Toward a Reduced-Form Factor Portfolio

TTIC 31220 Final Project Presentation

Kabir Sawhney and Joseph Denby

March 13, 2019

Motivation & Background

$$r_i = r_f + \beta_1(r_m - r_f) + \beta_2(SMB) + \beta_3(HML) + \varepsilon$$

- Different financial assets (stocks, bonds, etc.) are all influenced by a similar set of underlying risks (inflation, economic growth, etc.).
- In finance theory, the returns on these assets can be modeled as a linear combination of the returns to these risks, plus an error term.
- Some researchers have used linear PCA to derive a smaller set of risks that perform well in explaining equity returns. We sought to extend this work to non-linear dimensionality reduction.

Description of Data

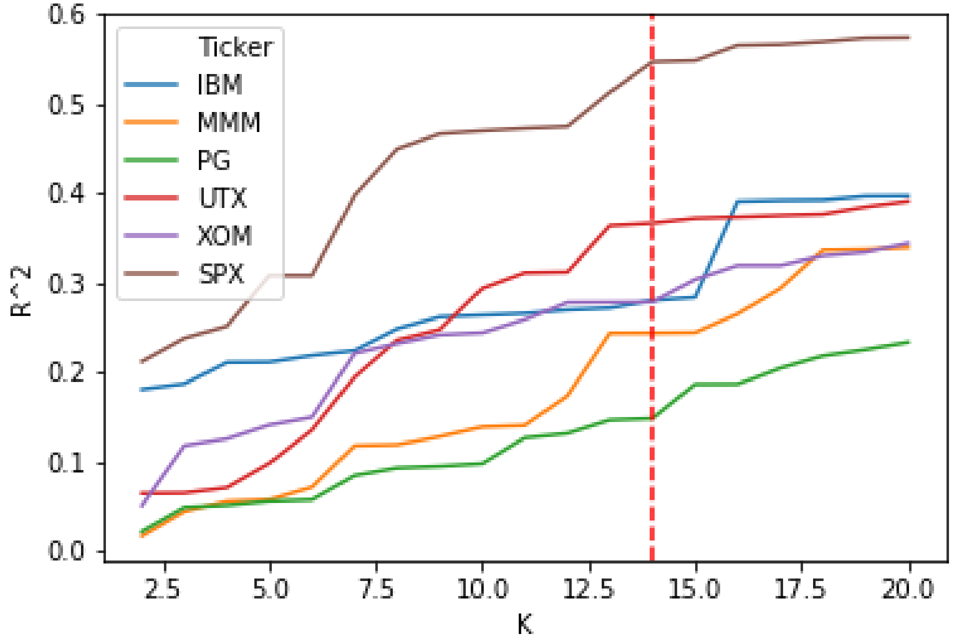
- Our data was monthly returns from 1963 to 2013.
- For our response variable, stock returns, we used the S&P 500 and the five longest-tenured stocks in the Dow Jones Industrial Average (ExxonMobil, Procter & Gamble, United Technologies, 3M, and IBM).
- For the risks, we used a set of factor strategies developed by Novy-Marx and Velikov. They collected the returns to rules-based strategies designed to provide exposure to different factors.
- We compared our results with the most popular theoretically-constructed risk factor portfolio, the Fama-French three-factor model.

Methodology

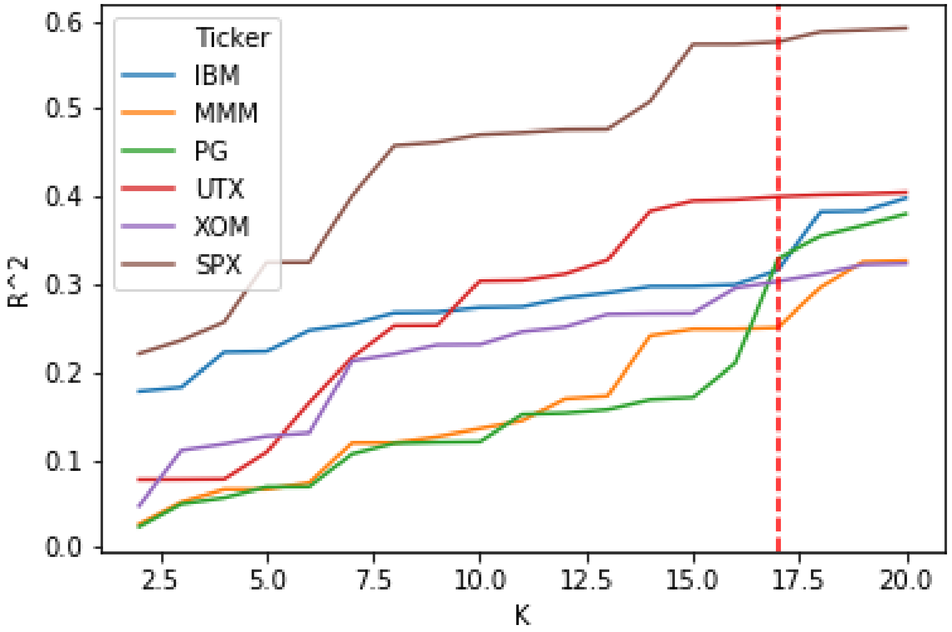
- Split the data into training, validation, and test sets in time.
- Tune hyperparameters for linear PCA, kernel PCA (with the polynomial and RBF kernels), Isomap, and Laplacian eigenmap.
- Choose the best-performing non-linear dimensionality reduction methods.
- Use the learned representation to fit a linear regression model.
- Compare performance on the test set with linear PCA and the Fama-French model.

