

Multi-Factor Models and Beyond

Version 1.0

Dr. Ye Luo
HKU Business School
Jan 2023

Fundamental Question in Financial Asset Pricing

- Huge number of assets: 5000+ stocks in Chinese Exchanges, 20000 to 30000 in US market. Other asset classes: Fixed income, Futures, Derivatives, Forex, etc.
- A Question from normative stand point: What explains asset returns?
- A Question from positive stand point: What predicts asset returns?

- Market is mostly efficient. In most of the scenarios and for most of the assets, future prices are hard to predict - often known as the efficient market hypothesis.
- Higher return, higher risk. In most of the scenarios, the only way to earn return higher than the risk free rate is to take more risk.
- Beating the market is extremely hard. 99% of the institutions can hardly beat market each year. In most of the scenarios, the market price move is simply like a random walk.

Random Walk View on the Price

- A simple model for a sequence of random walk can be viewed as:

$$Y_{t+1} = Y_t + \epsilon_t,$$

where $\mathbb{E}[\epsilon_t] = 0$.

- The random walk model produces wave like curves that oscillates back and forth - hard to predict, and therefore, hard to beat.



Figure 1: Exchange Rates

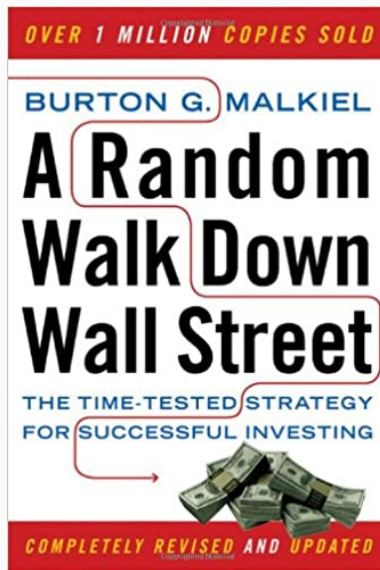


Figure 2: A Random Walk Perspective

A Probabilistic View on Financial Markets

- Suppose you have a dice, with 1 to 6 on its faces. You win 1 dollar when you roll 1-3, and loss 1 dollar when you roll 4-6. The expectation of the dice is 0.
- If the dice is fair, i.e., the probability of each face is $1/6$, it becomes impossible to beat the dice.
- Think market as a dice. You may only beat the market when you have an unfair dice. For example:
 1. (High Winning Probability) winning probability of 60% rather than 50%.
 2. (High Winning Edge) You win 2 dollars when the dice hits 5 or 6.
- The question is: if there is such unfair dice, would it be arbitrated away by other traders, even a long time ago?

The Two Fundamental Ways to Earn Abnormal Return

- Mis-Pricing: market could be inefficient at 5% of the time and 5% of the assets. Leading to a small room of arbitrage.
- Risk Premium: Market is still quite efficient, but you earn additional risk premium by taking exposure on certain type of risks. Leading to earn risk premiums when exposed to certain type of risks.
- You can diversify the risks by allocating your bets on different market. Diversification is probably the only free lunch in the financial market.

- Return R is a random variable. Expect return (mean return) is $\mathbb{E}[R]$, which can be estimated by the average return of the past history.
- Volatility σ is the standard deviation of R , i.e., $\sigma = \sqrt{\text{Var}(R)}$.
- Sharpe Ratio (SR) is defined as:

$$SR := \frac{\mathbb{E}[R] - r_f}{\sigma},$$

where r_f is the risk free rate.

- In many cases, industry uses the non-risk free rate adjusted Sharpe Ratio:

$$SR := \frac{\mathbb{E}[R]}{\sigma}.$$

- Sharpe Ratio SR measures the relative ratio between expected return and risk.
- The higher the Sharpe Ratio is, the better the strategy is.

Simple Facts on Market Benchmarks

- Annualized expected return of SP500 index is 4-8%.
- Annualized Volatility is around 20%.
- Sharpe Ratio is around 0.2 – 0.4.
- A good trading strategy needs to have high Sharpe Ratio (after transaction costs). Typically > 1.4 , $SR > 2$ or $SR > 3$ gives strong performances. Renaissance's Medallion fund has Sharpe Ratio ≈ 5 .

