Estimating Asset Pricing Factors from Large-Dimensional Panel Data

Markus Pelger ¹  Martin Lettau ²

¹Stanford University
²UC Berkeley

January 24th, 2017

UC Berkeley
Risk Seminar
Motivation: Cochrane (2011, Presidential Address)

**The Challenge of Cross-Sectional Asset Pricing**

- Fundamental insight: Arbitrage Pricing Theory: Expected return of assets should be explained by systematic risk factors.

- Problem: “Chaos” in asset pricing factors: Over 330 potential asset pricing factors published!

- Fundamental question: Which factors are really important in explaining expected returns? Which are subsumed by others?

**Goals of this paper:**

- Estimate “priced” factors:
  ⇒ Search for priced factors and separate them from unpriced factors

- Bring order into “factor chaos”
  ⇒ Summarize the pricing information in a small number of factors

- Illustrate and explain the flaws of statistical factors based on PCA
Contribution of this paper

Contribution

- New estimator for estimating *priced* latent factors (that can explain expected returns) from *large panel data sets*

Estimation theory:
- Asymptotic distribution theory for weak and strong factors
- Weak assumptions: Approximate factor model and arbitrage-pricing theory
- Estimator discovers “weak” factors with high Sharpe-ratios
- Strongly dominates PCA

Empirical results:
- New factors explain correlation structure and cross-sectional expected returns at the same time
- New factors have in *and out-of sample* smaller pricing errors and larger Sharpe-ratios than benchmark factors
Literature (partial list)

- Large-dimensional factor models with strong factors
  - Bai (2003): Distribution theory
  - Ahn and Horenstein (2013), Onatski (2010), Bai and Ng (2002): Determining the number of factors
  - Fan et al. (2013): Sparse matrices in factor modeling
  - Pelger (2016), Aït-Sahalia and Xiu (2015): High-frequency

- Large-dimensional factor models with weak factors (based on random matrix theory)
  - Onatski (2012): Phase transition phenomena
  - Benauch-Georges and Nadakuditi (2011): Perturbation of large random matrices

- Asset-pricing factors
  - Harvey and Liu (2015): Lucky factors
  - Clarke (2015): Level, slope and curvature for stocks
  - Kozak, Nagel and Santosh (2015): PCA based factors
  - Bryzgalova (2016): Spurious factors
Agenda

1. Introduction (√)
2. Factor model setup and illustration
3. Statistical model
   1. Weak factor model
   2. Strong factor model
4. Simulation
5. Empirical results
6. Conclusion
The Model

Approximate Factor Model

- Observe excess returns of $N$ assets over $T$ time periods:

$$X_{t,i} = F_t^T \Lambda_i + e_{t,i} \quad i = 1, \ldots, N \quad t = 1, \ldots, T$$

- Matrix notation

$$X = F \Lambda^T + e$$

- $N$ assets (large)
- $T$ time-series observation (large)
- $K$ systematic factors (fixed)

- $F$, $\Lambda$ and $e$ are unknown
The Model

Approximate Factor Model

- Systematic and non-systematic risk ($F$ and $e$ uncorrelated):

\[
\text{Var}(X) = \Lambda \text{Var}(F)\Lambda^\top + \text{Var}(e) = \underbrace{\text{systematic}} + \underbrace{\text{non-systematic}}
\]

⇒ Systematic factors should explain a large portion of the variance
⇒ Idiosyncratic risk can be weakly correlated

- Arbitrage-Pricing Theory (APT): The expected excess return is explained by the risk-premium of the factors:

\[
E[X_i] = E[F]\Lambda_i^\top
\]

⇒ Systematic factors should explain the cross-section of expected returns
The Model

Time-series objective function:

Minimize the unexplained variance:

\[
\min_{\Lambda, F} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{ti} - F_{t}\Lambda^{T}_{i})^2
\]

\[
= \min_{\Lambda} \frac{1}{NT} \text{trace} \left( (XM_{\Lambda})^{T} (XM_{\Lambda}) \right) \quad \text{s.t.} \quad F = X(\Lambda^{T}\Lambda)^{-1}\Lambda^{T}
\]

- Projection matrix \( M_{\Lambda} = I_{N} - \Lambda(\Lambda^{T}\Lambda)^{-1}\Lambda^{T} \)
- Error (non-systematic risk): \( e = X - F\Lambda^{T} = XM_{\Lambda} \)
- \( \Lambda \) proportional to eigenvectors of the first \( K \) largest eigenvalues of \( \frac{1}{NT} X^{T}X \) minimizes time-series objective function

\( \Rightarrow \) Motivation for PCA
The Model

Cross-sectional objective function:

Minimize cross-sectional expected pricing error:

\[
\frac{1}{N} \sum_{i=1}^{N} \left( \hat{E}[X_i] - \hat{E}[F] \Lambda_i^T \right)^2
\]

\[
= \frac{1}{N} \sum_{i=1}^{N} \left( \frac{1}{T} X_i^T \mathbb{1} - \frac{1}{T} \mathbb{1}^T F \Lambda_i^T \right)^2
\]

\[
= \frac{1}{N} \text{trace} \left( \left( \frac{1}{T} \mathbb{1}^T X \Lambda \right) \left( \frac{1}{T} \mathbb{1}^T X \Lambda \right)^T \right) \quad \text{s.t.} \quad F = X(\Lambda^T \Lambda)^{-1} \Lambda^T
\]

- \( \mathbb{1} \) is vector \( T \times 1 \) of 1’s and thus \( \frac{F^T \mathbb{1}}{T} \) estimates factor mean

- Why not estimate factors with cross-sectional objective function?
  - Factors not identified
  - Spurious factor detection (Bryzgalova (2016))
### The Model

**Combined objective function:**

\[
\min_{\Lambda, F} \frac{1}{NT} \text{trace} \left( \left( (XM_{\Lambda})^T (XM_{\Lambda}) \right) \right) + \gamma \frac{1}{N} \text{trace} \left( \left( \frac{1}{T} \mathbb{1}^T XM_{\Lambda} \right) \left( \frac{1}{T} \mathbb{1}^T XM_{\Lambda} \right)^T \right)
\]

\[
= \min_{\Lambda} \frac{1}{NT} \text{trace} \left( M_{\Lambda} X^T \left( I + \frac{\gamma}{T} \mathbb{1} \mathbb{1}^T \right) XM_{\Lambda} \right)
\]

s.t. \( F = X (\Lambda^T \Lambda)^{-1} \Lambda^T \)

- The objective function is minimized by the eigenvectors of the largest eigenvalues of \( \frac{1}{NT} X^T (I_T + \frac{\gamma}{T} \mathbb{1} \mathbb{1}^T) X \).
- \( \hat{\Lambda} \) estimator for loadings: proportional to eigenvectors of the first \( K \) eigenvalues of \( \frac{1}{NT} X^T (I_T + \frac{\gamma}{T} \mathbb{1} \mathbb{1}^T) X \)
- \( \hat{F} \) estimator for factors: \( \frac{1}{N} X \hat{\Lambda} = X (\hat{\Lambda}^T \hat{\Lambda})^{-1} \hat{\Lambda}^T \).
- Estimator for the common component \( C = FA \) is \( \hat{C} = \hat{F} \hat{\Lambda}^T \)
The Model

Weighted Combined objective function:

Straightforward extension to weighted objective function:

$$\min_{\Lambda, F} \frac{1}{NT} \text{trace}(Q^T (X - F\Lambda^T)^T (X - F\Lambda^T)Q)$$

$$+ \gamma \frac{1}{N} \text{trace} \left( (1^T (X - F\Lambda^T)QQ^T (X - F\Lambda^T)^T 1 \right)$$

$$= \min_{\Lambda} \text{trace} \left( M_\Lambda Q^T X^T \left( I + \frac{\gamma}{T}11^T \right) XQM_\Lambda \right) \quad \text{s.t. } F = X(\Lambda^T \Lambda)^{-1} \Lambda^T$$

- Cross-sectional weighting matrix $Q$
- Factors and loadings can be estimated by applying PCA to $Q^T X^T \left( I + \frac{\gamma}{T}11^T \right) XQ$.
- Today: Only $Q$ equal to inverse of a diagonal matrix of standard deviations. For $\gamma = -1$ corresponds to PCA of a correlation matrix.
- Optimal choice of $Q$: GLS type argument
**The Model**

<table>
<thead>
<tr>
<th>Interpretation of Risk-Premium-PCA (RP-PCA):</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong> Time- and cross-sectional regression: Combines the time- and cross-sectional criteria functions.</td>
</tr>
<tr>
<td>- Select factors with small cross-sectional alpha’s.</td>
</tr>
<tr>
<td>- Protects against spurious factor with vanishing loadings as it requires the time-series errors to be small as well.</td>
</tr>
<tr>
<td><strong>2</strong> High Sharpe ratio factors: Search for factors explaining the time-series but penalizes low Sharpe-ratios.</td>
</tr>
<tr>
<td><strong>3</strong> Information interpretation: (GMM interpretation)</td>
</tr>
<tr>
<td>- PCA of a covariance matrix uses only the second moment but ignores first moment</td>
</tr>
<tr>
<td>- Using more information leads to more efficient estimates. RP-PCA combines first and second moments efficiently.</td>
</tr>
</tbody>
</table>
Signal-strengthening: Intuitively the matrix \( \frac{1}{T} X^\top (I_T + \frac{\gamma}{T} 11^\top) X \) converges to

\[
\Lambda \left( \Sigma_F + (1 + \gamma) \mu_F \mu_F^\top \right) \Lambda^\top + \text{Var}(e)
\]

with \( \Sigma_F = \text{Var}(F) \) and \( \mu_F = E[F] \). The signal of weak factors with a small variance can be “pushed up” by their mean with the right \( \gamma \).
The Model