Results: Method Selection & Tuning

Tuning Number of Components for Linear PCA

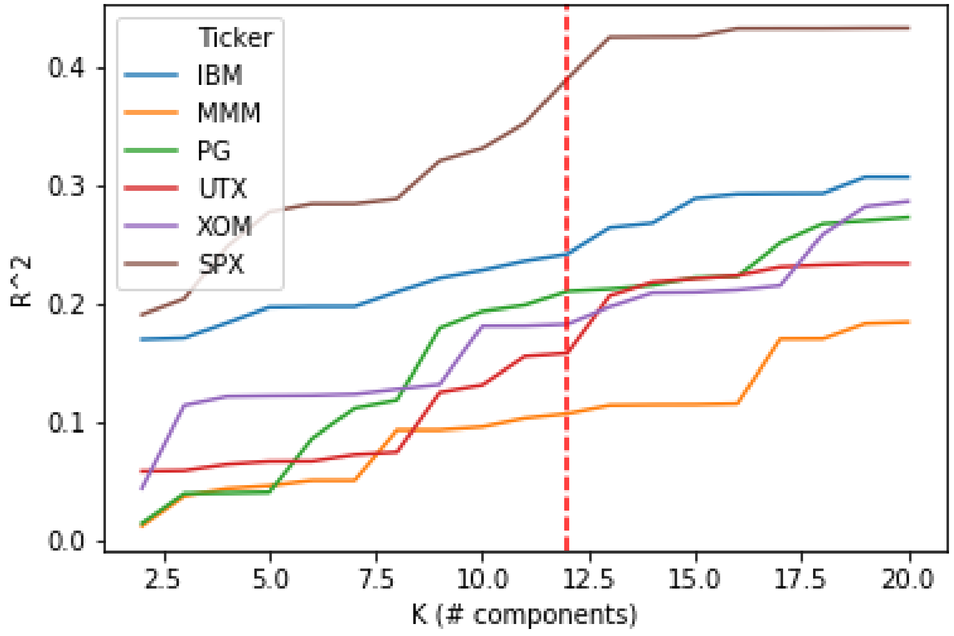


Tuning Number of Components for Kernel PCA (RBF Kernel, Gamma = $1x10^{-4}$)