- Standard Capital Asset Pricing Model: for asset i and time t ,

$$R_{it} - r_f = \alpha_i + \beta_i(R_{mt} - r_f) + \epsilon_{it},$$

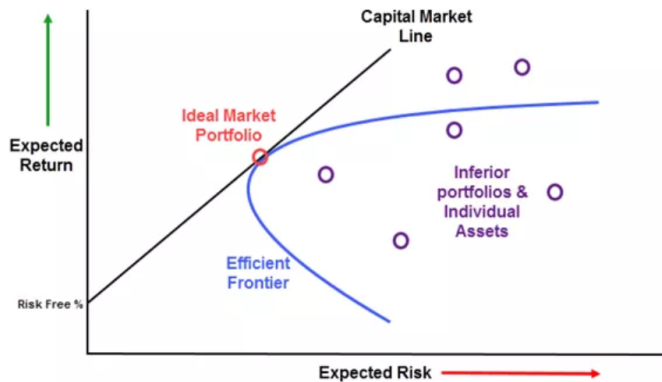
where R_{mt} is the return of the market (indices like Shanghai A or NASDAQ).

- $(R_{mt} - r_f)$ is known as the market factor.
- α_i : excess return of asset i - a constant estimated return of this particular asset.
- β_i : loading on market factor of asset i - a return that is associated with the move of the market.

The Risk Premium Perspective

- One way to construct high Sharpe Ratio strategy is to bet on Risk Premium.
- Risk Premium refers to the excess return exposed to certain type of risks, as market is close to efficient.
- Different risk premium carries over different types of risks and different excess returns.
- Allocate weights by diversifying on different risk premium may produce high Sharpe Ratio strategy - hold a portfolio of assets at a time helps to reduce risks.

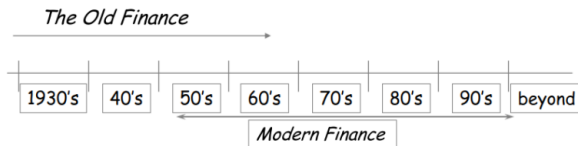
The CAPM



The Key Implications of CAPM

- Idiosyncratic risks ϵ_{it} can be washed away by diversification. Only “Market” risk matters.
- Market portfolio is always on the mean-variance frontier, and the market-capital line is tangent to the MVF.
- Should always hold a combination of market assets and risk free asset.
- It indicates that all α_i should be 0. However, is this really true?

Haugen's view: The Evolution of Academic Finance



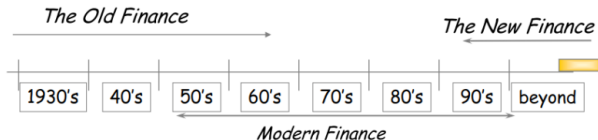
Modern Finance

Theme: Valuation Based on Rational Economic Behavior

Paradigms: Optimization Irrelevance CAPM EMH
(Markowitz) (Modigliani & Miller) (Sharpe, Lintner & Mossen) (Fama)

Foundation: Financial Economics

Haugen's View on Financial Economics



The New Finance

Theme: Inefficient Markets

Paradigms: Inductive *ad hoc* Factor Models
Expected Return Risk

Behavioral Models

Foundation: Statistics, Econometrics, and Psychology

How to dig excess returns in today's market?

- Arbitrage strategy: find mis-pricing patterns that constantly exist.
- α - strategy: find patterns of strong risk premium. The market β exposure is 0.
- Smart β - strategy: dynamically allocate portfolio on different factors/risk exposures.
- Hedging: reduce other unnecessary risk exposure by hedging.

- Many assets traded in the market. Denote N as the number of assets. $N \approx 20000$ in the US equity market and $N \approx 5000$ in the Chinese equity market.
- Without any structures, one needs to learn:
 1. Expected return $\mathbb{E}[R_i]$, $i = 1, 2, \dots, N$.
 2. Volatility σ_i , $i = 1, 2, \dots, N$.
 3. Co-variance $\text{Cov}(R_i, R_j)$, $i, j \in \{1, 2, \dots, N\}$
- In the multiple asset cases, the co-Variance $\text{Cov}(R_i, R_j)$ is difficult to estimate, and therefore, it is difficult to estimate the risk of a portfolio of assets.

