Data Driven Methods in Finance:

Economic Factor Models

Fall 2023: IEOR 4576

Naftali Cohen

Oct 22, 2023

Introduction

- The economic factor model is the counterpart to the fundamental factor model. It also combines relevant stock information in an efficient way, but with a different twist on the factor model framework.

- The structure of the model remains the same as we move from the fundamental factor model to the economic factor model. The model still expresses the central idea that stock returns are the payoff for taking risk.

- In the fundamental factor model, stock returns are determined by the product of factor exposures (i.e., exposures to risk) and factor premiums (i.e., payoffs for exposure to risk). In the economic factor model, however, the roles of the factor exposure and factor premium are, in a way, reversed.

- Recall that for the fundamental factor model, the factor exposure is observable in financial statements, whereas the factor premium must be estimated from a cross-sectional regression. In the economic factor model, the factor premium is the known value, whereas the factor exposure must be estimated by a regression of stock returns on factor premiums.

Industrial Engineering and Operations Research
COLUMBIA ENGINEERING

Factor Categories

The strength of the economic factor model is that it can include practically all kinds of factors. In terms of how the model treats them, there are **three categories of factors**:

1. **Economic/behavioral/market factors**: Gross domestic product (GDP), inflation, unemployment, interest rates, and other macroeconomic variables; consumer sentiment index, business confidence index, investor sentiment index, or other survey-based indexes; returns on broad market indexes such as the Standard & Poor's (S&P) 500 or returns on other market group/industry indexes

2. **Fundamental/technical/analyst factors**: Log of market capitalization, book-to-price ratio, earnings-to-price ratio, debt-to-equity ratio, and other firm characteristics available through financial statements; momentum, trading volume, and other information reflected in trading data; analyst rating changes, earnings revisions, or other information provided by analysts

3. **Statistical factors**: Factors obtained from principal-component analysis applied to historical returns

Factor Premium

In the economic factor model, **the factor premium is the known value** (as opposed to the factor exposure, which is a regression estimate). This does not mean that one can always observe a factor premium directly, though.

- For economic/behavioral/market factors, the computation is rather trivial.
- For fundamental/technical/analyst factors, though, the computation is somewhat more demanding, and
- for statistical factors, it poses quite a bit of a challenge.

Factor Premium: Economic/Behavioral/Market Factors

	Unemployment Rate	Consumer Sentiment Growth	Excess Market Return
Jan 2020	3.60	0.50	−0.17
Feb 2020	3.50	1.20	−8.35
Mar 2020	3.50	−11.78	−12.47
Apr 2020	4.40	−19.42	12.82
Apr 2020	14.80	0.70	4.75
Jun 2020	13.30	8.02	1.98
Jul 2020	11.10	−7.17	5.63
Aug 2020	10.20	2.21	7.18
Sep 2020	8.40	8.50	−3.81
Oct 2020	7.80	1.74	−2.67
Nov 2020	6.90	−5.99	10.94
Dec 2020	6.70	4.94	3.83

Factor Premium for Fundamental/Technical/Analyst Factors

The computation involves constructing **zero-investment portfolios** and calculating their **returns**. A zero-investment portfolio simultaneously takes a long position in a portfolio of stocks with high factor exposures and a short position in a portfolio of stocks with low factor exposures.

Procedure:

- **Stock Ranking**: Rank all the stocks at time t in terms of the factor.
- **Portfolio Creation**: Create high-exposure and low-exposure portfolios by equally weighting the stocks in the top 33% of the list and in the bottom 33% of the list. (A critical value other than 33% may be justifiable.)
- **Zero-Investment Portfolio & Factor Premium**: Calculate the zero-investment portfolio return as the difference between the returns on the high-exposure and low-exposure portfolios. The return on the zero-investment portfolio is the factor premium for time t.

Factor Premium for Fundamental/Technical/Analyst Factors

	Log of Market Capitalization (LOGSIZE)	Book-to-Price Ratio (B/P)
Jan 2020	3.67	−8.35
Feb 2020	−1.00	−3.34
Mar 2020	6.71	−10.83
Apr 2020	−6.55	2.99
May 2020	1.22	−7.85
Jun 2020	−5.01	−1.67
Jul 2020	2.66	−6.96
Aug 2020	−1.74	1.98
Sep 2020	1.79	−3.81
Oct 2020	−0.83	4.74

Factor Premium for Statistical Factors

Obtaining factor premiums for statistical factors involves a rather intensive computation. The computational procedure is known as **principal-component analysis** (PCA) and is available on standard computer software packages:

- Estimate the variance-covariance matrix (Σ) from N stock returns over T time periods.

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{r}_t\mathbf{r}_t' - \bar{\mathbf{r}}\,\bar{\mathbf{r}}'$$

Factor Premium for Statistical Factors

Obtaining factor premiums for statistical factors involves a rather intensive computation. The computational procedure is known as **principal-component analysis** (PCA) and is available on standard computer software packages:

- Estimate the variance-covariance matrix (Σ) from N stock returns over T time periods.
- Once we have the variance-covariance matrix, we "diagonalize" it by finding an orthogonal matrix **Q** (that is, $\mathbf{Q}^{-1} = \mathbf{Q}^{T}$) such that.