Strong vs. weak factor models

- **Strong factor model** \( \left( \frac{1}{N} \Lambda^T \Lambda \text{ bounded} \right) \)
  - Interpretation: strong factors affect most assets (proportional to \( N \)), e.g. market factor
  - \( \Rightarrow \) RP-PCA always more efficient than PCA
  - \( \Rightarrow \) optimal \( \gamma \) relatively small

- **Weak factor model** \( (\Lambda^T \Lambda \text{ bounded}) \)
  - Interpretation: weak factors affect a smaller fraction of assets, e.g. value factor
  - \( \Rightarrow \) RP-PCA detects weak factors which cannot be detected by PCA
  - There exists a critical variance level, such that factors with \( \sigma_F^2 < \sigma_{crit}^2 \) cannot be estimated at all with PCA, but can reliably be estimated with RP-PCA.
  - \( \Rightarrow \) optimal \( \gamma \) relatively large
The Model

Strong vs. weak factor models

- Consequences for eigenvalues of $\frac{1}{T}X^\top X$:
  - Strong factors lead to exploding eigenvalues
  - Weak factors lead to large but bounded eigenvalues

- Empirical evidence (equity data): Strong and weak factors:
  - 1st eigenvalue typically substantially larger than rest of spectrum (usually 10 \times larger than the 2nd)
  - 2nd and 3rd eigenvalues typically stand out, but similar magnitudes as the rest of the spectrum
Illustration: Anomaly-sorted portfolios (Size and accrual)

Factors

1. **PCA**: Estimation based on PCA of correlation matrix, \( K = 3 \)
2. **RP-PCA**: Estimation based on PCA of \( X^T (I + \frac{\gamma}{T} 11^T) X \) (normalized standard deviation of \( X \)), \( K = 3 \) and \( \gamma = 100 \)
3. **Fama-French 5** factor model: market, size, value, profitability and investment
4. **Specific** factors: market, size and accrual

Data

- Double-sorted portfolios according to size and accrual (from Kenneth French’s website)
- Monthly return data from July 1963 to December 2013 (\( T = 606 \)) for \( N = 25 \) portfolios
Comparison among estimators

Goodness-of-fit-measures:

- **SR**: Sharpe ratio of the stochastic discount factor: $\left(\sqrt{\mu_F^\top \Sigma_F^{-1} \mu_F}\right)$.

- **Cross-sectional pricing error $\alpha$**:
  - **Time-series estimator**: Intercept of regression: $X_i = \alpha_i + F \Lambda_i + e_i$
  - **Cross-sectional estimator**: Regression of $E[X] = E[F]\Lambda^\top + \alpha$
  - Results the same. This presentation: Time-series regression $\alpha$.

- **RMS $\alpha$**: Root-mean-squared pricing errors $\sqrt{\frac{1}{N} \sum_{i=1}^{N} \alpha_i^2}$

- **Out-of-sample estimation**: Rolling window of 10 years ($T=120$) to estimate loadings for next month: $\hat{\alpha}_{t,i} = X_{t,i} - \hat{\alpha}_t$ with $\hat{\alpha}_t = X_t (\Lambda_{t-1} (\Lambda_{t-1}^\top \Lambda_{t-1})^{-1} \Lambda_{t-1}^\top)$.

- **Fama-MacBeth test-statistic** (weighted sum of squared $\alpha'$s, with $\chi^2_{N-K}$ distribution under $H_0$).
Portfolio Data: In-sample (Size and accrual)

<table>
<thead>
<tr>
<th></th>
<th>SR</th>
<th>RMS $\alpha$</th>
<th>Fama-MacBeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.305</td>
<td>0.068</td>
<td>44.570</td>
</tr>
<tr>
<td>PCA</td>
<td>0.135</td>
<td>0.141</td>
<td>89.946</td>
</tr>
<tr>
<td>Fama-French</td>
<td>0.344</td>
<td>0.154</td>
<td>61.979</td>
</tr>
<tr>
<td>Specific</td>
<td>0.173</td>
<td>0.155</td>
<td>76.041</td>
</tr>
</tbody>
</table>

**Table**: Maximal Sharpe-ratios, root-mean-squared pricing errors and Fama-MacBeth test statistics. $K = 3$ statistical factors and risk-premium weight $\gamma = 100$.

⇒ RP-PCA significantly better than PCA and quantile-sorted factors.
Cross-sectional $\alpha'$s for sorted portfolios (Size and Accrual)

$\Rightarrow$ RP-PCA avoids large pricing errors due to penalty term.
Loadings for statistical factors (Size and Accrual)

⇒ RP-PCA detects accrual factor while 3rd PCA factor is noise.
Maximal Incremental Sharpe Ratio

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>RP-PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>0.134</td>
<td>0.137</td>
</tr>
<tr>
<td>2 Factors</td>
<td>0.135</td>
<td>0.139</td>
</tr>
<tr>
<td>3 Factors</td>
<td>0.135</td>
<td>0.305</td>
</tr>
</tbody>
</table>

**Table:** Maximal Sharpe-ratio by adding factors incrementally. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$.

⇒ 1st and 2nd PCA and RP-PCA factors the same.
⇒ Better performance of RP-PCA because of third accrual factor.
Portfolio Data: Objective function (Size and Accrual)

<table>
<thead>
<tr>
<th></th>
<th>PCA TS</th>
<th>RP-PCA TS</th>
<th>PCA XS</th>
<th>RP-PCA XS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>3.308</td>
<td>3.617</td>
<td>0.014</td>
<td>0.002</td>
</tr>
<tr>
<td>2 Factors</td>
<td>1.937</td>
<td>2.240</td>
<td>0.014</td>
<td>0.002</td>
</tr>
<tr>
<td>3 Factors</td>
<td>1.623</td>
<td>1.751</td>
<td>0.014</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Table:** Time-series and cross-sectional objective functions.

- RP-PCA and PCA explain the same amount of variation.
- PR-PCA explains cross-sectional pricing much better.
- Motivation for risk-premium weight $\gamma = 100$. 
Portfolio Data: Out-of-sample (Size and Accrual)

<table>
<thead>
<tr>
<th></th>
<th>Out-of-sample</th>
<th>In-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.097</td>
<td>0.090</td>
</tr>
<tr>
<td>PCA</td>
<td>0.128</td>
<td>0.146</td>
</tr>
<tr>
<td>Fama-French 5</td>
<td>0.111</td>
<td>0.102</td>
</tr>
<tr>
<td>Specific</td>
<td>0.134</td>
<td>0.126</td>
</tr>
</tbody>
</table>

Table: Root-mean-squared pricing errors. Out-of-sample factors are estimated with a rolling window. $K = 3$ statistical factors and risk-premium weight $\gamma = 100$.

⇒ RP-PCA performs better in- and out-of-sample.
Cross-sectional $\alpha$’s out-of-sample (Size and Accrual)

⇒ RP-PCA avoids large pricing errors due to penalty term.
Weak Factor Model

- Weak factors either have a small variance or affect a smaller fraction of assets:
  - $\Lambda^T \Lambda$ bounded (after normalizing factor variances)
- Statistical model: Spiked covariance models from random matrix theory
- Eigenvalues of sample covariance matrix separate into two areas:
  - The bulk, majority of eigenvalues
  - The extremes, a few large outliers
- Bulk spectrum converges to generalized Marchenko-Pastur distribution (under certain conditions)
Weak Factor Model

- Large eigenvalues converge either to
  - A biased value characterized by the Stieltjes transform of the bulk spectrum
  - To the bulk of the spectrum if the true eigenvalue is below some critical threshold
  - Phase transition phenomena: estimated eigenvectors orthogonal to true eigenvectors if eigenvalues too small

- Onatski (2012): Weak factor model with phase transition phenomena

- Problem: All models in the literature assume that random processes have mean zero

  - RP-PCA implicitly uses non-zero means of random variables

  - New tools necessary!
Assumption 1: Weak Factor Model

1. Residual matrix can be represented as $e = \epsilon \Sigma$ with $\epsilon_{t,i} \sim N(0, 1)$. The empirical eigenvalue distribution function of $\Sigma$ converges to a non-random spectral distribution function with compact support. The supremum of the support is $b$.

