Big Data Analysis

Lecture 13

2019/12/30

Factor Investing

- Relative returns across all tradables
  - Cross sectional
  - Long short portfolios
- Absolute return vs. Index enhancement
- Core task: forecast of expected returns
- Long those with high expected returns and short low expected returns

Source: High Flyer Quant

Expected Returns

- $E[R_i] = \beta_i \lambda + \alpha_i$
- To minimize α_i
  - Gain maximum difference between expected returns
  - Avoid uncertainty
- Key steps
  - Select factors and calculate the exposure β_i
  - Find the above linear relationship through regression
  - Calculate α_i and test whether all of them are jointly zero

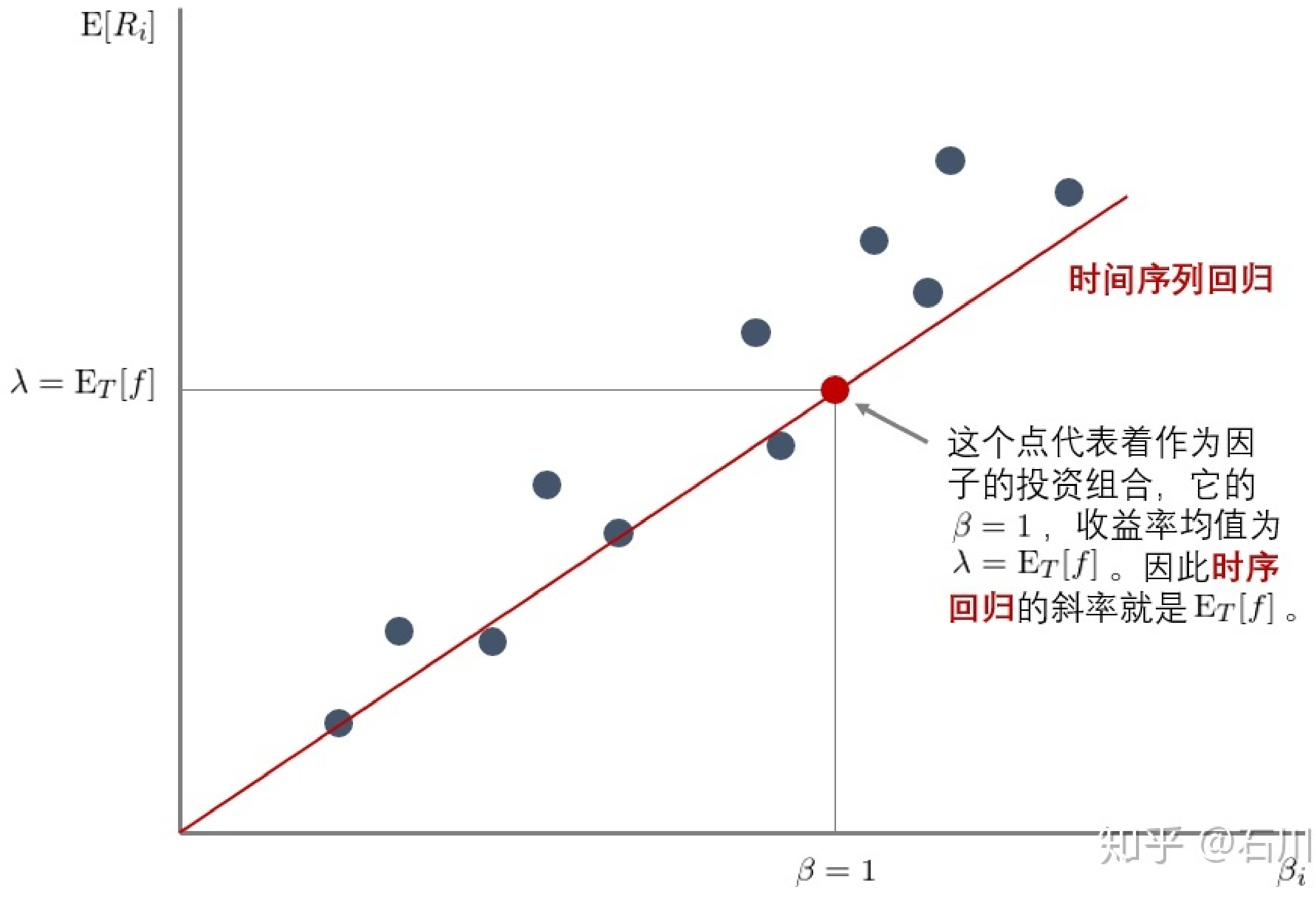
Ways to Estimate Expected Returns

- Time Series Regression
- Cross Sectional Regression
- Fama-MacBeth Regression
- Barra Models