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{r}_t \mathbf{r}_t' - \bar{\mathbf{r}}\,\bar{\mathbf{r}}' \qquad \mathbf{Q}'\,\hat{\Sigma}\,\mathbf{Q} = \mathbf{D} \qquad \mathbf{D} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$$

Factor Premium for Statistical Factors

Obtaining factor premiums for statistical factors involves a rather intensive computation. The computational procedure is known as **principal-component analysis** (PCA) and is available on standard computer software packages:

- Estimate the variance-covariance matrix (Σ) from N stock returns over T time periods.
- Once we have the variance-covariance matrix, we "diagonalize" it by finding an orthogonal matrix **Q** (that is, $\mathbf{Q}^{-1} = \mathbf{Q}^{T}$) such that.
- **D** is a diagonal matrix whose diagonal elements are eigenvalues of Σ. It turns out that each column of **Q** is an orthonormal (i.e., of unit length) eigenvector corresponding to eigenvalues of Σ.
- $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of Σ such that $\lambda_1 \geq \ldots \geq \lambda_N \geq 0$. (Since Σ is a positive definite matrix, all the eigenvalues are positive.)

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{r}_t \mathbf{r}_t' - \bar{\mathbf{r}}\,\bar{\mathbf{r}}' \qquad \mathbf{Q}'\,\hat{\Sigma}\,\mathbf{Q} = \mathbf{D} \qquad \mathbf{D} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$$

Factor Premium for Statistical Factors

Obtaining factor premiums for statistical factors involves a rather intensive computation. The computational procedure is known as **principal-component analysis** (PCA) and is available on standard computer software packages:

- Estimate the variance-covariance matrix (Σ) from N stock returns over T time periods.
- Once we have the variance-covariance matrix, we "diagonalize" it by finding an orthogonal matrix **Q** (that is, $\mathbf{Q}^{-1} = \mathbf{Q}^T$) such that.
- **D** is a diagonal matrix whose diagonal elements are eigenvalues of Σ. It turns out that each column of **Q** is an orthonormal (i.e., of unit length) eigenvector corresponding to eigenvalues of Σ.
- $\lambda_1, \dots, \lambda_N$ are the eigenvalues of Σ such that $\lambda_1 \geq \dots \geq \lambda_N \geq 0$. (Since Σ is a positive definite matrix, all the eigenvalues are positive.)
- Let $\mathbf{q}_1, \dots, \mathbf{q}_N$ be the orthonormal eigenvectors corresponding to $\lambda_1, \dots, \lambda_N$.
- If we want to find K factors, then we obtain K factor premiums by weighting individual stock returns using the first K columns of **Q**.
  That is, factor premiums $f_1, \dots, f_K$ are defined as

$$
\begin{aligned}
f_{1,t} &= \mathbf{q}_1' \mathbf{r}_t \\
&\vdots \\
f_{K,t} &= \mathbf{q}_K' \mathbf{r}_t
\end{aligned}
$$

Factor Exposure

In the economic factor model, factor exposures typically are determined from the time-series regression of stock returns on factor premiums.

We can estimate the following equation:

$$r_{it} = \alpha_i + \beta_i' \mathbf{f}_t + \epsilon_{it} \qquad t = 1, \ldots, T$$

where "coefficient" β_i is the factor exposure that we wish to discover.

Factor Exposure of Selected Stocks

Unemployment Factor Exposures:

- Negative: Microsoft, Johnson & Johnson, Amazon.
- Positive: Apple, Walmart (defensive stocks).

Value vs. Growth Classification:

- Value Stock (positive book-to-price exposure): Walmart.
- Growth Stocks (negative book-to-price exposure): Microsoft, Johnson & Johnson, Amazon, Apple.

Ticker	Unemployment Rate (UR)	Consumer Sentiment Growth (CSG)	Excess Market Returns (MKT)	Log of Market Capitalization (LOGSIZE)	Book-to-Market Ratio (B/P)
AAPL	0.215	0.226	1.418	−0.066	−0.647
	(0.411)	(0.191)	(0.218)	(0.354)	(0.310)
MSFT	−0.297	0.127	1.008	0.203	−0.320
	(0.233)	(0.108)	(0.124)	(0.201)	(0.176)
WMT	0.231	−0.302	0.473	0.660	0.244
	(0.280)	(0.130)	(0.148)	(0.241)	(0.211)
JNJ	−0.309	−0.141	0.768	0.191	−0.025
	(0.233)	(0.108)	(0.123)	(0.200)	(0.175)
AMZN	−0.356	−0.181	1.345	−0.088	−0.841
	(0.379)	(0.176)	(0.201)	(0.326)	(0.285)

When the Standard Approach Fails

To run time-series regressions, portfolio managers need to have enough data on stock returns and factor premiums to thoroughly cover a reasonable time period at regular time intervals. Recent initial public offerings (IPOs) or stocks of recently merged or divested companies lack sufficient data for meaningful regressions.

- **Merges**: find the weighted average of the factor exposures of the two premerger firms (s_A is the premerger market capitalization of firm A, and s_B is the premerger market capitalization of firm B)

$$\hat{\beta}_{AB} = \frac{s_A}{s_A + s_B}\hat{\beta}_A + \frac{s_B}{s_A + s_B}\hat{\beta}_B$$

- **Recent IPO**: find similar firms and take the average factor exposures of those similar firms

$$\hat{\beta}_C = \frac{1}{M}\left(\hat{\beta}_1 + \cdots + \hat{\beta}_M\right)$$

Disclaimer

This course is for educational purposes only and does not offer investment advice or pre-packaged trading algorithms. The views expressed herein are not representative of any affiliated organizations or agencies. The main objective is to explore the specific challenges that arise when applying Data Science and Machine Learning techniques to financial data. Such challenges include, but are not limited to, issues like short historical data, non-stationarity, regime changes, and low signal-to-noise ratios, all of which contribute to the difficulty in achieving consistently robust results. The topics covered aim to provide a framework for making more informed investment decisions through a systematic and scientifically-grounded approach.