2. The factors $F$ are uncorrelated among each other and are independent of $e$ and $\Lambda$ and have bounded first two moments.

\[
\hat{\mu}_F := \frac{1}{T} \sum_{t=1}^{T} F_t \xrightarrow{p} \mu_F \quad \hat{\Sigma}_F := \frac{1}{T} F_t F_t^T \xrightarrow{p} \Sigma_F = \begin{pmatrix} \sigma_{F_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{F_K}^2 \end{pmatrix}
\]

3. The column vectors of the loadings $\Lambda$ are orthogonally invariant and independent of $\epsilon$ and $F$ (e.g. $\Lambda_{i,k} \sim N(0, \frac{1}{N})$ and

\[
\Lambda^T \Lambda = I_K
\]

4. Assume that $\frac{N}{T} \to c$ with $0 < c < \infty$. 

Definition: Weak Factor Model

- Average idiosyncratic noise $\sigma^2_e := \text{trace}(\Sigma)/N$

- Denote by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$ the ordered eigenvalues of $\frac{1}{T} e^T e$. The Cauchy transform (also called Stieltjes transform) of the eigenvalues is the almost sure limit:

$$G(z) := \text{a.s.} \lim_{T \to \infty} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{z - \lambda_i} = \text{a.s.} \lim_{T \to \infty} \frac{1}{N} \text{trace} \left( (z I_N - \frac{1}{T} e^T e)^{-1} \right)$$

- $B$-function

$$B(z) := \text{a.s.} \lim_{T \to \infty} \frac{c}{N} \sum_{i=1}^{N} \frac{\lambda_i}{(z - \lambda_i)^2}$$

$$= \text{a.s.} \lim_{T \to \infty} \frac{c}{N} \text{trace} \left( \left( (z I_N - \frac{1}{T} e^T e)^{-2} \left( \frac{1}{T} e^T e \right) \right) \right)$$
Weak Factor Model

Estimator

- Risk-premium PCA (RP-PCA): Apply PCA estimation to
  \[ S_\gamma := \frac{1}{T} X^\top \left( I_T + \gamma \frac{11^T}{T} \right) X \]

- PCA: Apply PCA to estimated covariance matrix
  \[ S_{-1} := \frac{1}{T} X^\top \left( I_T - \frac{11^T}{T} \right) X, \text{ i.e. } \gamma = -1. \]

⇒ PCA special case of RP-PCA

“Signal” Matrix for Covariance PCA

\[ M_{Var} = \Sigma_F + c\sigma_e^2 I_K = \begin{pmatrix} \sigma_{F_1}^2 + c\sigma_e^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{F_K}^2 + c\sigma_e^2 \end{pmatrix} \]

⇒ Intuition: Largest \( K \) “true” eigenvalues of \( S_{-1} \).
Lemma: Covariance PCA

Assumption 1 holds. Define the critical value $\sigma^2_{\text{crit}} = \lim_{z \downarrow b} \frac{1}{G(z)}$. The first $K$ largest eigenvalues $\hat{\lambda}_i$ of $S^{-1}$ satisfy for $i = 1, \ldots, K$

$$\hat{\lambda}_i \xrightarrow{p} \begin{cases} G^{-1} \left( \frac{1}{\sigma^2_{F_i} + c\sigma^2_e} \right) & \text{if } \sigma^2_{F_i} + c\sigma^2_e > \sigma^2_{\text{crit}} \\ b & \text{otherwise} \end{cases}$$

The correlation between the estimated and true factors converges to

$$\tilde{\text{Corr}}(F, \hat{F}) \xrightarrow{p} \begin{pmatrix} \varrho_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \varrho_K \end{pmatrix}$$

with

$$\varrho^2_i \xrightarrow{p} \begin{cases} \frac{1}{1 + (\sigma^2_{F_i} + c\sigma^2_e)B(\hat{\lambda}_i)} & \text{if } \sigma^2_{F_i} + c\sigma^2_e > \sigma^2_{\text{crit}} \\ 0 & \text{otherwise} \end{cases}$$
Weak Factor Model

Corollary: Covariance PCA for i.i.d. errors

Assumption 1 holds, \( c \geq 1 \) and \( e_{t,i} \text{ i.i.d. } N(0, \sigma_e^2) \). The largest \( K \) eigenvalues of \( S_{-1} \) have the following limiting values:

\[
\hat{\lambda}_i \xrightarrow{p} \begin{cases} 
\sigma_{F_i}^2 + \frac{\sigma_e^2}{\sigma_{F_i}^2} (c + 1 + \sigma_e^2) & \text{if } \sigma_{F_i}^2 + c\sigma_e^2 > \sigma_{\text{crit}}^2 \Leftrightarrow \sigma_{F}^2 > \sqrt{c}\sigma_e^2 \\
\sigma_e^2(1 + \sqrt{c})^2 & \text{otherwise}
\end{cases}
\]

The correlation between the estimated and true factors converges to

\[
\widehat{\text{Corr}}(F, \hat{F}) \xrightarrow{p} \begin{pmatrix} \varrho_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \varrho_K \end{pmatrix}
\]

with

\[
\varrho_i^2 \xrightarrow{p} \begin{cases} 
1 - \frac{c\sigma_e^4}{\sigma_{F_i}^4} & \text{if } \sigma_{F_i}^2 + c\sigma_e^2 > \sigma_{\text{crit}}^2 \\
\frac{c\sigma_e^2}{\sigma_{F_i}^2} + \frac{\sigma_e^4}{\sigma_{F_i}^4} (c^2 - c) & \text{otherwise}
\end{cases}
\]
Weak Factor Model

"Signal" Matrix for RP-PCA

- Define $\tilde{\gamma} = \sqrt{\gamma + 1} - 1$ and note that $(1 + \tilde{\gamma})^2 = 1 + \gamma$.

$\Rightarrow$ Projection on $K$ demeaned factors and on mean operator.

- Denote by $\theta_1 \geq ... \geq \theta_{K+1}$ the eigenvalues of the "signal matrix" $M_{RP}$ and by $\tilde{U}$ the corresponding orthonormal eigenvectors:

$$\tilde{U}^\top M_{RP} \tilde{U} = \begin{pmatrix} \theta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \theta_{K+1} \end{pmatrix}$$

$\Rightarrow$ Intuition: $\theta_1, ..., \theta_{K+1}$ largest $K + 1$ "true" eigenvalues of $S_{\gamma}$. 

\[ M_{RP} = \begin{pmatrix} \Sigma_F + c\sigma^2_e & \Sigma_F^{1/2} \mu_F(1 + \tilde{\gamma}) \\ \mu_F^\top \Sigma_F^{1/2} (1 + \tilde{\gamma}) & (1 + \gamma)(\mu_F^\top \mu_F + c\sigma^2_e) \end{pmatrix} \]
Theorem 1: Risk-Premium PCA under weak factor model

Assumption 1 holds. The first $K$ largest eigenvalues $\hat{\theta}_i$, $i = 1, \ldots, K$ of $S_\gamma$ satisfy

$$\hat{\theta}_i \xrightarrow{p} \begin{cases} G^{-1} \left( \frac{1}{\hat{\theta}_i} \right) b & \text{if } \theta_i > \sigma_{crit}^2 = \lim_{z \downarrow b} \frac{1}{G(z)} \\ b & \text{otherwise} \end{cases}$$

The correlation of the estimated with the true factors converges to

$$\widehat{\text{Corr}}(F, \hat{F}) \xrightarrow{p} \begin{pmatrix} I_K & 0 \end{pmatrix} \tilde{U} \begin{pmatrix} \rho_1 & 0 & \cdots & 0 \\ 0 & \rho_2 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & \rho_K \end{pmatrix} D_K^{1/2} \hat{\Sigma}^{-1/2} \hat{f}$$

with

$$\rho_i^2 \xrightarrow{p} \begin{cases} 1 & \text{if } \theta_i > \sigma_{crit}^2 \\ \frac{1}{1 + \theta_i B(\hat{\theta}_i)} & \text{otherwise} \end{cases}$$
Theorem 1: continued

\[ \hat{\Sigma}_F = D_K^{1/2} \left( \begin{pmatrix} \rho_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_K \end{pmatrix}^T \bar{U}^T \begin{pmatrix} I_K & 0 \\ 0 & 0 \end{pmatrix} \bar{U} \begin{pmatrix} \rho_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_K \end{pmatrix} \right) 
\]

\[ + \begin{pmatrix} 1 - \rho_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 - \rho_K^2 \end{pmatrix} \right) D_K^{1/2} \]

\[ D_K = \text{diag} \left( (\hat{\theta}_1 \ \cdots \ \hat{\theta}_K) \right) \]
Lemma: Detection of weak factors

If $\gamma > -1$ and $\mu_F \neq 0$, then the first $K$ eigenvalues of $M_{RP}$ are strictly larger than the first $K$ eigenvalues of $M_{Var}$, i.e.

$$\theta_i > \sigma^2_{F_i} + c\sigma^2_e$$

For $\theta_i > \sigma^2_{crit}$ it holds that

$$\frac{\partial \hat{\theta}_i}{\partial \theta_i} > 0 \quad \frac{\partial \rho_i}{\partial \theta_i} > 0 \quad i = 1, ..., K$$

Thus, if $\gamma > -1$ and $\mu_F \neq 0$, then $\rho_i > \varrho_i$.