Time Series Regression

- Factors are portfolios returns, Fama and French (1993)
- Stocks are sorted and grouped to form long short portfolios
- With periodic rebalance, portfolio returns are achieved
- Let f_t be factor returns during t period,
$$R_{it} = \alpha_i + \beta_i f_t + \varepsilon_{it}, t = 1, 2, \dots, T$$
- Taking expectation: $E_T[R_i] = \beta_i E_T[f_t] + \alpha_i$
- Comparing to $E[R_i] = \beta_i \lambda + \alpha_i$, we have $\lambda = E_T[f_t]$

Time Series Regression



石川@知乎 https://www.zhihu.com/people/mitcshi/posts

Cross Sectional Regression

- Factors don't have to be portfolios returns
- Two step regression estimate
- Beta estimation

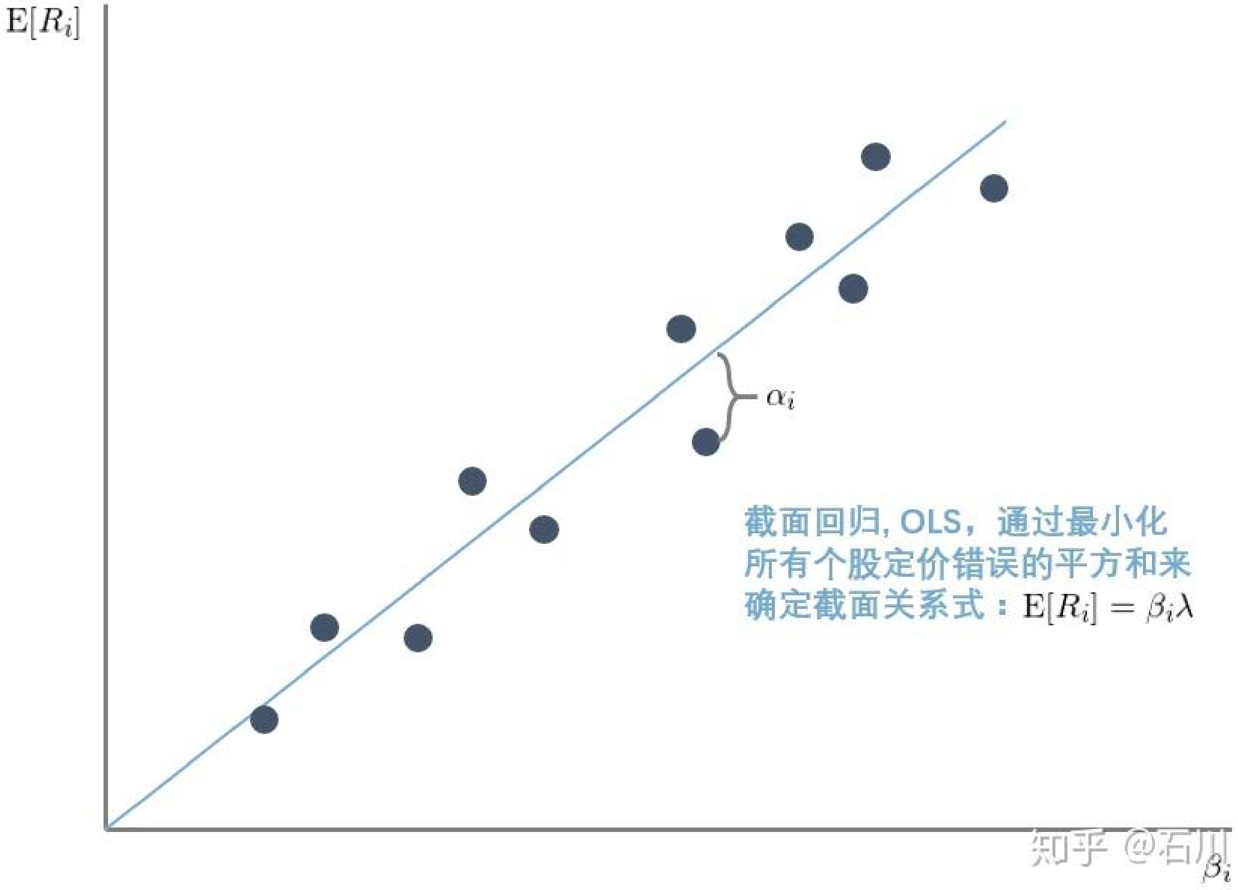
$$R_{it} = a_i + \beta_i f_t + \varepsilon_{it}, t = 1, 2, \ldots, T$$

