Shrinking Factor Dimension: A Reduced-Rank Approach

Dashan Huang
Singapore Management University

Jiaen Li    Guofu Zhou
Washington University in St. Louis

25 July 2019

Thirteenth Annual Risk Management Conference
Cochrane (2011): “a zoo of new variables”

- There seem to have hundreds of variables that can predict stock returns:
  5. ...

- New factors models have been developed in the past 5 years:
  1. Fama and French (FF, 2015): MKT, SMB, HML, RMW, CMA
  5. FF (2018): MKT, SMB, HML, RMW, CMA, MOM
  6. ...

Cochrane (2011) concludes:

“the world is once again descending into chaos.”
1. How many factors do we really need based on existing 300+?

2. Given the FF five factors, do the other factors provide incremental information?
The Literature vs. This Paper

- Kozak, Nagel, and Santosh (2018): at least 10
- Gu, Kelly, and Xiu (2018): at least 20
- Feng, Giglio, and Xiu (2017): 14
- Kelly, Pruitt, and Su (2018): 8
- DeMiguel, Martin-Utrera, Nogales, and Uppal (2018): 15

This paper proposes a reduced-rank approach, which is analytically solvable in the GMM framework of Hansen (1982).

Two main findings:

1. For practical purposes, a five-factor model works well in terms of pricing errors.
2. Extant factors do not provide much incremental information beyond the FF five factors.
Reduced-Rank Approach: RRA

- Suppose asset returns are governed by a $K$-factor model:

$$R_{it} = \alpha_i + \beta_{i1}f_{1t} + \cdots + \beta_{iK}f_{Kt} + u_{it}, \ i = 1, \cdots, N,$$

where, $f_1, \cdots, f_K$ are latent variables but are related to $L$ factor proxies:

$$f_{kt} = \phi_{k1}g_{1t} + \cdots + \phi_{kL}g_{Lt}, \ k = 1, \cdots, K.$$  

**e.g.,** $K = 5, \ L = 70$

- In terms of matrix notation:

$$R_t = \alpha + \beta'f_t + U_t,$$  

$$f_t = \Phi'g_t.$$  

- Then the parameters are $\alpha : N \times 1$, $\Phi : L \times K$ and $\beta : K \times N$. 

\[5/19\]
To estimate the parameters, we turn to the GMM of Hansen (1982). Let $U_t = R_t - \alpha - \beta'f_t$ and $Z_t = (1, g_t')'$, where $g_t$ is the $L$ factor candidates at $t$.

The moment condition is

$$E[h_t(\alpha, \Phi, \beta)] = 0, \quad h_t(\alpha, \Phi, \beta) \equiv U_t(\alpha, \Phi, \beta) \otimes Z_t.$$ 

The GMM estimator solves

$$\min_{\alpha, \Phi, \beta} Q \equiv h_T' W_T h_T,$$

(5)

where $W_T$ is the weighting matrix and $h_T$ is the sample mean of $h_t$:

$$h_T = \frac{1}{T} \sum_{t=1}^{T} h_t(\alpha, \Phi, \beta).$$
If the weighting matrix follows

$$W_T \equiv W_1 \otimes W_2,$$

(6)

\(\alpha, \Phi, \) and \(\beta\) can be solved analytically and \(\hat{\Phi}\) is given as:

$$\hat{\Phi} = \left( G'PG/T^2 \right)^{-1/2} E,$$

(7)

where \(P\) and \(E\) are analytical matrices.

Then, the RRA factor is

$$\hat{f}_t = \hat{\Phi}' g_t.$$

(8)

In our paper, we assume

$$W_T^{-1} = \hat{S}_T = \frac{1}{T} \sum \hat{U}_t \hat{U}_t' \otimes \frac{1}{T} \sum Z_t Z_t'$$

(9)

and use its diagonal matrix.
Alternative Dimension Reduction Approaches

1 Principal component analysis (PCA): \( f_t^{\text{PCA}} = (\Phi^{\text{PCA}})'g_t \)

\[
\max_{\phi_k^{\text{PCA}}} \text{Var}(g_t'\phi_k^{\text{PCA}}),
\]

such that the \( k \)-th component \((\phi_k^{\text{PCA}})'g_t\) is independent from the former.