$\Rightarrow$ For $\mu_F \neq 0$ RP-PCA always better than PCA.
Example: One-factor model

Assume that there is only one factor, i.e. \( K = 1 \). The “signal matrix” \( M_{RP} \) simplifies to

\[
M_{RP} = \begin{pmatrix}
    
    \sigma_F^2 + c\sigma_e^2 & \sigma_F \mu (1 + \tilde{\gamma}) \\
    \mu \sigma_F (1 + \tilde{\gamma}) & (\mu^2 + c\sigma_e^2)(1 + \gamma)
\end{pmatrix}
\]

and has the eigenvalues:

\[
\theta_{1,2} = \frac{1}{2} \sigma_F^2 + c\sigma_e^2 + (\mu^2 + c\sigma_e^2)(1 + \gamma)
\]

\[
\pm \frac{1}{2} \sqrt{(\sigma_F^2 + c\sigma_e^2 + (\mu^2 + c\sigma_e^2)(1 + \gamma))^2 - 4(1 + \gamma)c\sigma_e^2(\sigma_F^2 + \mu^2 + c\sigma_e^2)}
\]

The eigenvector of first eigenvalue \( \theta_1 \) has the components

\[
\tilde{U}_{1,1} = \frac{\mu \sigma_F (1 + \tilde{\gamma})}{\sqrt{(\theta_1 - (\sigma_F^2 + c\sigma_e^2))^2 + \mu^2\sigma_F^2(1 + \gamma)}}
\]

\[
\tilde{U}_{1,2} = \frac{\theta_1 - \sigma_F^2 + c\sigma_e^2}{\sqrt{(\theta_1 - (\sigma_F^2 + c\sigma_e^2))^2 + \mu^2\sigma_F^2(1 + \gamma)}}
\]
Corollary: One-factor model

The correlation between the estimated and true factor has the following limit:

$$\hat{\text{Corr}}(F, \hat{F}) \xrightarrow{p} \frac{\rho_1}{\sqrt{\rho_1^2 + (1 - \rho_1^2) \left(\frac{\theta_1 - (\sigma_F^2 + c\sigma_e^2))^2 + 1}{\mu^2\sigma_F^2(1+\gamma)}\right)}}$$
Strong Factor Model

- Strong factors affect most assets: e.g. market factor
- $\frac{1}{N} \Lambda^T \Lambda$ bounded (after normalizing factor variances)
- Statistical model: Bai and Ng (2002) and Bai (2003) framework
- Factors and loadings can be estimated consistently and are asymptotically normal distributed
- RP-PCA provides a more efficient estimator of the loadings
- Assumptions essentially identical to Bai (2003)
Strong Factor Model

Asymptotic Distribution (up to rotation)

- PCA under assumptions of Bai (2003):
  - Asymptotically \( \hat{\Lambda} \) behaves like OLS regression of \( F \) on \( X \).
  - Asymptotically \( \hat{\Lambda} \) behaves like OLS regression of \( \Lambda \) on \( X \).
- RP-PCA under slightly stronger assumptions as in Bai (2003):
  - Asymptotically \( \hat{\Lambda} \) behaves like OLS regression of \( FW \) on \( XW \)
    with \( W^2 = \left( I_T + \gamma \frac{11^\top}{T} \right) \).
  - Asymptotically \( \hat{\Lambda} \) behaves like OLS regression of \( \Lambda \) on \( X \).

Asymptotic Expansion

Asymptotic expansions (under slightly stronger assumptions as in Bai (2003)):

1. \[
\sqrt{T} \left( H^\top \hat{\Lambda}_i - \Lambda_i \right) = \left( \frac{1}{T} F^\top W^2 F \right)^{-1} \frac{1}{\sqrt{T}} F^\top W^2 e_i + O_p \left( \frac{\sqrt{T}}{N} \right) + o_p(1)
\]
2. \[
\sqrt{N} \left( H^\top \hat{\Lambda}_t - \Lambda_t \right) = \left( \frac{1}{N} \Lambda^\top \Lambda \right)^{-1} \frac{1}{\sqrt{N}} \Lambda^\top e_t + O_p \left( \frac{\sqrt{N}}{T} \right) + o_p(1)
\]
with known rotation matrix \( H \).
Assumption 2: Strong Factor Model

Assume the same assumptions as in Bai (2003) (Assumption A-G) hold and in addition

\[
\left( \frac{1}{\sqrt{T}} \sum_{t=1}^{T} F_t e_{t,i} \right) \overset{D}{\to} N(0, \Omega) \quad \Omega = \begin{pmatrix} \Omega_{1,1} & \Omega_{1,2} \\ \Omega_{2,1} & \Omega_{2,2} \end{pmatrix}
\]
Theorem 2: Strong Factor Model

Assumption 2 holds and $\gamma \in [-1, \infty)$. Then:

- For any choice of $\gamma$ the factors, loadings and common components can be estimated consistently pointwise.
- If $\frac{\sqrt{N}}{T} \to 0$ then $\sqrt{T} \left( H^T \hat{\Lambda}_i - \Lambda_i \right) \xrightarrow{D} N(0, \Phi)$

$$
\Phi = \left( \Sigma_F + (\gamma + 1) \mu_F \mu_F^T \right)^{-1} \left( \Omega_{1,1} + \gamma \mu_F \Omega_{2,1} + \gamma \Omega_{1,2} \mu_F + \gamma^2 \mu_F \Omega_{2,2} \mu_F \right) \\
\cdot \left( \Sigma_F + (\gamma + 1) \mu_F \mu_F^T \right)^{-1}
$$

For $\gamma = -1$ this simplifies to the conventional case $\Sigma_F^{-1} \Omega_{1,1} \Sigma_F^{-1}$.

- The asymptotic distribution of the factors is not affected by the choice of $\gamma$.
- The asymptotic distribution of the common component depends on $\gamma$ if and only if $\frac{N}{T}$ does not go to zero. For $\frac{T}{N} \to 0$ $\sqrt{T} \left( \hat{C}_{t,i} - C_{t,i} \right) \xrightarrow{D} N(0, F_t^T \Phi F_t)$
Example 2: Toy model with i.i.d. residuals and $K = 1$

Assume $K = 1$ and $e_{t,i} \sim^{i.i.d.} (0, \sigma^2_e)$. If Assumption 2 holds and $\frac{\sqrt{T}}{N} \to 0$, then

$$\sqrt{T} \left( \hat{\Lambda}_i - \Lambda_i \right) \overset{D}{\to} N(0, \Omega)$$

with

$$\Omega = \sigma_e^2 \frac{\left( \sigma_F^2 + \mu_F^2 (1 + \gamma)^2 \right)}{\left( \sigma_F^2 + \mu_F^2 (1 + \gamma) \right)^2}$$

⇒ Optimal choice minimizing the asymptotic variance is risk-premium weight $\gamma = 0$.

⇒ Choosing $\gamma = -1$, i.e. the covariance matrix for factor estimation, is not efficient.
Simulation

Simulation parameters

- \( N = 250 \) and \( T = 350 \).

- Factors: \( K = 4 \)
  - 1. Factor represent the market with \( N(1.2, 9) \): Sharpe-ratio of 0.4
  - 2. Factor represents an industry factors following \( N(0.1, 1) \): Sharpe-ratio of 0.1.
  - 3. Factor follows \( N(0.4, 1) \): Sharpe-ratio of 0.4.
  - 4. Factor has a small variance but high Sharpe-ratio. It follows \( N(0.4, 0.16) \): Sharpe-ratio of 1.

- Loadings normalized such that \( \frac{1}{N} \Lambda^\top \Lambda \).
  \( \Lambda_{i,1} = 1 \) and \( \Lambda_{i,k} \sim N(0, 1) \) for \( k = 2, 3, 4 \).

- Errors: Cross-sectional and time-series correlation and heteroskedasticity in the residuals. Half of the variation due to non-systematic risk.
**Simulation**

![Graph](image)

**Figure:** Sample path of the first four factors and the estimated factor processes. $\gamma = 50$. 

---

**Intro**  **Model**  **Illustration**  **Weak Factors**  **Strong F.**  **Simulation**  **Empirical Results**  **Extension**  **Conclusion**  **Appendix**
### Simulation

#### Table: Average root-mean-squared (RMS) errors of estimated factors relative to the true factor processes. $\gamma = 50.$

<table>
<thead>
<tr>
<th>Factor</th>
<th>PCA Var</th>
<th>PCA Corr</th>
<th>RP-PCA</th>
<th>RP-PCA Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0.094</td>
<td>0.086</td>
<td>0.042</td>
<td>0.040</td>
</tr>
<tr>
<td>2.</td>
<td>0.023</td>
<td>0.022</td>
<td>0.025</td>
<td>0.022</td>
</tr>
<tr>
<td>3.</td>
<td>0.100</td>
<td>0.095</td>
<td>0.079</td>
<td>0.074</td>
</tr>
<tr>
<td>4. <strong>Factor</strong></td>
<td><strong>0.312</strong></td>
<td><strong>0.312</strong></td>
<td><strong>0.183</strong></td>
<td><strong>0.170</strong></td>
</tr>
</tbody>
</table>
Squared correlations between estimated and true factor based on the weak factor model prediction and Monte-Carlo simulations for different variances of the factor. The Sharpe-ratio of the factor is 1, i.e. the mean equals $\mu_F = \sigma_F$. The normalized variance of the factors is $\sigma_F^2 \cdot N$. 
Weak Factor Model: Dependent residuals

Figure: Values of $\rho_i^2 \left( \frac{1}{1 + \theta_i B(\hat{\theta}_i)} \right)$ if $\theta_i > \sigma^2_{\text{crit}}$ and 0 otherwise) for different signals $\theta_i$. The average noise level is normalized in both cases to $\sigma^2_e = 1$ and $c = 1$. For the correlated residuals we assume that $\Sigma^{1/2}$ is a Toeplitz matrix with $\beta, \beta, \beta, \beta^2$ on the right four off-diagonals with $\beta = 0.7$. 
Empirical Results

