University of Bristol

2020

Replicating Anomalies

Ziyi XI

Supervisor: Yaqiong Yao

Catalogue

01.Introduction

Introduction

Part1 Replication

This paper replicate of the bulk of the published anomalies literature in finance and accounting with 447 anomaly variables.

The variables include six categories:
momentum, value-versus-growth, investment, profitability, intangibles, and trading frictions

The authors use NYSE breakpoints to divide the stocks into three categories: Big, small and micro

The authors treat an anomaly as a replication success if the average return of its high-minus-low decile is significant at the 5% level ($t \geq 1.96$)

Conclusions

The authors find that widespread P-hacking exists in the anomalies literature.

In the replication, out of 447 anomalies, 286 (64%) are insignificant at the 5% level.

e.g. Trading frictions category

95 out of 102 variables (93%) are insignificant

Significant anomalies,

their magnitudes are often much lower than originally reported.

Introduction

Why does the replication differ so much from ordinary studies?

The key word is microcaps

Many studies overweight microcaps with equal-weighted returns.

Microcaps not only have the highest equal-weighted returns, and also largest cross-sectional standard deviations in returns and anomaly variables among microcaps, small stocks, and big stocks.

In the replication, the authors apply value-weighted return of portfolios.

Part 2. q-factor model

Then paper uses q-factor model to explain those 161 significant anomalies.
But there are still 46 significant anomalies can not be explained by q-factor.

02.Motivation

Motivation

Finance:

Data mining is a serious problem in Finance academic.

Researchers tend to try different model and use different data to make their research results significant.

Some researches point out that two publication biases are likely responsible for the high percentage of false discoveries.

The first: it is difficult to publish a negative result in top academic journals.
The second: it is difficult to publish replication studies in finance and economics.

P-hacking problem!

Motivation

Economics:

Replication is hard to success, because the key results will be effected by small specification changes. In addition, different software packages often produce very different estimate.

03.Methodology

Data Source

Monthly returns

CRSP

Accounting information

Compustat Annual and Quarterly Fundamental Files

Sample

from January 1967 to December 2014

Why Portfolio Sorts with NYSE Breakpoints and Value-weighted Returns

NYSE Breakpoints: (breakpoints are used to group stocks)

The cross-sectional dispersion of anomaly variables is the largest among microcaps.

With NYSE-Amex-NASDAQ breakpoints, microcaps account for more than 60% of the stocks in extreme deciles. Therefore, these microcaps can greatly inflate the anomalies.

But using NYSE breakpoints assigns a fair number of small and big stocks into extreme deciles, alleviating the impact of microcaps.

Why Portfolio Sorts with NYSE Breakpoints and Value-weighted Returns

Value-weighted Returns:

1. value-weights accurately reflect the wealth effect experienced by investors
2. microcaps are influential in equal-weighted returns. (3% of market value, 60% of stock numbers)

NYSE Breakpoints:

The cross-sectional dispersion of anomaly variables is the largest among microcaps. With NYSE-Amex-NASDAQ breakpoints, microcaps typically account for more than 60% of the stocks in extreme deciles. Microcaps can greatly inflate anomalies. But in NYSE, the distribution of the small and big stocks in extreme deciles is more reasonable.

What is replication?

1. Pure replication is redoing a prior study in exactly the same way;

2. Statistical replication is the same statistical model but different sample from the same underlying population;

3. Scientific replication is different sample, different population, and similar but not identical statistical model.

Replication results:

Anomalies That Cannot be Replicated

286 out of 447 anomaly variables (64%) earn insignificant average return spreads, including 20, 37, 11, 46, 77, and 95 anomalies from the momentum, value-versus-growth, investment, profitability, intangibles, and trading frictions categories, respectively.

Replicated Anomalies That Are Significant at the 5% Level

Their magnitudes are often much lower than those reported in their original studies.

For example, abnormal returns around earnings announcements (Abr)
Replication:0.3%, 0.22% per month across 6 and 12 months; Original study: 5.9%, 8.3%

Other replication results:

Average Return Spreads in the Original Samples

Using same sample to test anomalies, the authors found that only 36% of anomalies can be replicated, which means sampling variation plays a limited role.

Average Return Spreads with NYSE-Amex-NASDAQ Breakpoints and Equal- weighted Returns

181 (40%) are insignificant at the 5% level , therefore overweighting microcaps can inflate anomalies.

Q-factor model:

$$r_t^i - r_t^f = \alpha^j + \beta_{MKT}^i \left(r_{Mt} - r_{ft} \right) + \beta_{ME}^i r_{ME,t} + \beta_{I/A}^i r_{I/A,t} + \beta_{ROE}^i r_{ROE,t} + e_t^i$$

The model says that the expected return of an asset in excess of the risk-free rate is described by its sensitivities to the market factor, a size factor, an investment factor, and a return on equity factor.

Methodology

Replication

q-factor model

momentum anomalies

The Roe factor is the main source of the model’s performance

value-versus-growth anomalies

The investment factor is the main source of the model’s performance

profitability anomalies

The Roe factor is the main source of the model’s performance

investment anomalies

The investment factor is the main source of the model’s performance

The weakness of the q-factor model:

There are still many anomalies that the q-factor model cannot explain.

For example:

9 momentum anomalies
6 value-versus-growth anomalies
7 investment anomalies
9 profitability anomalies
11 intangible anomalies
seasonality anomalies
4 friction anomalies

04.Results

Results

Contribution:

Replication indicates widespread p-hacking, mainly by overweighting microcaps.

The authors recommend authors, referees, and editors should be keenly aware of the complex agency problem that can arise from financial conflicts of interest and publication biases.

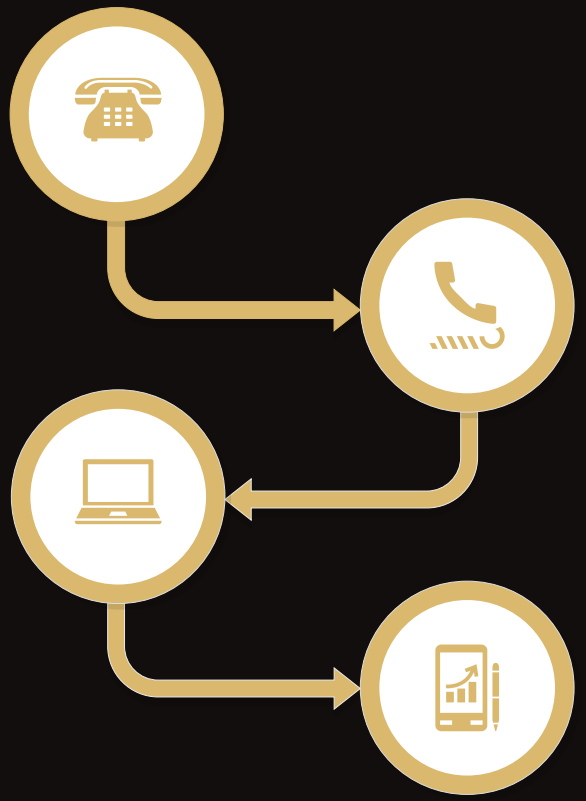
On the sample size, using global data instead of U.S.-centric CRSP-Compustat data can improve the quality of the anomalies literature.

05.Improvements

Improvements

Using global data can avoid over

data mining



Adding a factor into q-factor model:

Expected growth factor

University of Bristol

2020

Thank you

Ziyi XI

Supervisor: Yaqiong Yao