2 Partial least squares (PLS): \( f_t^{\text{PLS}} = (\Phi^{\text{PLS}})'g_t \)

\[
\max_{\psi_k^{\text{PLS}}, \phi_k^{\text{PLS}}} \text{Cov}(R_t'\psi_k^{\text{PLS}}, g_t'\phi_k^{\text{PLS}}).
\]

**RRA vs. PCA vs. PLS:**

- **PCA** does not consider the statistical objective: explaining returns.
- **PLS** focuses on the covariance, rather than the average returns.
Performance Measures

1 Root-mean-squared (RMS) alpha:

\[\text{RMS } \alpha = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \hat{\alpha}_i^2}\]

2 Total adj-\(R^2\) (Kelly, Pruitt, and Su, 2018):

\[\text{Total adj-}R^2 = 1 - \frac{\sum_{i,t}(R_{it} - \hat{\alpha}_i - \hat{\beta}_{i1}\hat{f}_{1t} - \cdots - \hat{\beta}_{iK}\hat{f}_{Kt})^2}{\sum_{i,t} R_{it}^2} \times \frac{T_i-1}{T_i-K-1}\]
Data

- **Target assets**: proxy for the cross section of stock returns
  1. Fama-French 48 industry portfolios
  2. 202 characteristic portfolios (Giglio and Xiu, 2018):
     - 25 by size and b/m, 17 industry, 24 by profitability and investment, 25 by size and variance, 35 by size and net issuance, 25 by size and accruals, 25 by size and beta, and 25 by size and momentum

- **70 factor candidates**
  1. FF five factors, momentum, Pastor-Stambaugh liquidity, Hou-Xue-Zhang ROE
  2. 62 CAPM anomalies (value-weighted decile spreads)

- **Testing assets**
  1. 48 industry portfolios
  2. 202 characteristic portfolios
  3. all individual stocks
  4. all-but-micro stocks

- **Sample period**: 1974:01–2016:12
Both Target and Testing Assets are 48 Industry Portfolios

<table>
<thead>
<tr>
<th></th>
<th>1 factor</th>
<th>3 factors</th>
<th>5 factors</th>
<th>6 factors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Root-mean-squared alpha (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>0.24</td>
<td>0.29</td>
<td>0.34</td>
<td>0.31</td>
</tr>
<tr>
<td>PCA</td>
<td>1.00</td>
<td>1.06</td>
<td>0.88</td>
<td>0.99</td>
</tr>
<tr>
<td>PLS</td>
<td>1.03</td>
<td>0.54</td>
<td>0.33</td>
<td>0.24</td>
</tr>
<tr>
<td>RRA</td>
<td>0.23</td>
<td>0.20</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Total adj-$R^2$ (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>51.39</td>
<td>55.57</td>
<td>57.77</td>
<td>58.34</td>
</tr>
<tr>
<td>PCA</td>
<td>16.74</td>
<td>20.49</td>
<td>29.92</td>
<td>33.14</td>
</tr>
<tr>
<td>PLS</td>
<td>23.42</td>
<td>47.19</td>
<td>58.97</td>
<td>61.11</td>
</tr>
<tr>
<td>RRA</td>
<td>54.60</td>
<td>61.11</td>
<td>64.75</td>
<td>65.38</td>
</tr>
<tr>
<td></td>
<td>1 factor</td>
<td>3 factors</td>
<td>5 factors</td>
<td>6 factors</td>
</tr>
<tr>
<td>------------------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td><strong>Root-mean-squared alpha (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>2.53</td>
<td>2.61</td>
<td>3.08</td>
<td>3.19</td>
</tr>
<tr>
<td>PCA</td>
<td>2.93</td>
<td>3.10</td>
<td>3.15</td>
<td>3.18</td>
</tr>
<tr>
<td>PLS</td>
<td>2.93</td>
<td>2.92</td>
<td>2.98</td>
<td>3.05</td>
</tr>
<tr>
<td>RRA</td>
<td>2.50</td>
<td>2.59</td>
<td>2.65</td>
<td>2.82</td>
</tr>
<tr>
<td><strong>Total adj-$R^2$ (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>9.37</td>
<td>13.64</td>
<td>14.70</td>
<td>15.55</td>
</tr>
<tr>
<td>PCA</td>
<td>8.88</td>
<td>10.65</td>
<td>12.39</td>
<td>12.82</td>
</tr>
<tr>
<td>PLS</td>
<td>10.02</td>
<td>12.97</td>
<td>14.35</td>
<td>15.10</td>
</tr>
<tr>
<td>RRA</td>
<td>9.89</td>
<td>12.83</td>
<td>14.43</td>
<td>15.19</td>
</tr>
</tbody>
</table>
Optimal Number of Factors

- Statistically, a model with up to 10 factors is rejected by the GMM.
  —more than 10 factors are needed to fit the moment conditions.

- Economically, however,
  1. The MKT is the most important factor (Harvey and Liu, 2018).
  2. When the testing assets are the target, the RRA outperforms the FF.
  3. When the testing assets are different, the RRA performs similarly as the FF, and both of them slightly perform better than the PCA and PLS.
  4. When increasing the number of factors from 5 to 6 to 10, the pricing power does not increase very much.

It appears that a five-factor model, like the FF5, works well for practical purposes.
Do The Other Factors Provide Incremental Information Relative to the FF Five Factors?