- Average stock returns to estimate factor return using GLS

$$E_T[R_i] = \beta_i \lambda + \alpha_i, i = 1, 2, \ldots, N$$

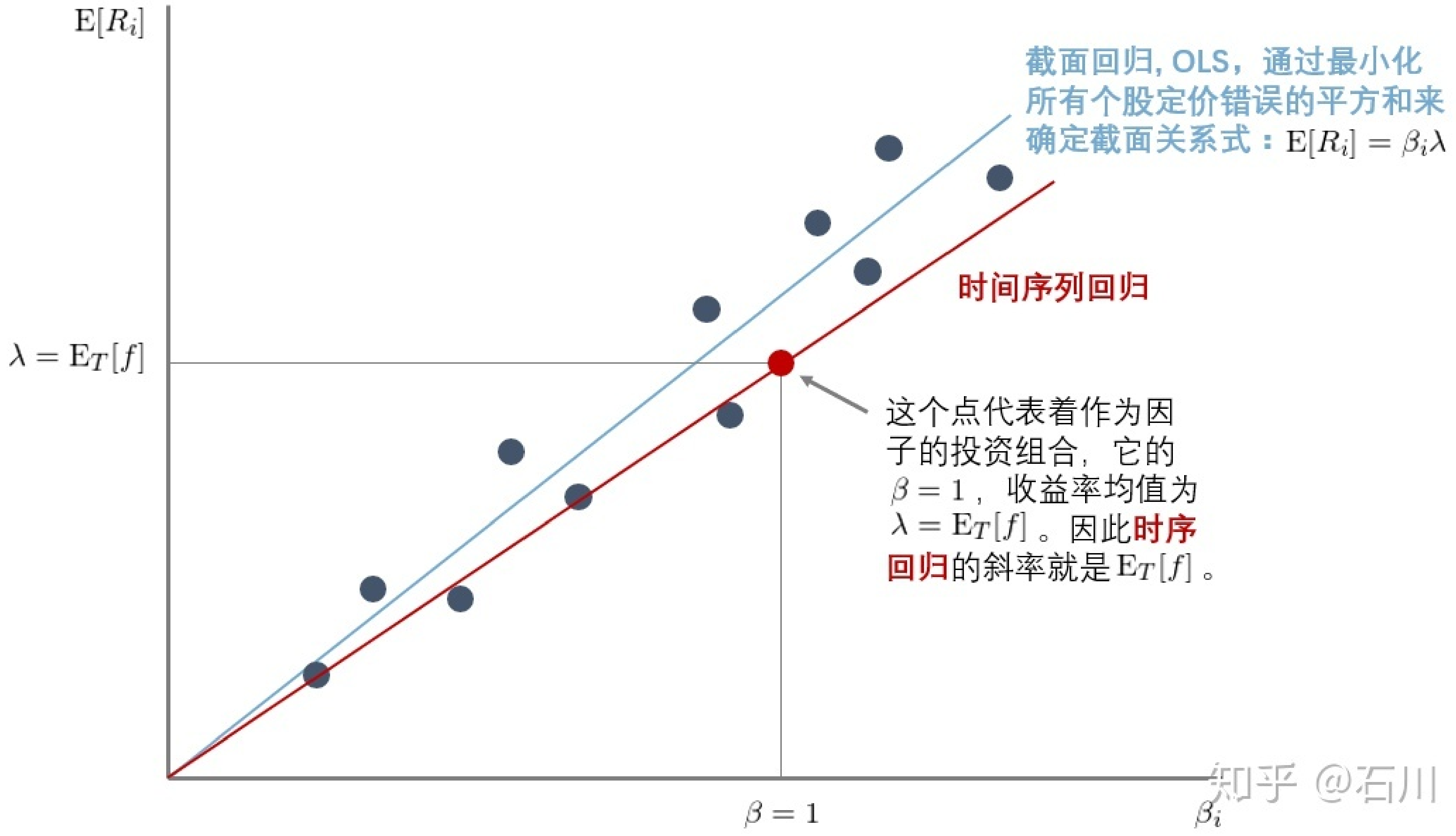
- Note that β_i are estimated, therefore the covariance matrix of α_i are not easy to estimate. GMM is needed to test α_i