Portfolio Data

- Data
  - Monthly return data from July 1963 to December 2013 ($T = 606$)
  - 13 double sorted portfolios (consisting of 25 portfolios) from Kenneth French’s website and 49 industry portfolios

- Factors
  1. **PCA**: $K = 3$
  2. **RP-PCA**: $K = 3$ and $\gamma = 100$
  3. **Fama-French 5** factor model: market, size, value, profitability and investment
  4. **Specific** factors: market + two specific anomaly long-short factors
## Empirical Results

### Pricing errors $\alpha$ (in-sample)

<table>
<thead>
<tr>
<th></th>
<th>RP-PCA</th>
<th>PCA</th>
<th>FF 5</th>
<th>Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size and BM</td>
<td>0.13</td>
<td>0.14</td>
<td>0.12</td>
<td>0.20</td>
</tr>
<tr>
<td><strong>BM and Investment</strong></td>
<td><strong>0.07</strong></td>
<td><strong>0.12</strong></td>
<td><strong>0.14</strong></td>
<td><strong>0.13</strong></td>
</tr>
<tr>
<td>BM and Operating Profits</td>
<td>0.11</td>
<td>0.12</td>
<td>0.14</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Size and Accrual</strong></td>
<td><strong>0.07</strong></td>
<td><strong>0.14</strong></td>
<td><strong>0.15</strong></td>
<td><strong>0.16</strong></td>
</tr>
<tr>
<td>Size and Beta</td>
<td>0.06</td>
<td>0.07</td>
<td>0.08</td>
<td>0.17</td>
</tr>
<tr>
<td>Size and Investment</td>
<td>0.11</td>
<td>0.13</td>
<td>0.11</td>
<td>0.20</td>
</tr>
<tr>
<td>Size and Operating Profits</td>
<td>0.06</td>
<td>0.07</td>
<td>0.08</td>
<td>0.16</td>
</tr>
<tr>
<td>Size and Short-Term Reversal</td>
<td>0.15</td>
<td>0.16</td>
<td>0.24</td>
<td>0.33</td>
</tr>
<tr>
<td>Size and Long-Term Reversal</td>
<td>0.11</td>
<td>0.13</td>
<td>0.09</td>
<td>0.20</td>
</tr>
<tr>
<td>Size and Res. Var.</td>
<td>0.17</td>
<td>0.18</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>Size and Total Var.</td>
<td>0.18</td>
<td>0.19</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td><strong>Operating Profits and Investment</strong></td>
<td><strong>0.11</strong></td>
<td><strong>0.14</strong></td>
<td><strong>0.12</strong></td>
<td><strong>0.14</strong></td>
</tr>
<tr>
<td>Size and Net Share Iss.</td>
<td>0.14</td>
<td>0.16</td>
<td>0.13</td>
<td>0.17</td>
</tr>
<tr>
<td>49 Industries</td>
<td>0.14</td>
<td>0.16</td>
<td>0.13</td>
<td>0.29</td>
</tr>
</tbody>
</table>
## Empirical Results

### Pricing errors $\alpha$ (out-of-sample)

<table>
<thead>
<tr>
<th></th>
<th>RP-PCA</th>
<th>PCA</th>
<th>FF 5</th>
<th>Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size and BM</td>
<td>0.17</td>
<td>0.19</td>
<td>0.14</td>
<td>0.21</td>
</tr>
<tr>
<td><strong>BM and Investment</strong></td>
<td><strong>0.12</strong></td>
<td><strong>0.16</strong></td>
<td><strong>0.11</strong></td>
<td><strong>0.14</strong></td>
</tr>
<tr>
<td>BM and Operating Profits</td>
<td>0.15</td>
<td>0.18</td>
<td>0.15</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Size and Accrual</strong></td>
<td><strong>0.10</strong></td>
<td><strong>0.13</strong></td>
<td><strong>0.11</strong></td>
<td><strong>0.13</strong></td>
</tr>
<tr>
<td>Size and Beta</td>
<td>0.09</td>
<td>0.10</td>
<td>0.07</td>
<td>0.09</td>
</tr>
<tr>
<td>Size and Investment</td>
<td>0.14</td>
<td>0.17</td>
<td>0.12</td>
<td>0.19</td>
</tr>
<tr>
<td>Size and Operating Profits</td>
<td>0.09</td>
<td>0.12</td>
<td>0.09</td>
<td>0.18</td>
</tr>
<tr>
<td>Size and Short-Term Reversal</td>
<td>0.17</td>
<td>0.19</td>
<td>0.09</td>
<td>0.18</td>
</tr>
<tr>
<td>Size and Long-Term Reversal</td>
<td>0.13</td>
<td>0.14</td>
<td>0.09</td>
<td>0.14</td>
</tr>
<tr>
<td>Size and Res. Var.</td>
<td>0.17</td>
<td>0.20</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td>Size and Total Var.</td>
<td>0.17</td>
<td>0.21</td>
<td>0.20</td>
<td>0.26</td>
</tr>
<tr>
<td><strong>Operating Profits and Investment</strong></td>
<td><strong>0.13</strong></td>
<td><strong>0.17</strong></td>
<td><strong>0.13</strong></td>
<td><strong>0.16</strong></td>
</tr>
<tr>
<td>Size and Net Share Iss.</td>
<td>0.14</td>
<td>0.21</td>
<td>0.16</td>
<td>0.18</td>
</tr>
<tr>
<td>49 Industries</td>
<td>0.26</td>
<td>0.24</td>
<td>0.21</td>
<td>0.25</td>
</tr>
</tbody>
</table>
### Empirical Results

#### Maximum Sharpe-Ratios

<table>
<thead>
<tr>
<th></th>
<th>RP-PCA</th>
<th>PCA</th>
<th>Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size and BM</td>
<td>0.25</td>
<td>0.22</td>
<td>0.16</td>
</tr>
<tr>
<td><strong>BM and Investment</strong></td>
<td><strong>0.26</strong></td>
<td><strong>0.17</strong></td>
<td><strong>0.24</strong></td>
</tr>
<tr>
<td>BM and Operating Profits</td>
<td>0.24</td>
<td>0.22</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>Size and Accrual</strong></td>
<td><strong>0.30</strong></td>
<td><strong>0.13</strong></td>
<td><strong>0.17</strong></td>
</tr>
<tr>
<td>Size and Beta</td>
<td>0.23</td>
<td>0.23</td>
<td>0.17</td>
</tr>
<tr>
<td>Size and Investment</td>
<td>0.30</td>
<td>0.26</td>
<td>0.23</td>
</tr>
<tr>
<td>Size and Operating Profits</td>
<td>0.22</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td>Size and Short-Term Reversal</td>
<td>0.26</td>
<td>0.20</td>
<td>0.25</td>
</tr>
<tr>
<td>Size and Long-Term Reversal</td>
<td>0.23</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td>Size and Res. Var.</td>
<td>0.33</td>
<td>0.30</td>
<td>0.32</td>
</tr>
<tr>
<td>Size and Total Var.</td>
<td>0.32</td>
<td>0.28</td>
<td>0.32</td>
</tr>
<tr>
<td><strong>Operating Profits and Investment</strong></td>
<td><strong>0.31</strong></td>
<td><strong>0.24</strong></td>
<td><strong>0.34</strong></td>
</tr>
<tr>
<td><strong>Size and Net Share Iss.</strong></td>
<td><strong>0.33</strong></td>
<td><strong>0.25</strong></td>
<td><strong>0.35</strong></td>
</tr>
<tr>
<td>49 Industries</td>
<td>0.35</td>
<td>0.25</td>
<td>0.11</td>
</tr>
</tbody>
</table>
Empirical Results

Portfolio Data

- Monthly return data from July 1963 to December 2013 ($T = 606$) for $N = 199$ portfolios
  - Novy-Marx and Velikov (2014) data: 150 portfolios sorted according to 15 anomalies (same data as in Kozak, Nagel and Santosh (2015))
  - 49 industry portfolios from Kenneth French’s website

1. **Fama-French 5**: The five factor model of Fama-French (market, size, value, investment and operating profitability, all from Kenneth French’s website).
2. **Specific**: Market, value, value-momentum-profitibility and volatility factors.

- Number of statistical factors $K = 4$ and $\gamma = 100$. 
Empirical Results

Portfolio Data I: 15 Novy-Marx factors and portfolios

- Size
- Gross Profitability
- Value
- Value Prof
- Accruals
- Net Issuance
- Asset Growth
- Investment
- Piotrotski F-Score
- ValMomProf
- ValMom
- Idiosyncratic Vol
- Momentum
- Long Run Reversal
- Beta Arbitrage.
Empirical Results

Portfolio Data: In-sample

<table>
<thead>
<tr>
<th></th>
<th>SR</th>
<th>RMS</th>
<th>Fama-MacBeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.417</td>
<td>0.135</td>
<td>729.944</td>
</tr>
<tr>
<td>PCA</td>
<td>0.155</td>
<td>0.213</td>
<td>820.804</td>
</tr>
<tr>
<td>Fama-French</td>
<td>0.344</td>
<td>0.225</td>
<td>801.013</td>
</tr>
<tr>
<td>Specific</td>
<td>0.413</td>
<td>0.152</td>
<td>731.392</td>
</tr>
</tbody>
</table>

Table: Maximal Sharpe-ratios, root-mean-squared pricing errors and Fama-MacBeth test statistics. $K = 4$ statistical factors and risk-premium weight $\gamma = 100$.

