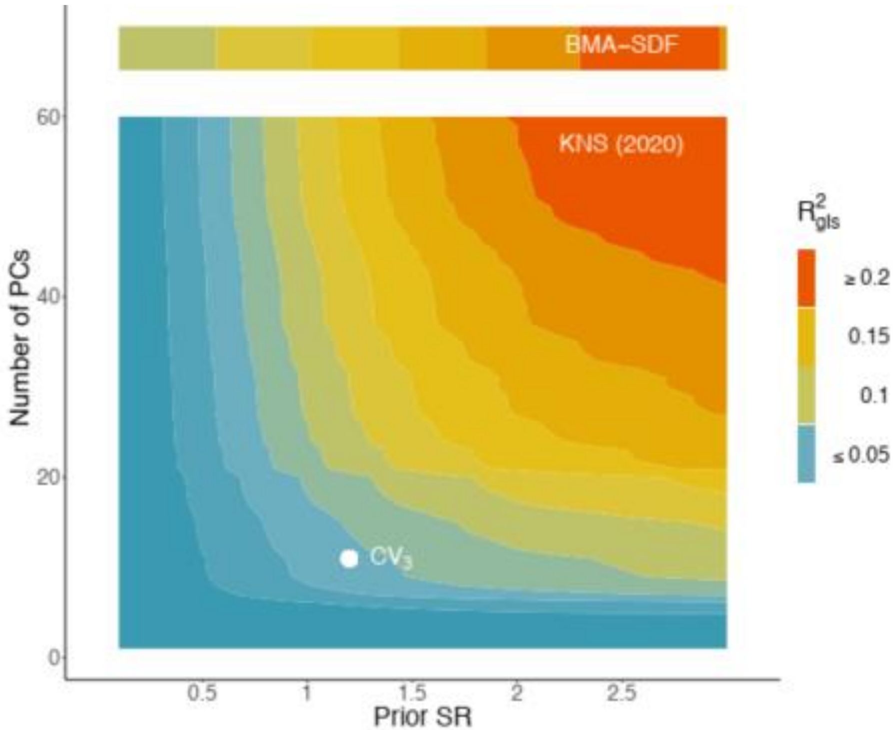
Panel B: RP-Principal Components (Lettau and Pelger (2020)) as Factors

RP-PC1	0.600	0.631	0.640	0.634	0.592	0.448	-0.016	-0.030	-0.043	-0.056	-0.066	-0.067
RP-PC3	0.548	0.597	0.645	0.661	0.651	0.529	-0.004	-0.009	-0.017	-0.024	-0.032	-0.035
BEH_PEAD	0.540	0.585	0.628	0.681	0.709	0.630	0.014	0.032	0.058	0.097	0.149	0.185
CMA*	0.510	0.523	0.542	0.571	0.616	0.531	0.009	0.020	0.037	0.062	0.104	0.129
RMW*	0.500	0.504	0.517	0.547	0.583	0.466	0.007	0.017	0.033	0.059	0.101	0.112
MKT	0.507	0.518	0.525	0.516	0.493	0.391	0.013	0.028	0.044	0.061	0.081	0.103
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Useless I	0.499	0.499	0.500	0.500	0.499	0.497	0.000	0.000	0.000	0.000	0.001	0.007
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Useless II	0.495	0.495	0.495	0.498	0.496	0.499	0.000	0.000	0.000	0.001	0.002	0.010
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
RP-PC5	0.481	0.487	0.488	0.484	0.459	0.338	0.001	0.003	0.005	0.008	0.011	0.012
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
RP-PC4	0.494	0.487	0.479	0.459	0.433	0.303	0.002	0.003	0.004	0.005	0.005	0.005
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
RP-PC2	0.479	0.464	0.458	0.439	0.403	0.267	0.000	-0.001	-0.001	-0.001	-0.001	0.000

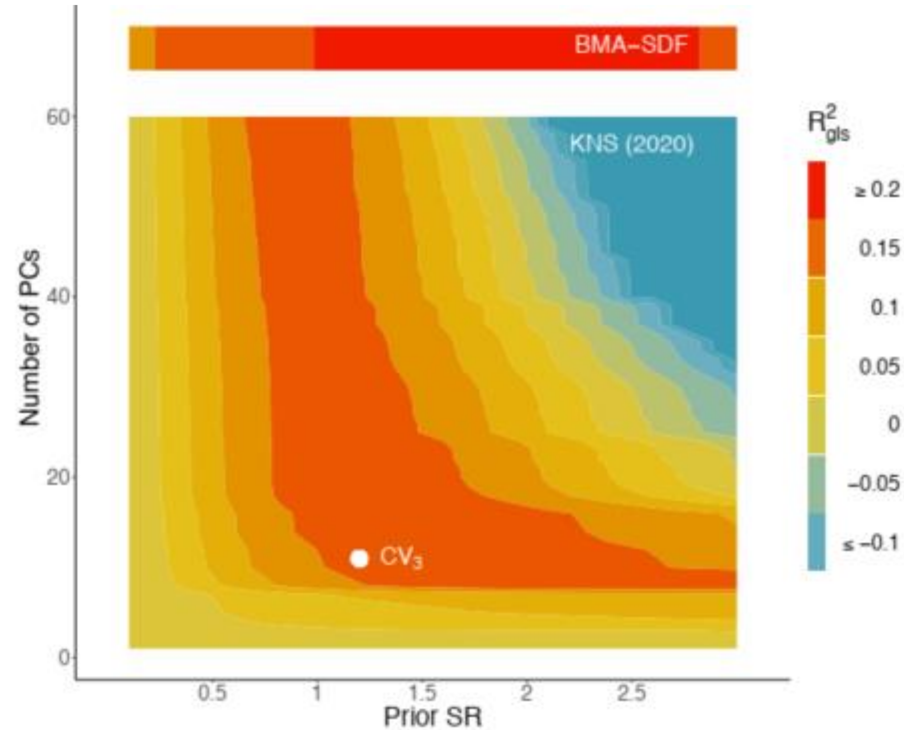
- The underlying SDF would be best described by a combination of both observable factors and (some) latent variables

3.2. Empirical Analysis

- **Selection or Aggregation**



(c) R^2_{gls}, estimation period: 1995/06-2016/12, evaluation period: 1973/10-1995/05



(d) R^2_{gls}, estimation period: 1973/10-1995/05, evaluation period: 1995/06-2016/12

- For the same value of the prior SR, BMA-SDF tends to outperform the cross-validated estimates (CV3) of KNS