Cross Sectional Regression

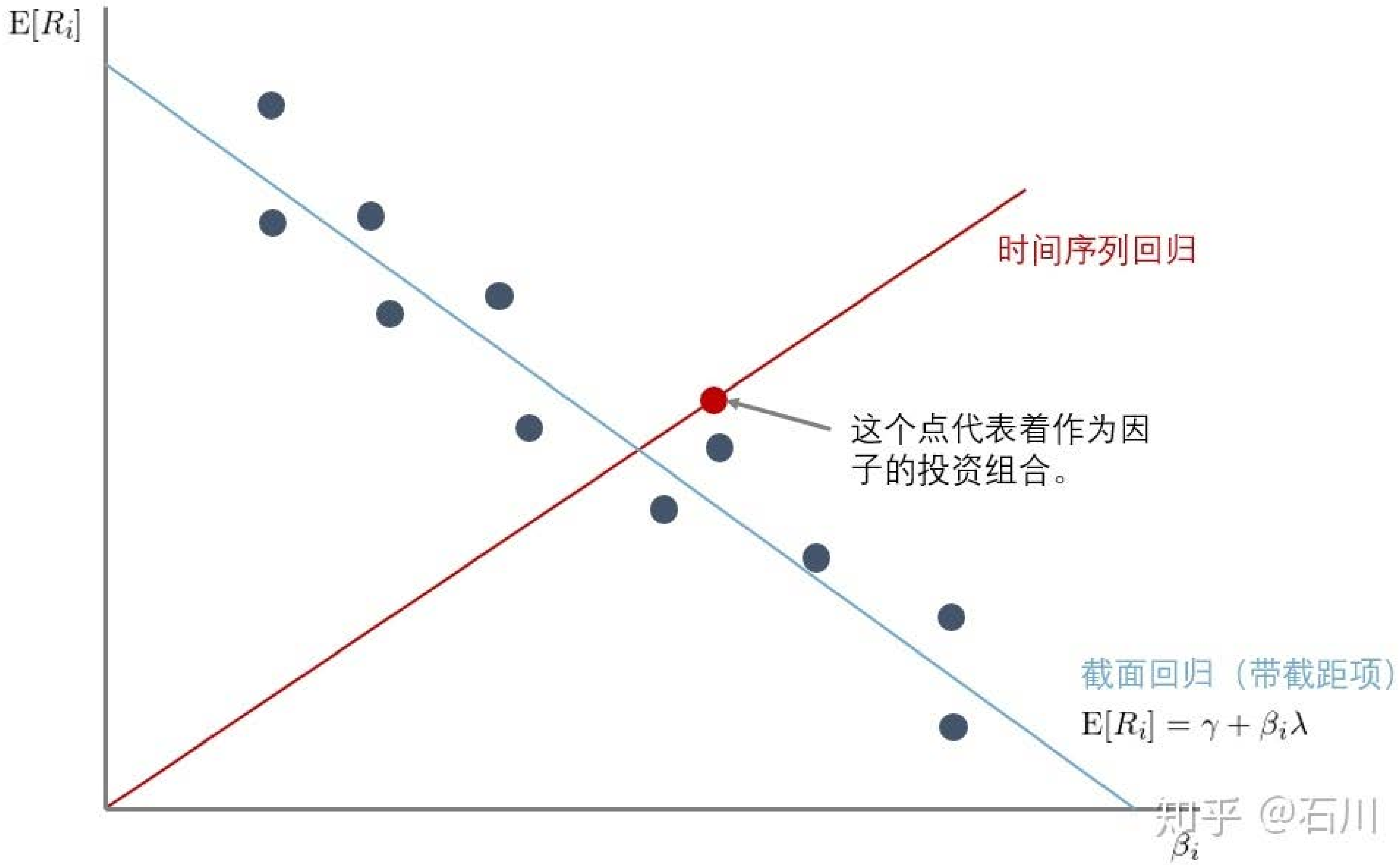


石川@知乎 https://www.zhihu.com/people/mitcshi/posts

Time Series vs. Cross Sectional



Time Series vs. Cross Sectional



石川@知乎 https://www.zhihu.com/people/mitcshi/posts

Fama MacBeth Regression

- Commonly used empirical model for asset pricing
- Portfolios instead of individual stocks are listed on the LHS