- RP-PCA strongly dominates PCA and Fama-French 5 factors
- Specific factors (Market, Value, Value-Momentum-Profitibility and Volatility) perform similar to RP-PCA.
### Empirical Results

#### Portfolio Data: Out-of-sample

<table>
<thead>
<tr>
<th></th>
<th>Out-of-sample</th>
<th>In-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.178</td>
<td>0.145</td>
</tr>
<tr>
<td>PCA</td>
<td>0.202</td>
<td>0.208</td>
</tr>
<tr>
<td>Fama-French 5</td>
<td>0.182</td>
<td>0.182</td>
</tr>
<tr>
<td>Specific</td>
<td>0.154</td>
<td>0.137</td>
</tr>
</tbody>
</table>

**Table:** Root-mean-squared pricing errors. Out-of-sample factors are estimated with a rolling window. $K = 4$ statistical factors and risk-premium weight $\gamma = 100$.

⇒ RP-PCA performs well in- and out-of-sample.
Portfolio Data: Interpreting factors

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>RP-PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Gen. Corr.</td>
<td>0.997</td>
<td>0.997</td>
</tr>
<tr>
<td>2. Gen. Corr.</td>
<td>0.898</td>
<td>0.925</td>
</tr>
<tr>
<td>3. Gen. Corr.</td>
<td>0.809</td>
<td>0.888</td>
</tr>
<tr>
<td>4. Gen. Corr.</td>
<td>0.032</td>
<td>0.741</td>
</tr>
</tbody>
</table>

Table: Generalized Correlations between specific factors and statistical factors.

- Problem in interpreting factors: Factors only identified up to invertible linear transformations.
- Generalized correlations close to 1 measure of how many factors two sets have in common.
- Specific factors: Market, Value, Value-Momentum-Profitability and Volatility factors.

⇒ Specific factors approximate RP-PCA factors.
Empirical Results

Maximal Incremental Sharpe Ratio

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>RP-PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>0.127</td>
<td>0.137</td>
</tr>
<tr>
<td>2 Factors</td>
<td>0.149</td>
<td>0.381</td>
</tr>
<tr>
<td>3 Factors</td>
<td>0.153</td>
<td>0.412</td>
</tr>
</tbody>
</table>

Table: Maximal Sharpe-ratio by adding factors incrementally. $K = 4$ statistical factors and risk-premium weight $\gamma = 100$. 
Empirical Results

Portfolio Data: Objective function

<table>
<thead>
<tr>
<th></th>
<th>PCA TS</th>
<th>RP-PCA TS</th>
<th>PCA XS</th>
<th>RP-PCA XS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>44.771</td>
<td>51.623</td>
<td>0.298</td>
<td>0.037</td>
</tr>
<tr>
<td>2 Factors</td>
<td>39.846</td>
<td>42.326</td>
<td>0.268</td>
<td>0.001</td>
</tr>
<tr>
<td>3 Factors</td>
<td>36.112</td>
<td>37.849</td>
<td>0.263</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table: Time-series and cross-sectional objective functions.

⇒ RP-PCA and PCA explain the same amount of variation.
⇒ PR-PCA explains cross-sectional pricing much better.
⇒ Motivation for risk-premium weight $\gamma = 100$. 
Empirical Results

Cumulative returns of factors

![Graph showing cumulative returns of factors over the years 1963 to 2014. The y-axis represents the cumulative return, and the x-axis represents the years. The graph compares various models such as Mkt-Rf, RP-PCA 1, RP-PCA 2, RP-PCA 3, and RP-PCA 4. The returns are plotted against each other to illustrate their performance over time.](image-url)
Extension: Time-varying loadings

Model with time-varying loadings

- Observe panel of excess returns and $L$ covariates $Z_{i,t-1,l}$:
  \[
  X_{t,i} = F_t^\top g (Z_{i,t-1,1}, \ldots, Z_{i,t-1,L}) + e_{t,i}
  \]

- Loadings are function of $L$ covariates $Z_{i,t-1,l}$ with $l = 1, \ldots, L$
  e.g. characteristics like size, book-to-market ratio, past returns, ...

- Factors and loading function are latent

Literature (partial list)

- Dynamic semiparametric factor model: Park, Mammen, Härdle and Borak (2009)
- Nonparametric regression model: Connor and Linton (2007)
Extension: Time-varying loadings

Projected RP-PCA (work in progress)

- Assume additive nonparametric loading model:

\[ g_k(Z_{i,t-1}) = \sum_{l=1}^{L} g_{k,l}(Z_{i,t-1,l}) \]

- Each additive component of \( g_k \) is estimated by the sieve method.

- Choose appropriate basis functions \( \phi_1(.) , ..., \phi_D(.) \) (e.g. splines, polynomial series, kernels, etc.)

- Define projection \( P_{t-1} \) as regression on \( L \cdot D \times N \) matrix \( \phi(Z_{t-1}) \) with elements \( \phi_d(Z_{i,t-1,l}) , i = 1, ..., N , l = 1, ..., L , d = 1, ..., D. \)

- Apply RP-PCA to projected data \( \tilde{X}_t = X_t P_{t-1}. \)

- Empirical results promising: We recover size, value, momentum and volatility factors from individual stock price data
Conclusion

Methodology

- New estimator for estimating priced latent factors from large data sets
- Combines time-series and cross-sectional criterion function
- Asymptotic theory under weak and strong factor model assumption
- Detects weak factors with high Sharpe-ratio
- More efficient than conventional PCA

Empirical Results

- Strongly dominates estimation based on PCA of the covariance matrix
- Potential to provide benchmark factors for horse races.
- Promising empirical results.
Predicted excess return in-sample (Size and Accrual)
Predicted excess return out-of-sample (Size and Accrual)

- **OLS out-of-sample RP-PCA**
- **OLS out-of-sample PCA**
- **OLS out-of-sample 5 Fama-French factors**
- **OLS out-of-sample Specific factors**
Asymptotic expansions (under slightly stronger assumptions as in Bai (2003)):

1. \[ \sqrt{T} \left( H^T \hat{\Lambda}_i - \Lambda_i \right) = \left( \frac{1}{T} F^T W^2 F \right)^{-1} \frac{1}{\sqrt{T}} F^T W^2 e_i + O_p \left( \frac{\sqrt{T}}{N} \right) + o_p(1) \]

2. \[ \sqrt{N} \left( H^{-1} \hat{F}_t - F_t \right) = \left( \frac{1}{N} \Lambda^\top \Lambda \right)^{-1} \frac{1}{\sqrt{N}} \Lambda^\top e^\top_t + O_p \left( \frac{\sqrt{N}}{T} \right) + o_p(1) \]

3. \[ \sqrt{\delta} \left( \hat{C}_{t,i} - C_{t,i} \right) = \sqrt{\delta} \frac{1}{\sqrt{T}} F^T_t \left( \frac{1}{T} F^T W^2 F \right)^{-1} \frac{1}{\sqrt{T}} F^T W^2 e_i + \sqrt{\delta} \frac{1}{\sqrt{N}} \Lambda^\top_i \left( \frac{1}{N} \Lambda^\top \Lambda \right)^{-1} \frac{1}{\sqrt{N}} \Lambda^\top e^\top_t + o_p(1) \]

with \( H = \left( \frac{1}{T} F^T W^2 F \right) \left( \frac{1}{N} \Lambda^\top \hat{\Lambda} \right) V_{TN}^{-1} \), \( \delta = \min(N, T) \) and \( W^2 = \left( I_T + \gamma \frac{11^\top}{T} \right) \).
Simulation parameters

**Errors**

Residuals are modeled as $e = \sigma_e D_T A_T \epsilon A_N D_N$:

- $\epsilon$ is a $T \times N$ matrix and follows a multivariate standard normal distribution
- Time-series correlation in errors: $A_T$ creates an AR(1) model with parameter $\rho = 0.1$
- Cross-sectional correlation in errors: $A_N$ is a Toeplitz-matrix with $(\beta, \beta, \beta, \beta^2)$ on the right four off-diagonals with $\beta = 0.7$
- Cross-sectional heteroskedasticity: $D_N$ is a diagonal matrix with independent elements following $N(1, 0.2)$
- Time-series heteroskedasticity: $D_T$ is a diagonal matrix with independent elements following $N(1, 0.2)$
- Signal-to-noise ratio: $\sigma_e^2 = 10$
- Parameters produce eigenvalues that are consistent with the data.
## Simulation

<table>
<thead>
<tr>
<th>True Factors</th>
<th>PCA Var</th>
<th>PCA Corr</th>
<th>RP-PCA</th>
<th>PR-PCA Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR</td>
<td>1.330</td>
<td>0.515</td>
<td>0.517</td>
<td><strong>0.865</strong></td>
</tr>
</tbody>
</table>

**Table**: Maximal Sharpe Ratio with $K = 4$ factors. $\gamma = 50$.

<table>
<thead>
<tr>
<th>True</th>
<th>PCA Var</th>
<th>PCA Corr</th>
<th>RP-PCA</th>
<th>RP-PCA Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Factor</td>
<td>1.20</td>
<td>1.10</td>
<td>1.11</td>
<td>1.16</td>
</tr>
<tr>
<td>2. Factor</td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td>3. Factor</td>
<td>0.40</td>
<td>0.31</td>
<td>0.31</td>
<td>0.49</td>
</tr>
<tr>
<td>4. Factor</td>
<td><strong>0.40</strong></td>
<td><strong>0.08</strong></td>
<td><strong>0.08</strong></td>
<td><strong>0.21</strong></td>
</tr>
</tbody>
</table>

