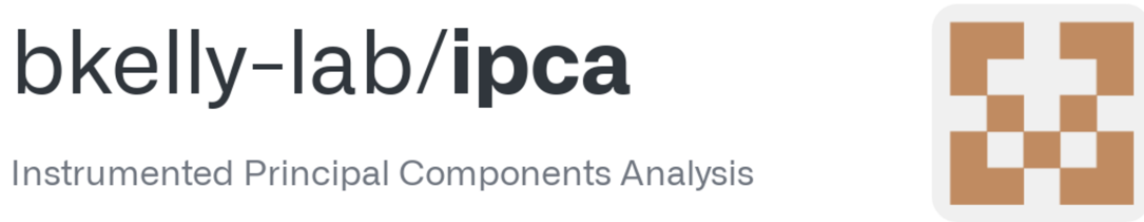
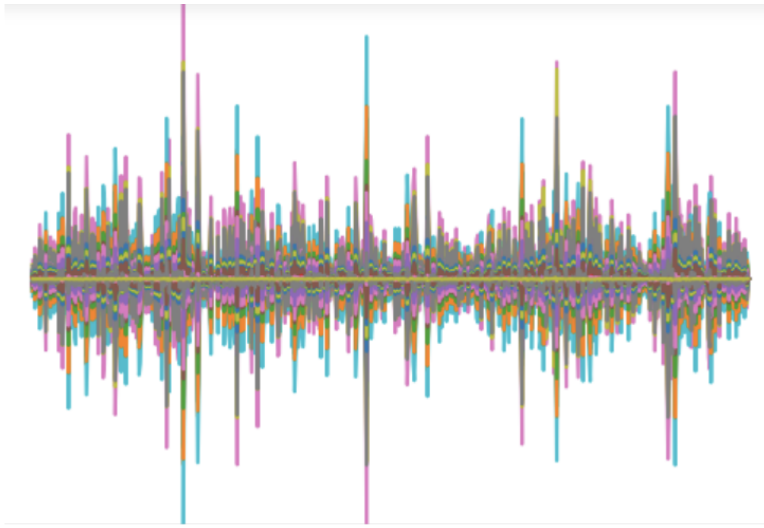


Risk & Return: Cond. FM / IPCA

Created	@February 5, 2023 3:06 PM
Period	@November 1, 2022 → March 1, 2023
Project Leaders	Cerina Yao

💡 Abstract: We study a new modeling approach for the cross-section of returns. Instrumented Principal Component Analysis (IPCA) allows for latent factors and time-varying loadings by introducing observable characteristics that instrument for the unobservable dynamic loadings. Studying returns and characteristics at the stock level, we find that five IPCA factors explain the cross-section of average returns significantly more accurately than existing factor models and produce characteristic-associated anomaly intercepts that are small and statistically insignificant.



Current Approach

Fama French

- Advantage: factors are fully observable
- Disadvantage: requires a previous understanding of the cross-section of average returns

Principal components analysis (PCA)

- Advantage: simultaneously estimate the factors and betas from the panel of realized returns, so requiring no ex-ante knowledge of the structure of average returns
- Disadvantage: only accommodate static loadings

Proposed New Approach: IPCA

- Allows factor loadings depend on observable asset characteristics that instrumented latent conditional loadings. Characteristics line up with average returns because they proxy for loadings on common risk factors
- If the “characteristics/expected return” relationship is driven by the compensation for exposure to latent risk factors, IPCA will identify the corresponding latent factors and betas.