$$R_{it} = a_i + \beta_i f_t + \varepsilon_{it}, t = 1, 2, \dots, T$$

- Typically Size B/M double sorted 5x5=25 portfolios are used
- Cross sectional regression is done for every t

$$E_t[R_{it}] = \beta_i \lambda_t + \alpha_i, i = 1, 2, \dots, N$$

- Then

$$\lambda = E_T\ [\lambda_t]$$

- Fama-MacBeth takes care of errors-in-variables partially

Mitchell A. Petersen

- Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, The Review of Financial Studies, 2009

- Although the literature has used an assortment of methods to estimate standard errors in panel data sets, the chosen method is often incorrect and the literature provides little guidance to researchers as to which method should be used. In addition, some of the advice in the literature is simply wrong. Since the methods sometimes produce incorrect estimates, it is important to understand how the methods compare and how to select the correct one. This is the paper's objective.

- https://www.nber.org/papers/w11280

Barra Model

- First started as a risk model, not for alpha model
- Standard model for risk decomposition
- Good at estimating the covariance matrix of individual stocks

- Barra model is also a cross sectional regression model
- Time series regression introduce delay and is subject to slow adjustment, while cross sectional regression is more responsive
- The key difference is how to calculate the exposure β_i
- Instead of using an estimated β_i, Barra takes the characteristics of stocks directly as exposure

Barra Model – Pure Factor Model

- Barra Risk Model Handbook (2007). MSCI
- The exposure in Barra model needs normalization

$$E[R_i] = \beta_i f_t + \alpha_i$$

- To resolve the weight matrix of stocks on pure factor portfolios
- Constraints: a pure factor model
  - Country factor: fully invested, market value weighted
  - Industry factor: long industry and short country
  - Style factor: long high exposure and short low exposure
- Industry exposure of individual stocks are relatively coarse
- Refer to the guest lecture (lecture 9) from China Scope

From Academic View

- Empirical asset pricing models
- Key questions to answer
  - What are anomalies and factors?
  - What is the difference between anomalies and factors?
  - How many factors to be included in a multi-factor model?
  - What are the main stream multi-factor models in academia?
  - How to compare and choose multi-factor models?

Anomalies

- An asset pricing anomaly refers to the statistically significant difference between two returns, the predicated returns of a given asset pricing model, and the returns of long/short portfolio formed based on some characteristics
- Time series regression

$$R_{it} = \alpha_i + \beta_i f_t + \varepsilon_{it}, t = 1, 2, \dots, T$$

- Portfolio returns ($i = 1, 2, \dots, N$) listed on the LHS and factor returns on the RHS
- Check whether $\alpha_N - \alpha_1$ is statistically significantly different from 0

Anomalies – Example

Lee, C. M. C., S. T. Sun, R. Wang and R. Zhang (2018). Technological Links and Return Predictability. Journal of Financial Economics, forthcoming.

Panel A: Portfolio returns

Decile	Excess returns (%)	CAPM alpha (%)	3-Factor alpha (%)	4-Factor alpha (%)	5-Factor alpha (%)	6-Factor alpha (%)
1	0.42	-0.16	-0.33	-0.13	-0.27	-0.12
(Low)	(1.46)	(-1.04)	(-2.66)	(-1.09)	(-2.17)	(-0.99)
10	1.59	1.05	0.93	0.95	1.10	1.09
(High)	(5.38)	(5.37)	(6.41)	(6.36)	(8.11)	(7.97)
L/S	**1.17**	**1.22**	**1.26**	**1.08**	**1.37**	**1.21**
(Equal weights)	(5.47)	(5.70)	(5.88)	(4.98)	(6.49)	(5.76)
L/S	**0.69**	**0.74**	**0.80**	**0.65**	**0.86**	**0.73**
(Value-weights)	(3.19)	(3.40)	(3.62)	(2.91)	(3.81)	(3.24)

出处： Lee et al. (2018)

Anomalies Everywhere?