**Table**: Estimated mean of factors. $\gamma = 50$. 
## Simulation

<table>
<thead>
<tr>
<th></th>
<th>True</th>
<th>PCA Var</th>
<th>PCA Corr</th>
<th>RP-PCA</th>
<th>RP-PCA Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Factor</td>
<td>9.000</td>
<td>8.608</td>
<td>8.615</td>
<td>8.494</td>
<td>8.510</td>
</tr>
<tr>
<td>2. Factor</td>
<td>1.000</td>
<td>0.697</td>
<td>0.716</td>
<td>0.683</td>
<td>0.706</td>
</tr>
<tr>
<td>3. Factor</td>
<td>1.000</td>
<td>0.801</td>
<td>0.820</td>
<td>0.674</td>
<td>0.690</td>
</tr>
<tr>
<td>4. Factor</td>
<td>0.160</td>
<td>0.028</td>
<td>0.028</td>
<td>0.066</td>
<td>0.070</td>
</tr>
</tbody>
</table>

**Table:** Estimated variance of factors. $\gamma = 50$. 
## Fama-MacBeth Test-Statistics: $\chi^2_{22}: 34(95\%)$  

<table>
<thead>
<tr>
<th>Factor Pairs</th>
<th>RP-PCA</th>
<th>PCA</th>
<th>FF 5</th>
<th>Specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size and BM</td>
<td>85.66</td>
<td>94.50</td>
<td>79.99</td>
<td>105.15</td>
</tr>
<tr>
<td>BM and Investment</td>
<td>14.52</td>
<td>37.04</td>
<td>26.14</td>
<td>31.61</td>
</tr>
<tr>
<td>BM and Operating Profits</td>
<td>19.45</td>
<td>25.95</td>
<td>15.40</td>
<td>21.92</td>
</tr>
<tr>
<td>Size and Accrual</td>
<td>44.57</td>
<td>89.95</td>
<td>61.98</td>
<td>76.04</td>
</tr>
<tr>
<td>Size and Beta</td>
<td>30.74</td>
<td>32.90</td>
<td>31.76</td>
<td>31.96</td>
</tr>
<tr>
<td>Size and Investment</td>
<td>87.89</td>
<td>104.53</td>
<td>93.88</td>
<td>103.60</td>
</tr>
<tr>
<td>Size and Operating Profits</td>
<td>29.17</td>
<td>32.98</td>
<td>29.16</td>
<td>42.32</td>
</tr>
<tr>
<td>Size and Short-Term Reversal</td>
<td>87.70</td>
<td>103.35</td>
<td>88.86</td>
<td>108.31</td>
</tr>
<tr>
<td>Size and Long-Term Reversal</td>
<td>53.92</td>
<td>65.07</td>
<td>44.09</td>
<td>68.69</td>
</tr>
<tr>
<td>Size and Res. Var.</td>
<td>134.57</td>
<td>147.18</td>
<td>125.28</td>
<td>163.77</td>
</tr>
<tr>
<td>Size and Total Var.</td>
<td>120.14</td>
<td>133.46</td>
<td>120.71</td>
<td>143.01</td>
</tr>
<tr>
<td>Operating Profits and Investment</td>
<td>29.21</td>
<td>51.63</td>
<td>34.38</td>
<td>35.89</td>
</tr>
<tr>
<td>Size and Net Share Iss.</td>
<td>121.13</td>
<td>149.78</td>
<td>119.91</td>
<td>126.64</td>
</tr>
<tr>
<td>49 Industries</td>
<td>140.76</td>
<td>175.77</td>
<td>140.59</td>
<td>206.47</td>
</tr>
</tbody>
</table>
Predicted excess return in-sample

![Graphs showing predicted and expected excess returns for different models and factors.](image)
Predicted excess return out-of-sample

- OLS out-of-sample RP-PCA
- OLS out-of-sample PCA
- OLS out-of-sample 5 Fama-French factors
- OLS out-of-sample Specific factors
Cumulative returns of optimal portfolios

RP-PCA Optimal Portfolios

PCA Optimal Portfolios
Portfolio Data: In-sample (BM and Investment)

<table>
<thead>
<tr>
<th></th>
<th>SR</th>
<th>RMS $\alpha$</th>
<th>Fama-MacBeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.256</td>
<td>0.074</td>
<td>14.520</td>
</tr>
<tr>
<td>PCA</td>
<td>0.169</td>
<td>0.123</td>
<td>37.038</td>
</tr>
<tr>
<td>Fama-French</td>
<td>0.344</td>
<td>0.140</td>
<td>26.144</td>
</tr>
<tr>
<td>Specific</td>
<td>0.236</td>
<td>0.127</td>
<td>31.611</td>
</tr>
</tbody>
</table>

Table: Maximal Sharpe-ratios, root-mean-squared pricing errors and Fama-MacBeth test statistics for different set of factors. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Cross-sectional $\alpha$’s for sorted portfolios (BM and Investment)
Loadings for statistical factors (BM and Investment)

- **Loadings of 1. PCA factor**
  - Portfolio
  - Loadings
  - Loadings range from 0.15 to 0.25

- **Loadings of 2. PCA factor**
  - Portfolio
  - Loadings
  - Loadings range from -0.4 to 0.2

- **Loadings of 3. PCA factor**
  - Portfolio
  - Loadings
  - Loadings range from -0.6 to 0.4

- **Loadings of 1. RP-PCA factor**
  - Portfolio
  - Loadings
  - Loadings range from 0.15 to 0.3

- **Loadings of 2. RP-PCA factor**
  - Portfolio
  - Loadings
  - Loadings range from -0.4 to 0.2

- **Loadings of 3. RP-PCA factor**
  - Portfolio
  - Loadings
  - Loadings range from -0.6 to 0.4
Maximal Incremental Sharpe Ratio (BM and Investment)

<table>
<thead>
<tr>
<th>Factors</th>
<th>PCA</th>
<th>RP-PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>0.144</td>
<td>0.149</td>
</tr>
<tr>
<td>2 Factors</td>
<td>0.167</td>
<td>0.193</td>
</tr>
<tr>
<td>3 Factors</td>
<td>0.169</td>
<td>0.256</td>
</tr>
</tbody>
</table>

**Table:** Maximal Sharpe-ratio by adding factors incrementally. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
### Portfolio Data: Objective function (BM and Investment)

<table>
<thead>
<tr>
<th></th>
<th>PCA TS</th>
<th>RP-PCA TS</th>
<th>PCA XS</th>
<th>RP-PCA XS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>5.543</td>
<td>5.989</td>
<td>0.021</td>
<td>0.002</td>
</tr>
<tr>
<td>2 Factors</td>
<td>4.416</td>
<td>4.647</td>
<td>0.014</td>
<td>0.001</td>
</tr>
<tr>
<td>3 Factors</td>
<td>3.944</td>
<td>4.098</td>
<td>0.013</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Table**: Time-series and cross-sectional objective functions.
Portfolio Data: Out-of-sample (BM and Investment)

<table>
<thead>
<tr>
<th></th>
<th>Out-of-sample</th>
<th>In-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.123</td>
<td>0.065</td>
</tr>
<tr>
<td>PCA</td>
<td>0.157</td>
<td>0.156</td>
</tr>
<tr>
<td>Fama-French 5</td>
<td>0.111</td>
<td>0.103</td>
</tr>
<tr>
<td>Specific</td>
<td>0.138</td>
<td>0.138</td>
</tr>
</tbody>
</table>

**Table:** Root-mean-squared pricing errors for different set of factors. Out-of-sample factors are estimated with a rolling window. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Cross-sectional $\alpha$’s out-of-sample (BM and Investment)
Predicted excess return in-sample (BM and Investment)
Predicted excess return out-of-sample (BM and Invest.)

- OLS out-of-sample RP-PCA
- OLS out-of-sample PCA
- OLS out-of-sample 5 Fama-French factors
- OLS out-of-sample Specific factors
Portfolio Data: In-sample (Size and BM)

<table>
<thead>
<tr>
<th></th>
<th>SR</th>
<th>RMS α</th>
<th>Fama-MacBeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.248</td>
<td>0.126</td>
<td>85.664</td>
</tr>
<tr>
<td>PCA</td>
<td>0.217</td>
<td>0.137</td>
<td>94.505</td>
</tr>
<tr>
<td>Fama-French</td>
<td>0.344</td>
<td>0.116</td>
<td>79.990</td>
</tr>
<tr>
<td>Specific</td>
<td>0.163</td>
<td>0.197</td>
<td>105.153</td>
</tr>
</tbody>
</table>

**Table**: Maximal Sharpe-ratios, root-mean-squared pricing errors and Fama-MacBeth test statistics for different set of factors. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Cross-sectional $\alpha$'s for sorted portfolios (Size and BM)
Loadings for statistical factors (Size and BM)
Maximal Incremental Sharpe Ratio

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>RP-PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>0.148</td>
<td>0.156</td>
</tr>
<tr>
<td>2 Factors</td>
<td>0.155</td>
<td>0.212</td>
</tr>
<tr>
<td>3 Factors</td>
<td>0.217</td>
<td>0.248</td>
</tr>
</tbody>
</table>