Suppose there are $K$ more factors, in addition to the FF five factors

$$R_{it} = \alpha_i + \beta'_0 FF5_t + \beta_i f_{1t} + \cdots + \beta_iK f_{Kt} + \epsilon_{it},$$

where $f_1, \cdots, f_K$ are latent factors and related to

$$f_{kt} = \phi_{k1} g_{1t} + \cdots + \phi_{kL} g_{Lt}, \quad k = 1, \cdots, K.$$

The GMM estimator can be extended to this case to obtain additional factors from existing proxies.
Both Target and Testing Assets are 48 Industry Portfolios

<table>
<thead>
<tr>
<th></th>
<th>5 factor</th>
<th>6 factors</th>
<th>7 factors</th>
<th>8 factors</th>
<th>10 factors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Root-mean-squared alpha (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>0.34</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PCA</td>
<td>–</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.28</td>
</tr>
<tr>
<td>PLS</td>
<td>–</td>
<td>0.34</td>
<td>0.30</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>RRA</td>
<td>–</td>
<td>0.34</td>
<td>0.30</td>
<td>0.21</td>
<td>0.21</td>
</tr>
<tr>
<td><strong>Total adj-$R^2$ (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>57.57</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>PCA</td>
<td>–</td>
<td>58.80</td>
<td>59.23</td>
<td>61.00</td>
<td>61.85</td>
</tr>
<tr>
<td>PLS</td>
<td>–</td>
<td>60.81</td>
<td>61.60</td>
<td>62.38</td>
<td>63.79</td>
</tr>
<tr>
<td>RRA</td>
<td>–</td>
<td>61.33</td>
<td>63.14</td>
<td>64.65</td>
<td>65.80</td>
</tr>
</tbody>
</table>
Target Assets Are 48 Industry Portfolios and Testing Assets Are All Stocks

<table>
<thead>
<tr>
<th></th>
<th>RMS alpha (%)</th>
<th></th>
<th>Total adj-$R^2$ (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5 factors</td>
<td>6 factors</td>
<td>10 factors</td>
<td></td>
</tr>
<tr>
<td>FF</td>
<td>3.08</td>
<td>–</td>
<td>–</td>
<td>14.70</td>
</tr>
<tr>
<td>FF5+PCA</td>
<td>–</td>
<td>3.22</td>
<td>3.61</td>
<td>–</td>
</tr>
<tr>
<td>FF5+PLS</td>
<td>–</td>
<td>3.18</td>
<td>3.63</td>
<td>–</td>
</tr>
<tr>
<td>FF5+RRA</td>
<td>–</td>
<td>3.18</td>
<td>3.41</td>
<td>–</td>
</tr>
</tbody>
</table>

Additional factors improve little beyond the FF5, implying that the other proxies do not contain much new information.
Robustness 1: Does A Mispricing Restriction Matter?

- In our analysis, there is no restriction on $\alpha_i$:

$$R_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + u_{it}.$$ 

- What if a degree of mispricing is allowed?

$$\alpha_i = \eta \sigma_i,$$  

  e.g., $\eta = 1/12$.

- Target assets: 48 industry portfolios; testing assets: all stocks

<table>
<thead>
<tr>
<th></th>
<th>1 factor</th>
<th>3 factors</th>
<th>5 factors</th>
<th>6 factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>no constraint</td>
<td>0.23</td>
<td>0.20</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>$\alpha_i = 0$</td>
<td>0.23</td>
<td>0.19</td>
<td>0.16</td>
<td>0.16</td>
</tr>
<tr>
<td>$\alpha_i = \frac{\sigma_i}{12} (\approx 0.5%)$</td>
<td>0.44</td>
<td>0.41</td>
<td>0.39</td>
<td>0.38</td>
</tr>
</tbody>
</table>
Robustness 2: Can We Extract Factors From the Target Directly?

Suppose the target assets are governed by

$$R_{it} = \alpha_i + b_{i1} f_{i1}^e + \cdots + b_{iJ} f_{iJ}^e + e_{it}, \quad i = 1, \cdots, N,$$

(11)

where each factor $j_{i,t}^e$ follows:

$$f_{jt}^e = c_{j1} R_{1t} + \cdots + c_{jN} R_{Nt}, \quad j = 1, \cdots, J.$$

(12)

Based on Bai (2003) and Balvers and Stivers (2018),

$$\hat{f}_t^e = \hat{C}' R_t, \quad i.e., \text{target-based factors.}$$

(13)

A more general case is:

$$R_t = \alpha + B\hat{f}_t^e + \beta' f_t + \varepsilon_t, \quad f_t \text{ are the RRA factors.}$$

(14)

The results show that the performance is no better than the FF and RRA models.
Conclusion

1. Statistically, the number of factors we need may exceed 10; but for practical use, a five-factor model, say the FF5, works well.

2. Given the FF five factors, the other factors do not provide much incremental information.