- There are 400+ anomalies found in US stock market, see Hou, Xue and Zhang (2017), Replicating Anomalies.
- Possible reasons for these many anomalies
  - Data mining
  - The choice of asset pricing models on the RHS
- Some references
  - Harvey, Liu and Zhu (2016)
  - Hou, Xue and Zhang (2015, 2017)
  - Linnainmaa and Roberts (2018)

P-Value: 2 or 3?

- According to Harvey, Liu and Zhu (2016)
  - # of T-Stat within [2, 2.57] ≈ # of T-Stat within [2.57, 3.14]
  - T-Stat = 2.57 (p-value = 0.005) while T-Stat = 3 (p-value = 0.001)
  - Publication bias
  - P-hacking
- HARKing: Hypothesizing after the results are known
- Caution on P-Value
  - False discovery rate
  - Bayesianized p-value

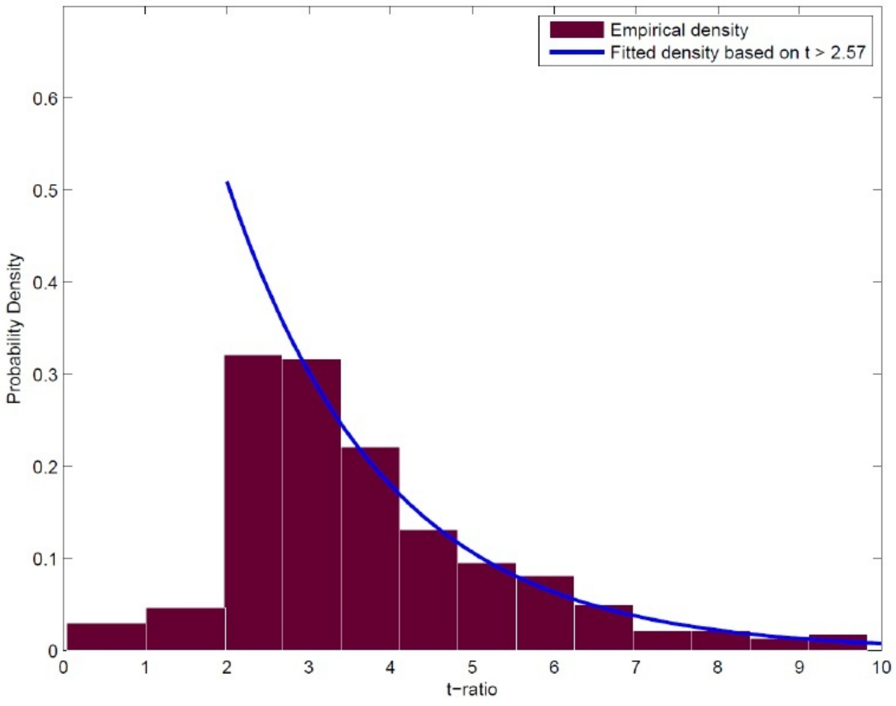


Figure 1. Distribution of reported *t*-statistics from factor studies, 1963 to 2012. Evidence from Harvey, Liu, and Zhu (2016).

Replicating Anomalies

- According to Hou, Xue and Zhang (2017), of 447 anomalies
  - 286 (64%) due to microcap stocks
  - 380 (85%) with t-stat less than 3.0
  - 436 (98%) insignificant if using 4 factor model Hou, Xue and Zhang (2015)

Table 1 : List of Anomaly Variables

The anomalies are grouped into six categories: (i) momentum; (ii) value-versus-growth; (iii) investment; (iv) profitability; (v) intangibles; and (vi) trading frictions. The number in parenthesis in the title of a panel is the number of anomalies in the category. The total number of anomalies is 447. For each anomaly variable, we list its symbol, brief description, and its academic source. Appendix A details variable definition and portfolio construction.