- Recall that CAPM does:

$$R_{it} - r_f = \alpha_i + \beta_i(R_{mt} - r_f) + \epsilon_{it}.$$

- Systematic risk: all assets co-move with R_{mt} . Idiosyncratic risk: ϵ_{it} .
- In CAPM: asset i and asset j are only correlated through the market factor $R_{mt} - r_f$.
- Therefore, the covariance of asset i and j equals to

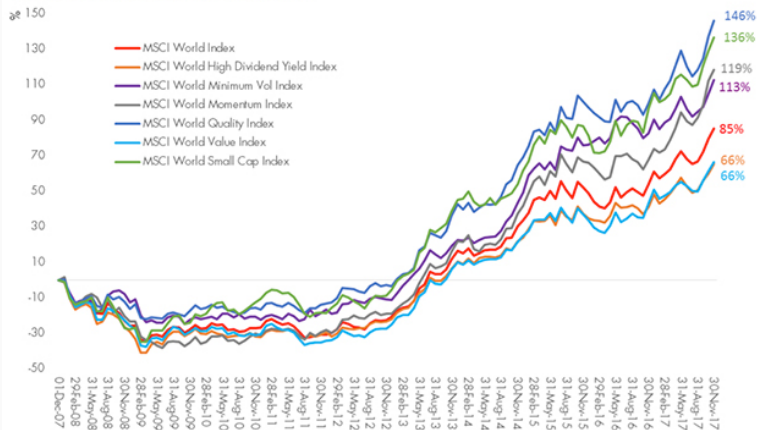
$$\beta_i \beta_j \text{Var}(R_{mt}).$$

Leads to dimensionality reduction.

Factors Beyond The Market

Performance of MSCI World Indices

1 December 2007 to 30 November 2017



Size Factor in the MV Space

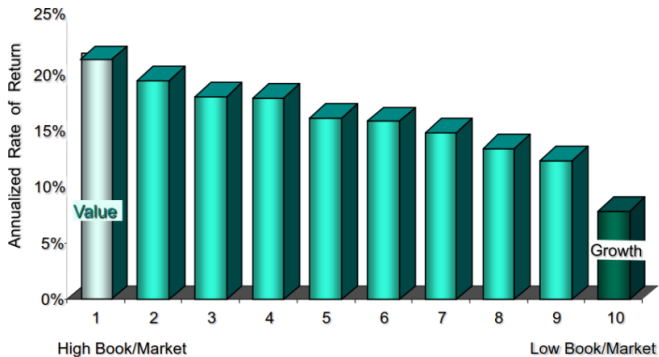
Risk Return of MSCI World Indices

1 December 2007 to 30 November 2017



- It is well-known that small cap has a stronger return over time compared to large cap.
- As a result, this additional premium is different from Market risk premium - you can consider a long short portfolio in which you hold +1 dollar in small caps and hold -1 dollar on large caps.
- Why this premium exists? This premium delivers mostly positive return, but can suffer from huge losses sometimes, e.g., economic recessions.

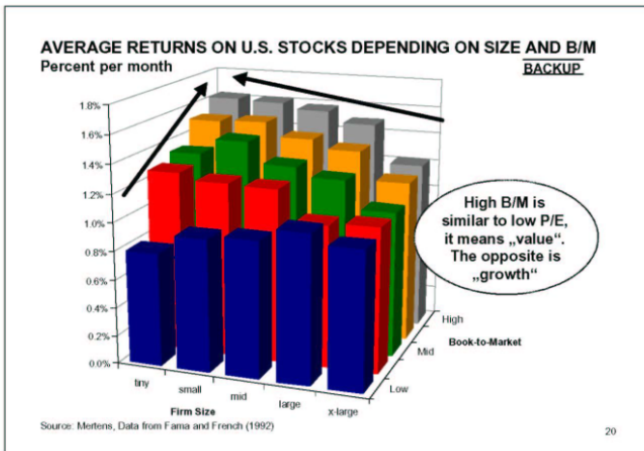
Book to Market as a Predictor of Return



- It started with Fama and French (1994)'s work in three factor models. This work wins the Nobel's prize in economics later in 2013.
- The three factors involves:
 1. CAPM's market excess return $R_m - r_f$ (this can be considered as a long-short portfolio of the market and risk-free bond).
 2. Book to market ratio (B/M). High book to market tends to have higher returns.
 3. Firm size. Larger firms tend to have smaller returns.