- When a new anomaly characteristic is proposed, it can be included in an IPCA specification that also includes the long list of characteristics from past studies. Then, IPCA can estimate the proposed characteristic's marginal contribution to the model's factor loadings and, if need be, its anomaly intercepts, after controlling for other characteristics in a complete multivariate analysis.
- Evaluate large numbers of characteristic predictors with the minimal computational burden

Theoretical Framework:

Restricted model ($\Gamma\alpha = 0$):

Estimation objective is to minimize the sum of squared composite model errors.

$$r_{i,t+1} = z_{i,t}' \Gamma_\beta f_{t+1} + \epsilon_{i,t+1}^*$$

$$\min_{\Gamma_\beta, F} \sum_{t=1}^{T-1} \left(r_{t+1} - Z_t \Gamma_\beta f_{t+1}\right)' \left(r_{t+1} - Z_t \Gamma_\beta f_{t+1}\right).$$

Which means *ft*+1 and Γβ that minimize the error satisfy the first-order conditions.

$$\hat{f}_{t+1} = \left(\hat{\Gamma}_\beta' Z_t' Z_t \hat{\Gamma}_\beta\right)^{-1} \hat{\Gamma}_\beta' Z_t' r_{t+1}, \quad \forall t$$

Which is equivalent to

$$\max_{\Gamma_\beta} \ \mathrm{tr}\left(\sum_{t=1}^{T-1} \left(\Gamma_\beta' Z_t' Z_t \Gamma_\beta\right)^{-1} \Gamma_\beta' Z_t' r_{t+1} r_{t+1}' Z_t \Gamma_\beta\right)$$

This objective maximizes a sum of so-called "Rayleigh quotients" that all have the same denominator. In this special case, the well-known PCA solution is given by the first *K* eigenvectors of the sample second moment matrix of returns, *r r*′.

However, this concentrated IPCA objective is more challenging because the Rayleigh quotient denominators are different for each element of the sum and thus there is no analogous eigenvector solution Γβ, so we *approximately* solve IPCA by applying the singular value decomposition not to raw returns, but to returns interacted with instruments.

$$x_{t+1} = \frac{Z_t' r_{t+1}}{N_{t+1}}$$

In this case, we approximate the Rayleigh quotient denominators with a constant by replacing dynamic instruments by average instruments, then the solution Γβ would be first K eigenvectors of the sample second moment matrix of managed portfolio returns, *X′X*

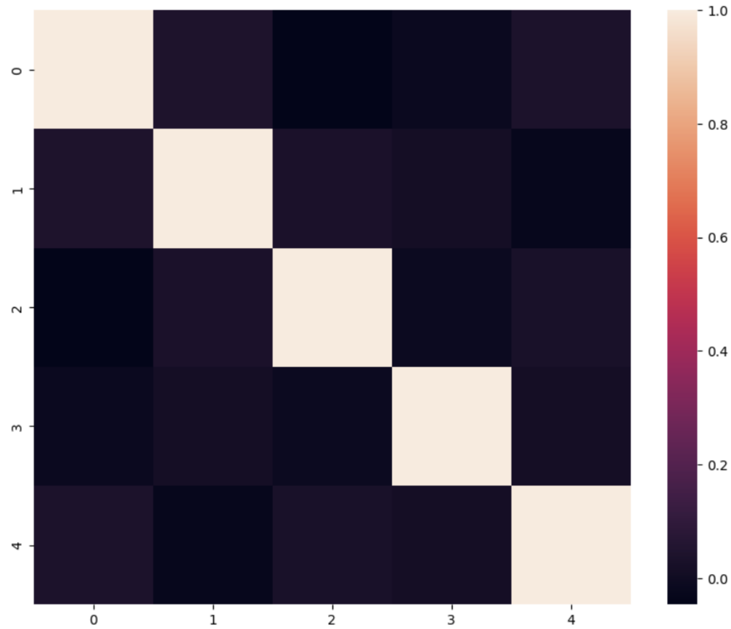
Unrestricted model (Γα≠0):

$$r_{i,t+1} = z_{i,t}' \Gamma_\alpha + z_{i,t}' \Gamma_\beta f_{t+1} + \epsilon_{i,t+1}^*$$