Panel A: Momentum (57)			
Sue1	Earnings surprise (1-month holding period), Foster, Olsen, and Shevlin (1984)	Sue6	Earnings surprise (6-month holding period), Foster, Olsen, and Shevlin (1984)
Sue12	Earnings surprise (12-month holding period),	Abr1	Cumulative abnormal stock returns around earnings announcements

Factors

- Not all anomalies are factors
- Note different definitions of factors
  - Barra model
  - Statistic factors (PCA factors)
- Risk factors vs. return factors
  - Return factors: persistent long term risk premium
  - Not all risk factors bear risk premium
  - Return factors usually correspond to resources of returns

Multi-Factor Models

- Commonly used benchmark models
  - Fama-French Three factor model (Fama and French 1993)
    - MKT, HML, SMB
  - Carhart Four factor model (Carhart 1997)
    - FF3, MOM
  - Fama-French Five factor model (Fama and French 2015)
    - FF3, CMA (investment), RMW(profitability)
- How to compare factor models?
  - GRS tests (Gibbons, Ross and Shanken 1989)
  - Mean-Variance Spanning tests (Huberman and Kandel 1987)
  - Bayesian approach (Stambaugh and Yuan 2016)

The Law of Parsimony

- Daniel, Hirshleifer and Sun (2018)
- Parsimony Index I = 0 - # of characteristics
- Parsimony Index II = 0 - # of characteristics - # of factors

模型	简称	Parsimony Index I	Parsimony Index II
Fama-French 三因子	FF3	-5	-5
Carhart 四因子	Carhart	-6	-7
Novy-Marx 四因子	NM4	-9	-9
Fama-French 五因子	FF5	-11	-9
Hou-Xue-Zhang 四因子	HXZ	-9	-7
Stambaugh-Yuan 四因子	SY4	-25	-16
Daniel-Hirshleifer-Sun 三因子	BF3	-4	-6

Factor Zoo to Factor War

- Factor zoo (John Cochrane 2011)
  - Which factors are important?
  - Which factors are independent?
  - Why do factors move prices?
- Taming the Factor Zoo (Feng, Giglio, Xiu 2019)
- https://www.aqr.com/About-Us/AQR-Insight-Award/2018/Taming-the-Factor-Zoo
- Factor war
  - Q-Factor vs FF5 (Zhang 2016)
  - Zhang, L. (2016). Factors war. Tsinghua Financial Review, Vol. 37, 101 – 104, in Chinese.

From Manager's View

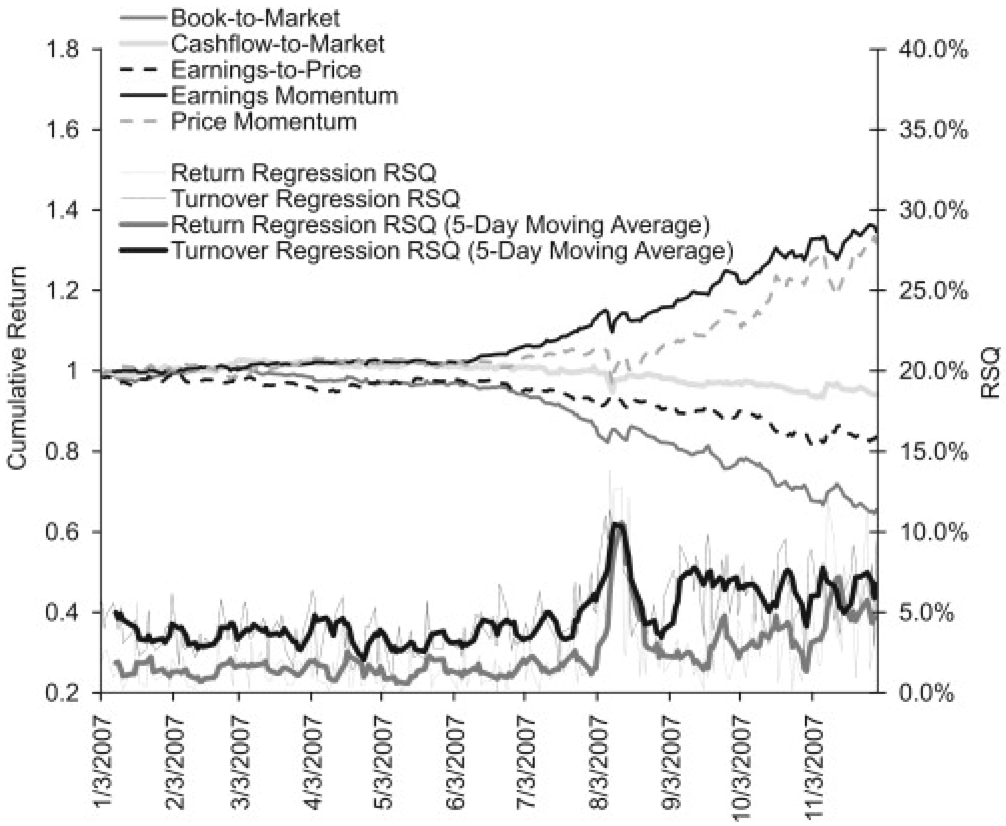
- Objective function is to maximize realized returns under given risk budget within a given investment horizon
- Factor returns are time varying
- When to include or exclude some factors?
- Many more factors are used to model risks
- How to use factors to manage risks, in particular, the tail risks?