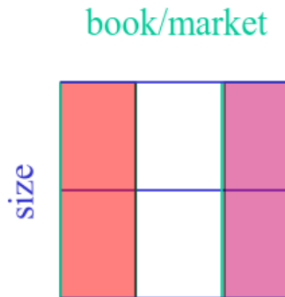
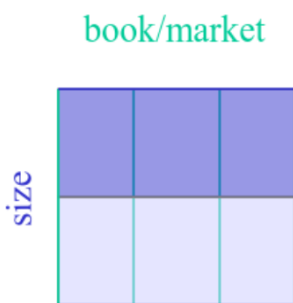
Book to Market and Size

Small „value“ companies have higher returns



The Sorting Approach

- We can rank assets based on size or B/M ratio, from low to high.
- The factor is constructed by a long-short portfolio, i.e., buy the high ranks and sell the low ranks or vice versa.



The three factor model

- SMB: long small size stocks and short large size stocks.
- HML: long high B/M ratio stocks and short low B/M ratio stocks.
- Three factor model:

$$R_{it} - r_f = \alpha_i + \beta_1(R_{mt} - r_f) + \beta_2SMB + \beta_3HML + \epsilon_{it}.$$

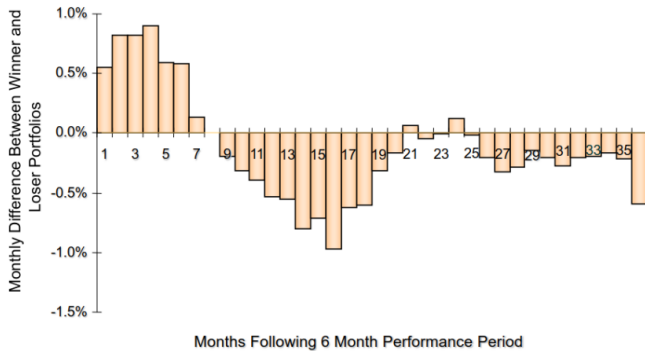
How to interpret the three factor model?

- The CAPM model

$$R_{it} - r_f = \alpha_i + \beta(R_{mt} - r_f) + \epsilon_{it}.$$

- The CAPM α is partially explained by the risk premium of SMB and HML.
- If we increase the factors, would the α be completely explained? - Factor represents the fundamentals of the individual assets, as different assets have different β loadings on factors.

Beyond three factors



- K factors f_{1t}, \dots, f_{Kt} . Vectorized:

$$f_t = \begin{pmatrix} f_{1t} \\ f_{2t} \\ \dots \\ f_{Kt} \end{pmatrix}$$

- Return

$$R_{it} - r_f = \alpha_i + \beta_i^\top f_t + \epsilon_{it},$$

where $\beta = (\beta_1, \dots, \beta_K)$ is a vector.

- The regression R^2 measures how well the model fits into the returns.

Estimating α , β and beyond

- Factor regression: $R_{it} - r_f = \alpha_i + \beta_i^\top f_t + \epsilon_{it}$.
- Once $\hat{\alpha}_i$ and $\hat{\beta}_i$ are estimated at time t , denoted as $\hat{\alpha}_{it}$ and $\hat{\beta}_{it}$.
- Fama-Macbeth Regression: Regress $R_{i,t+1}$ on $\hat{\beta}_i$ over all i, t .

$$R_{i,t+1} = \tilde{\alpha}_i + \hat{\beta}_i \tilde{F} + \eta_{i,t+1}$$

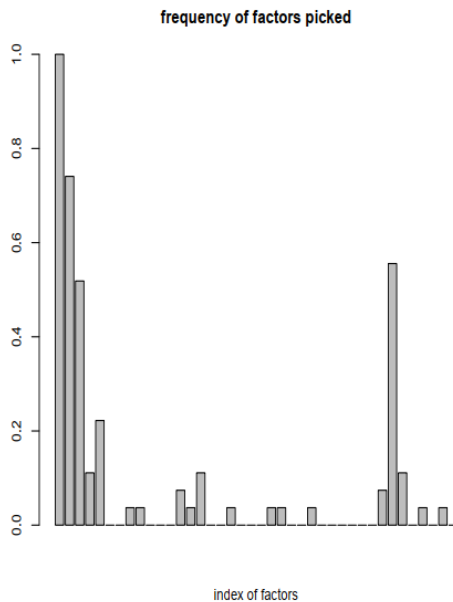
- The regression creates an estimate on \tilde{F} , which is an estimate of the $\mathbb{E}[f_t]$, the risk premium carried by the factor.
- The Fama-Macbeth regression estimates the risk premium vector \tilde{F} that is carried by the factor vector f .