We rewrite it into the following form $\tilde{\Gamma} \equiv [\Gamma_\alpha, \Gamma_\beta]$, $\tilde{f}_{t+1} \equiv [1, f'_{t+1}]'$ so that it is in the same form as the Restricted model.

$$r_{i,t+1} = z_{i,t}' \tilde{\Gamma} \tilde{f}_{t+1} + \epsilon_{i,t+1}^*$$

To achieve econometric identification, we impose the additional identification assumption that Γα and Γβ are orthogonal subspaces. By imposing the assumption, we allow risk loadings to explain as much of assets' mean returns as possible. Of the total return predictability possessed by the instruments, only the orthogonal residual left unexplained by factor loadings is assigned to the intercept.



From the heat map, we can see that the final chosen 5 characteristics have a very low correlation. The low correlation among the factors indicates that each factor has a distinct effect on the returns of the assets in the portfolio. This is important because it means that the factors identified by IPCA can be used to construct a well-diversified portfolio that is exposed to a range of unique sources of risk. This can lead to higher returns and lower risk compared to a less diversified portfolio.

Overall, the low correlation among the factors identified by IPCA is a positive sign, as it suggests that the model is effectively capturing the unique sources of risk that drive asset returns and can be used to construct a well-diversified portfolio that is more likely to achieve the investor's objectives

Real Data Analysis

We cross-sectionally transform instruments period by period. In particular, we calculate stocks' ranks for each characteristic, then divide ranks by the number of non-missing observations and subtract 0.5. This maps characteristics into the [−0.5,+0.5] interval and focuses on their ordering as opposed to the magnitude. We use this standardization for its insensitivity to outliers and find in unreported robustness analyses that results are qualitatively unchanged with no characteristic transformation.

We estimate the K-factor IPCA model for various choices of K and consider both restricted ($\Gamma\alpha = 0$) and unrestricted versions ($\Gamma\alpha \neq 0$) of each specification. Two R2 statistics measure model performance.

- The first we refer to as the ”total R2” and define it as

$$\text{Total } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - z'_{i,t}(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{f}_{t+1})\right)^2}{\sum_{i,t} r^2_{i,t+1}}.$$

It represents the fraction of return variance explained by both the dynamic behavior of conditional loadings (and alphas in the unrestricted model), as well as by the contemporaneous factor realizations, aggregated over all assets and all time periods. The total R2 summarizes how well the systematic factor risk in a given model specification describes the realized riskiness in the panel of individual stocks.

- The second measure we refer to as the ”predictive R2” and define it as

$$\text{Predictive } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t+1} - z'_{i,t}(\hat{\Gamma}_\alpha + \hat{\Gamma}_\beta \hat{\lambda})\right)^2}{\sum_{i,t} r^2_{i,t+1}}.$$

It represents the fraction of realized return variation explained by the model’s description of conditional expected returns. IPCA’s return predictions are based on dynamics in factor loadings (and alphas in the unrestricted model). In theory, expected returns can also vary because risk prices vary. Without further model structure, IPCA does not separately identify risk price dynamics. Hence, we hold estimated risk prices constant, and predictive information enters return forecasts only through the instrumented loadings. When $\Gamma\alpha = 0$ is imposed, the predictive R2 summarizes the model’s ability to describe risk compensation solely through exposure to systematic risk. For the unrestricted model, the predictive R2 describes how well characteristics explain expected returns in any form, through loadings or through anomaly intercepts.

Setting optimal K=5, the results of the total R squared of the restricted model is 0.186, while the predictive R squared is 0.0069. The result of the total R squared of the unrestricted model is 0.187, while the predictive R squared is 0.0074.

Reference

Characteristics are covariances: A unified model of risk and return

We propose a new modeling approach for the cross section of returns. Our method, Instrumented Principal Component Analysis (IPCA), allows for latent f…

https://www.sciencedirect.com/science/article/abs/pii/S0304405X19301151?via=ihub#bib0021