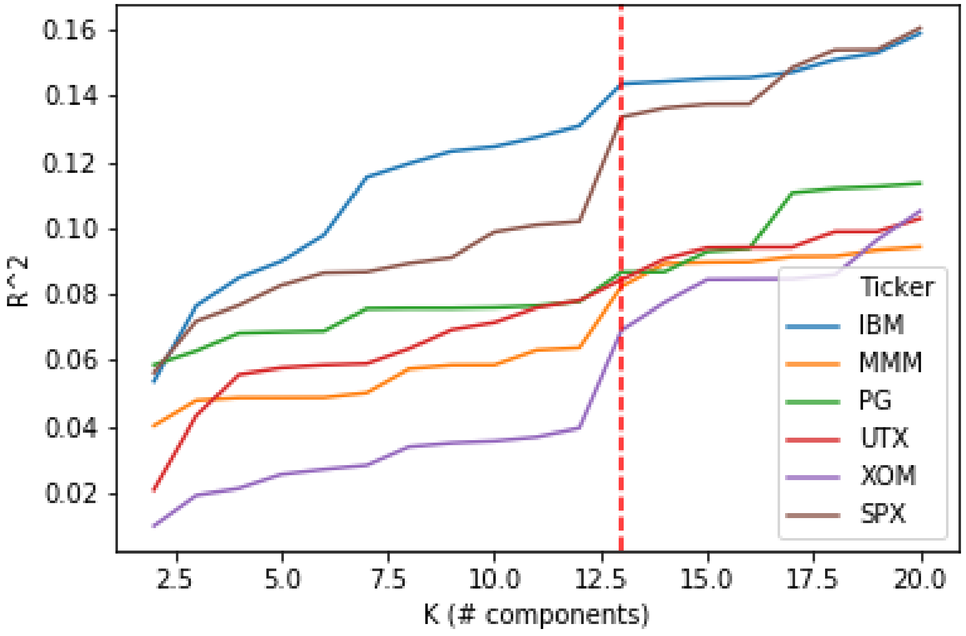


Results: Method Selection & Tuning

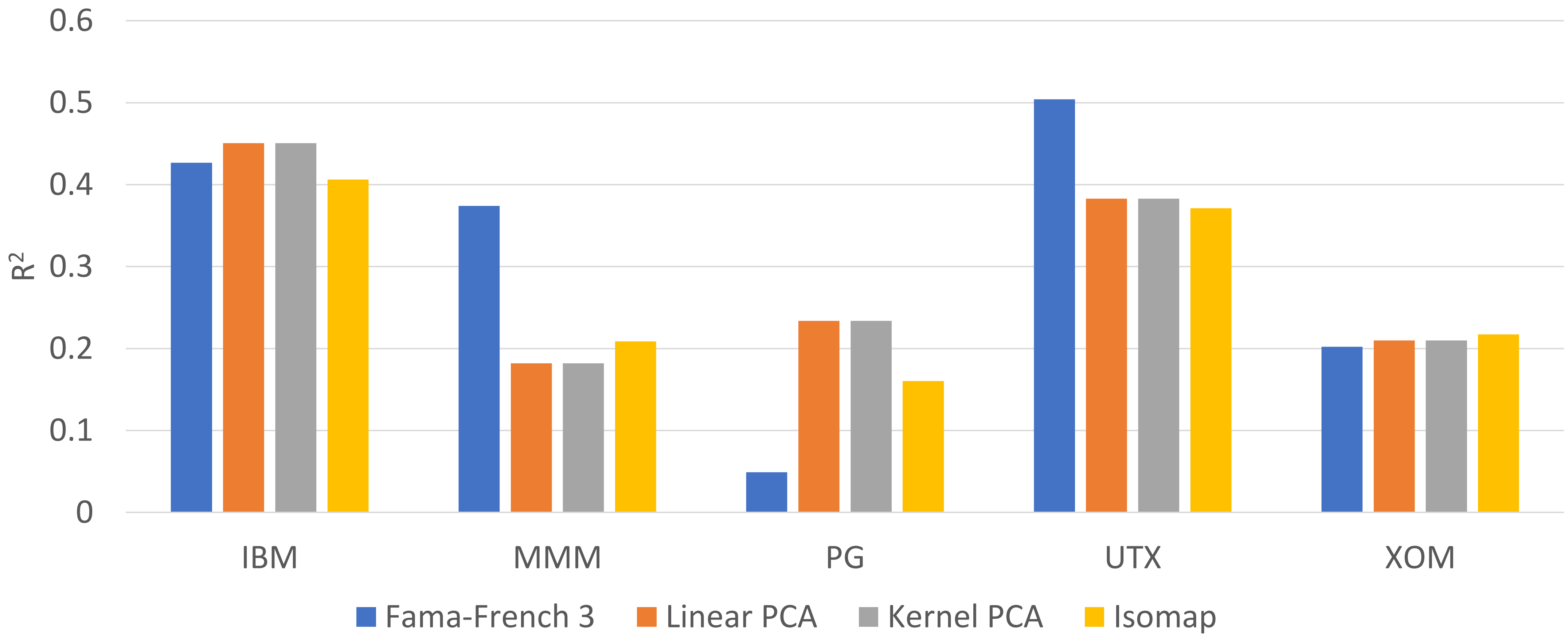
Tuning Number of Components for Isomap (kNN graph, nearest neighbors = 24)



Tuning Number of Components for Spectral Embedding (kNN graph, nearest neighbors = 6)



Results: Explanation of Returns



Conclusions & Future Work

- For explaining individual stock returns, non-linear dimensionality reduction methods don’t offer an improvement over standard linear PCA. This result suggests that the perspective that asset returns are a linear combination of returns to risk factors is broadly accurate.
- None of our learned representations performed better than the Fama-French model, except for one stock (Procter & Gamble). However, it’s hard to make a general conclusion, as this result might have been different if we used a different set of underlying factors.
- Future work could include doing the same analysis on different factor and/or response data, and using daily instead of monthly returns.

Questions?