Case Study: how to evaluate the Fama-French Three Factor Models?

- A long history of researches on digging factors. Large pool of papers. Modern approach: use machine learning to find factors.
- Overall 289 factors have been proposed, and “being showed” to be effective in pricing stocks.
- Data: all stocks traded in NYSE and NASDAQ from 1988 to 2015, monthly data. All 289 factors are considered. Total number of observation = 270k. Rolling window of 5 years.
- Question to answer: which factors are important and really explains asset returns?

- Factor selection by machine learning: G-Boosting Algorithms.
- The algorithm uses a voting mechanism: each stock creates a ranking for all the existing factors. Then, a score is computed by taking average of their rankings over the pool of stocks.
- The predictor with the highest ranking is chosen.
- Idea: the stocks vote for the “best predictor” they likes. After that, they vote for the next one.
- Stops at iteration K - which are the K most favorite factors that NYSE and NASDAQ stocks are explained by?

Factor Selection Frequency over time



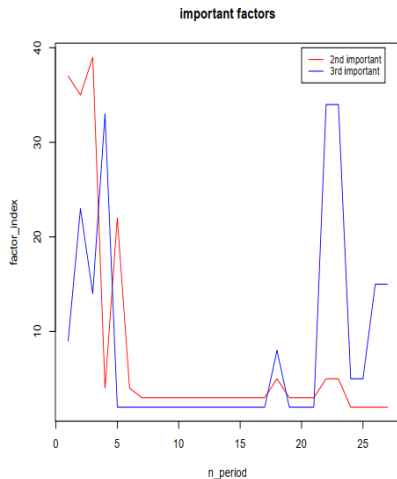
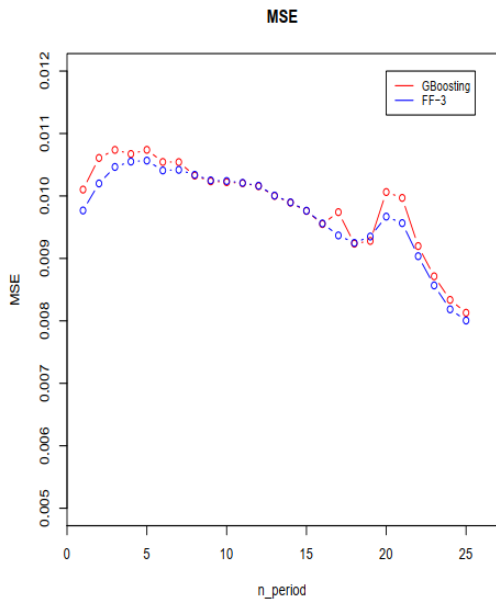
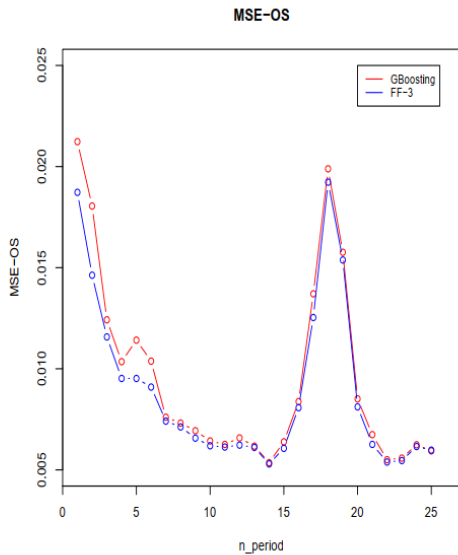


Figure: 2nd and 3rd most important factors in the history

Average In Sample MSE



Average Out-of Sample MSE



- Fama-French factors lose explainability power in financial crisis.
- Market pattern switches from growth to risk sheltering. Gold prices go up, bond yield goes down, stock prices goes down.
- FF model only explains the stock returns. It does not predict the stock returns. How do we utilize such explainability?
- When hedging these factors, portfolio becomes closer to pure α returns - less risk exposure but may lead to less returns as well. Portfolio construction and optimization? Next lecture.