Table: Maximal Sharpe-ratio by adding factors incrementally. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
### Portfolio Data: Objective function (Size and BM)

<table>
<thead>
<tr>
<th></th>
<th>PCA TS</th>
<th>RP-PCA TS</th>
<th>PCA XS</th>
<th>RP-PCA XS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>4.263</td>
<td>4.981</td>
<td>0.035</td>
<td>0.003</td>
</tr>
<tr>
<td>2 Factors</td>
<td>2.663</td>
<td>3.213</td>
<td>0.032</td>
<td>0.001</td>
</tr>
<tr>
<td>3 Factors</td>
<td>1.756</td>
<td>1.889</td>
<td>0.011</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Table:** Time-series and cross-sectional objective functions.
### Portfolio Data: Out-of-sample (Size and BM)

<table>
<thead>
<tr>
<th></th>
<th>Out-of-sample</th>
<th>In-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.171</td>
<td>0.160</td>
</tr>
<tr>
<td>PCA</td>
<td>0.187</td>
<td>0.180</td>
</tr>
<tr>
<td>Fama-French 5</td>
<td>0.141</td>
<td>0.140</td>
</tr>
<tr>
<td>Specific</td>
<td>0.212</td>
<td>0.196</td>
</tr>
</tbody>
</table>

**Table:** Root-mean-squared pricing errors for different set of factors. Out-of-sample factors are estimated with a rolling window. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Cross-sectional $\alpha$’s out-of-sample (Size and BM)
Generalized correlations for time-varying loadings (Size and BM)
Generalized correlations for time-varying loadings (Size and BM)
Time-varying loadings (Size and BM)
Predicted excess return in-sample (Size and BM)

In-sample RP-PCA

In-sample PCA

In-sample 5 Fama-French factors

In-sample Specific factors
Predicted excess return out-of-sample (Size and BM)

- **OLS out-of-sample RP-PCA**
- **OLS out-of-sample PCA**
- **OLS out-of-sample 5 Fama-French factors**
- **OLS out-of-sample Specific factors**
## Portfolio Data: In-sample (Size and Momentum)

<table>
<thead>
<tr>
<th></th>
<th>SR</th>
<th>RMS</th>
<th>α</th>
<th>Fama-MacBeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.255</td>
<td>0.146</td>
<td>87.702</td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>0.199</td>
<td>0.160</td>
<td>103.350</td>
<td></td>
</tr>
<tr>
<td>Fama-French</td>
<td>0.344</td>
<td>0.238</td>
<td>88.855</td>
<td></td>
</tr>
<tr>
<td>Specific</td>
<td>0.253</td>
<td>0.329</td>
<td>108.315</td>
<td></td>
</tr>
</tbody>
</table>

**Table:** Maximal Sharpe-ratios, root-mean-squared pricing errors and Fama-MacBeth test statistics for different set of factors. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Cross-sectional $\alpha$’s for sorted portfolios (Size and Momentum)

![Graph showing cross-sectional $\alpha$’s for sorted portfolios (Size and Momentum)](image)

- PCA
- RP-PCA
- Fama-French 5
- Specific

Pricing Errors Size and Short-Term Reversal

A 33
Loadings for statistical factors (Size and Momentum)
**Portfolio Data: Out-of-sample (Size and Momentum)**

<table>
<thead>
<tr>
<th></th>
<th>Out-of-sample</th>
<th>In-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.171</td>
<td>0.148</td>
</tr>
<tr>
<td>PCA</td>
<td>0.193</td>
<td>0.187</td>
</tr>
<tr>
<td>Fama-French 5</td>
<td>0.090</td>
<td>0.106</td>
</tr>
<tr>
<td>Specific</td>
<td>0.181</td>
<td>0.201</td>
</tr>
</tbody>
</table>

**Table:** Root-mean-squared pricing errors for different set of factors. Out-of-sample factors are estimated with a rolling window. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Cross-sectional $\alpha$’s out-of-sample (Size and Momentum)
Predicted excess return in-sample (Size and Momentum)

In-sample RP-PCA

In-sample PCA

In-sample 5 Fama-French factors

In-sample Specific factors
Predicted excess return out-of-sample (Size and Moment.)

- OLS out-of-sample RP-PCA
- OLS out-of-sample PCA
- OLS out-of-sample 5 Fama-French factors
- OLS out-of-sample Specific factors
## Portfolio Data: Objective function (Size and Moment.)

<table>
<thead>
<tr>
<th></th>
<th>PCA TS</th>
<th>RP-PCA TS</th>
<th>PCA XS</th>
<th>RP-PCA XS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>3.850</td>
<td>4.717</td>
<td>0.038</td>
<td>0.004</td>
</tr>
<tr>
<td>2 Factors</td>
<td>2.618</td>
<td>3.099</td>
<td>0.038</td>
<td>0.000</td>
</tr>
<tr>
<td>3 Factors</td>
<td>1.674</td>
<td>1.872</td>
<td>0.017</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Table:** Time-series and cross-sectional objective functions.
Maximal Incremental Sharpe Ratio (Size and Moment.)

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>RP-PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>0.129</td>
<td>0.138</td>
</tr>
<tr>
<td>2 Factors</td>
<td>0.130</td>
<td>0.255</td>
</tr>
<tr>
<td>3 Factors</td>
<td>0.199</td>
<td>0.255</td>
</tr>
</tbody>
</table>

Table: Maximal Sharpe-ratio by adding factors incrementally. For the statistical factor estimators we use \( K = 3 \) factors and \( \gamma = 100 \).
Portfolio Data: In-sample (Size and Net Share Iss.)

<table>
<thead>
<tr>
<th></th>
<th>SR</th>
<th>RMS $\alpha$</th>
<th>Fama-MacBeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.328</td>
<td>0.142</td>
<td>121.135</td>
</tr>
<tr>
<td>PCA</td>
<td>0.246</td>
<td>0.163</td>
<td>149.778</td>
</tr>
<tr>
<td>Fama-French</td>
<td>0.344</td>
<td>0.129</td>
<td>119.912</td>
</tr>
<tr>
<td>Specific</td>
<td>0.349</td>
<td>0.167</td>
<td>126.635</td>
</tr>
</tbody>
</table>

**Table:** Maximal Sharpe-ratios, root-mean-squared pricing errors and Fama-MacBeth test statistics for different set of factors. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Cross-sectional $\alpha$’s for sorted portfolios (Size and Net Share Iss.)
Loadings for statistical factors (Size and Net Share Iss.)

- **Loadings of 1. PCA factor**
- **Loadings of 2. PCA factor**
- **Loadings of 3. PCA factor**
- **Loadings of 1. RP-PCA factor**
- **Loadings of 2. RP-PCA factor**
- **Loadings of 3. RP-PCA factor**
Cross-sectional $\alpha$’s out-of-sample (Size and Net Share Iss.)
Portfolio Data: Out-of-sample (Size and Net Share Iss.)

<table>
<thead>
<tr>
<th></th>
<th>Out-of-sample</th>
<th>In-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>RP-PCA</td>
<td>0.142</td>
<td>0.151</td>
</tr>
<tr>
<td>PCA</td>
<td>0.206</td>
<td>0.204</td>
</tr>
<tr>
<td>Fama-French 5</td>
<td>0.164</td>
<td>0.152</td>
</tr>
<tr>
<td>Specific</td>
<td>0.183</td>
<td>0.156</td>
</tr>
</tbody>
</table>

**Table:** Root-mean-squared pricing errors for different set of factors. Out-of-sample factors are estimated with a rolling window. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 
Predicted excess return in-sample (Size and Shares)

- **In-sample RP-PCA**
- **In-sample PCA**
- **In-sample 5 Fama-French factors**
- **In-sample Specific factors**
Predicted excess return out-of-sample (Size and Shares)

- **OLS out-of-sample RP-PCA**
- **OLS out-of-sample PCA**
- **OLS out-of-sample 5 Fama-French factors**
- **OLS out-of-sample Specific factors**
Portfolio Data: Objective function (Size and Shares)

<table>
<thead>
<tr>
<th></th>
<th>PCA TS</th>
<th>PCA XS</th>
<th>RP-PCA TS</th>
<th>RP-PCA XS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>6.421</td>
<td>7.707</td>
<td>0.060</td>
<td>0.006</td>
</tr>
<tr>
<td>2 Factors</td>
<td>4.502</td>
<td>5.729</td>
<td>0.060</td>
<td>0.000</td>
</tr>
<tr>
<td>3 Factors</td>
<td>3.628</td>
<td>3.813</td>
<td>0.025</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Table:** Time-series and cross-sectional objective functions.
Maximal Incremental Sharpe Ratio (Size and Shares)

<table>
<thead>
<tr>
<th></th>
<th>PCA</th>
<th>RP-PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Factor</td>
<td>0.138</td>
<td>0.148</td>
</tr>
<tr>
<td>2 Factors</td>
<td>0.138</td>
<td>0.327</td>
</tr>
<tr>
<td>3 Factors</td>
<td>0.246</td>
<td>0.328</td>
</tr>
</tbody>
</table>

**Table:** Maximal Sharpe-ratio by adding factors incrementally. For the statistical factor estimators we use $K = 3$ factors and $\gamma = 100$. 

**PCA RP-PCA**

1 Factor 0.138 0.148
2 Factors 0.138 0.327
3 Factors 0.246 0.328