Imperial College
London

5 May 2023

Pelger's Factor Model

Wasim Ahmed
Matteo Di Venti
Toshiko Matsui

Previous Factor Models

Three common approaches to identify the factors that describe the *systematic risk* - factors are based either on:

- Theory and economic intuition
  e.g.) CAPM (market is the only common factor)
  $$E(R_i) = \alpha_i + \beta_i(E(R_M) - R_f)$$
- Firm characteristics
  e.g.) Fama-French 3 Factor (5 Factor) model
  $$ri - rf = \beta_{MKT,i}(r_M - r_f) + \beta_{SMB,i}SMB + \beta_{HML,i}HML + \beta_{RMW,i}RMW + \beta_{CMA,i}CMA + \epsilon_i$$
- Statistical approach - This paper's focus

*SMB (small minus big: size of firms), HML (high minus low: book-to-market values), RMW (robust minus weak: profitability), CMA (conservative minus aggressive: investment)

Motivation behind Pelger's Model

Pelger analysis presents three major innovations:

- The use of statistical methods to determine factors rather than a priori assumptions (and potentially wrong) assumptions on the market
- The use of high-frequency data that allows to analyse very short time periods independently.
- Separates high frequency-returns in continuous intraday, intraday and overnight jumps which allows to better understand systematic risk.

A surprising result is that the portfolio weights are stable over time and that the estimated factors seem to be stable when considering different time scales.

Factor estimation

General method outlined in his paper, Pelger(2019), PCA applied to volatility and jump covariance matrix.

$$R = \Delta F \Lambda^T + \Delta e$$

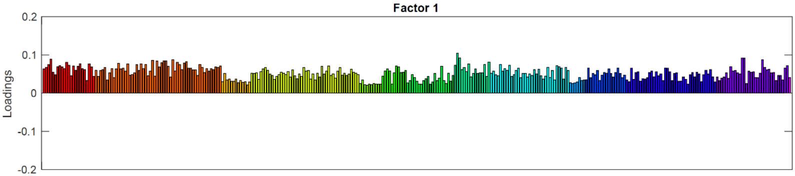
where R is the panel HF log return matrix $\Delta F, (M \times K)$ and $\Lambda^T, (K \times N)$ represent the loadings and the increments of the factors.

1. Estimate the loading matrix through PCA as the eigenvectors of the K largest eigenvalues of $\frac{R^T R}{N}$ multiplied by $\sqrt{N}$
2. The estimated loadings measure exposure to risk factor and build the continuous portfolio weights $\omega^C = \frac{\hat{\Lambda}^C}{\sqrt{N}}$
3. Finally the continuous factors are given by $R^{time}\omega^C$

Imperial College London

Proxy Factor 1

We wish to find general economic factors that can replicate these statistical factors.

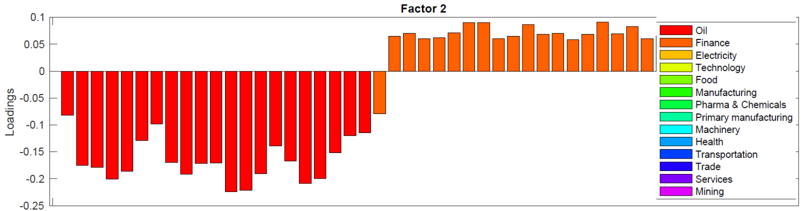


- These are the portfolio weights for all of the stocks in the balanced portfolio
- There is no clear industry this factor corresponds to, so Pelger chooses an equally weighted first factor

Imperial College London

Proxy Factor 2

Here are stocks (sorted by industry) for the largest 15% of portfolio weights corresponding to 2nd PCA factor.

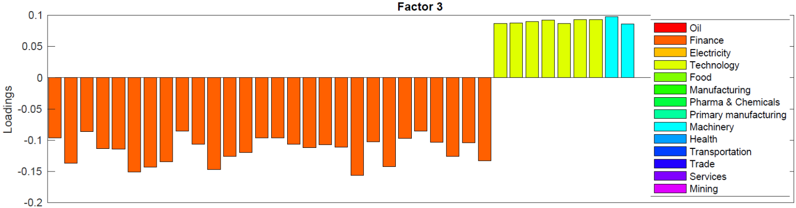


- The largest weights are in the oil and finance industries, with significantly larger weights in the oil industry
- So, the second proxy factor is an equally weighted oil factor

Imperial College
London

Proxy Factor 3

Similarly, here are the stocks for the largest 15% of portfolio weights corresponding to 3rd PCA factor.

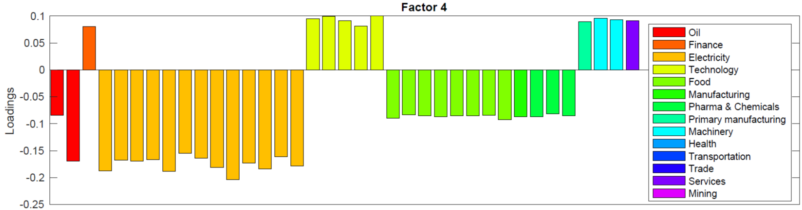


- Clearly, the largest weights are in the finance industry
- An equally weighted finance factor is chosen for the third proxy factor

Imperial College London

Proxy Factor 4

The fourth factor is far less clear, and so we look at stocks corresponding to just the top 11% of weights



- Although there are more industries represented, as electricity has the largest weighs, it is chosen for the 4th proxy factor
- As the generalised correlations are $\{1, 0.99, 0.95, 0.91\}$, these proxy factors do a good job of representing the PCA factors