Different Investment Horizons

- Long term (years or more): Strategic asset allocation
- Medium term (months or quarters): Tactical asset allocation
- Short term (weeks or days): Trading
- High frequency

Things to Keep in Mind

- Factor crowding
  - Quant meltdown in August 2007 (Khandani and Lo 2011)
  - Flash crash of US stock market on May 6 2010
- Factor timing
  - factor momentum vs. factor valuation
- Factor returns
  - Alpha or smart beta



Things to Keep in Mind

- Factor crowding
  - Quant meltdown in August 2007 (Khandani and Lo 2011)
  - Flash crash of US stock market on May 6 2010
- Factor timing
  - factor momentum vs. factor valuation
- Factor returns
  - Alpha or smart beta

The joint 2010 report detailed how a large mutual fund firm selling an unusually large number of E-Mini S&P contracts first exhausted available buyers, and then how high-frequency traders (HFT) started aggressively selling, accelerating the effect of the mutual fund's selling and contributing to the sharp price declines that day.

Source: Wikipedia

Factor Timing

- The promises and pitfalls of factor timing（Bender et al. 2017）
- Timing 'Smart Beta' Strategies? Of Course! Buy Low, Sell High!（Arnott et al. 2016）
  - Arnott propose to use factor valuation
- Forecasting Factor and Smart Beta Returns (Hint: History Is Worse than Useless)（Arnott et al. 2017）
  - Selecting strategies or factors based on past performance, regardless of the length of the sample, will not help investors earn a superior return and is actually more likely to hurt them. —— Arnott et al. (2017)
- Contrarian Factor Timing is Deceptively Difficult（Asness et al. 2017）
  - At first glance, valuation-based timing of styles appears promising, which is not surprising because it is a simple consequence of the efficacy of the value strategy itself. Yet when the authors implement value timing in a multi-style framework that already includes the value style, they find somewhat disappointing results. Because value timing of factors is correlated to the standard value factor, it adds further value exposure.
- Factor Momentum Everywhere , Tarun Gupta 和 Bryan Kelly, JPM 2019
  - Factor momentum adds significant incremental performance to investment strategies that employ traditional momentum, industry momentum, value, and other commonly studied factors.

Defensive Factor Timing

- BlackRock (Fergis et al. 2019)
- The purpose is to manage tail risks

Macroeconomic Factor Definitions

Factor	Economic Rationale	Factor Mimicking Portfolio	
Economic Growth	Reward for taking exposure to the global economy	Long: Short:	Equity futures, listed real estate (real estate investment trusts), commodities Cash
Real Rates	Reward for taking exposure to the risk of movements in interest rates	Long: Short:	Basket of sovereign inflation-linked bonds Cash
Inflation	Reward for taking exposure to changes in prices	Long: Short:	Basket of nominal sovereign bonds Basket of inflation-linked sovereigns of matching maturity
Credit	Reward for lending to corporations rather than governments	Long: Short:	Investment-grade bonds, high-yield bonds Government bonds
Emerging Markets	Reward for taking exposure to the additional political risk from emerging markets	Long: Short:	Emerging market equity, emerging market debt Developed market equity, developed government bonds
Liquidity	Reward for taking exposure to illiquid assets	Long: Short:	Small-cap equity Large-cap equity, volatility futures

Notes: For illustrative purposes only using hypothetical factors and not representative of an actual investment or account. As of June 30, 2018.

出处：Fergis et al. (2019)