Essays on Empirical Asset Pricing

By

Yixin Chen

Master of Finance
Massachusetts Institute of Technology, 2012

SUBMITTED TO THE SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN MANAGEMENT

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2018

Signature redacted

Signature of Author: ____________________
Department of Management
April 17, 2018

Signature redacted

Certified by: ____________________
Leonid Kogan
Nippon Telegraph and Telephone Professor of Management
Thesis Supervisor

Signature redacted

Certified by: ____________________
Jonathan A. Parker
Robert C. Merton (1970) Professor of Finance
Thesis Supervisor

Signature redacted

Accepted by: ____________________
Catherine Tucker
Sloan Distinguished Professor of Management
Professor of Marketing
Chair, MIT Sloan PhD Program

Essays on Empirical Asset Pricing

by

Yixin Chen

Submitted to the Alfred P. Sloan School of Management
on May 1, 2018, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This dissertation consists of three chapters.

Chapter 1 shows that, for active mutual funds, historical in-sample alpha is a poor predictor of out-of-sample alpha. However, by focusing on a subset of skilled managers who are able to generate positive alpha via profitable bets on firm specific risks (stock-picking), I show that a new first-order stochastic dominance (FSD) condition can be employed as an additional search criterion to identify such skilled stock-pickers. I implement an FSD filter to select funds by bootstrapping the return distribution in a given period associated with a random stock-picking strategy that has a given factor exposure and degree of diversification. Simulations show that the identification of funds as skilled by the FSD filter performs well in finite samples, in the face of heteroscedasticity and benchmark mis-specification. With the new FSD filter, I identify a group of active funds that are able to outperform the Carhart benchmark by 2.04% (t=2.78) per year before fees (0.78% (t=1.07) per year after fees) out of sample. Moreover, in this sample of funds, in-sample alpha is significantly predictive of out-of-sample alpha: the top quintile of stock-picking mutual funds deliver an out-of-sample alpha of 3.55% (t=3.24) per year before fees (2.24% (t=2.05) per year after fees). These outperforming funds tend to be more aggressive stock-pickers (hold more concentrated portfolios), charge higher fees, and attract more fund flows.

By exploring mutual fund managers' Herding tendency and Trading Intensity, Chapter 2 develops a systematic approach to identify mutual fund managers with the Warren Buffett style, i.e. managers who are fundamental, long-term, value investors. Using data during 1995-2015, I further show that the group of such managers outperformed the Carhart four-factor benchmark by 3.06% (t = 3.58) per year before fees (1.94% (t = 2.35) per year after fees). Moreover, these managers have both statistically and economically high exposures to AQR's Quality Minus Junk (QMJ) factor. Last but not least, I show that their before-fees performances can be almost perfectly replicated by an investor who implements the strategy of investing in the lagged portfolio holdings of these managers when they become publicly available.

Chapter 3 proposes a methodology to recover countries' stochastic discount factors (SDFs) from exchange rates under three assumptions: 1) the Euler equation holds internationally; 2) there is a factor structure among exchange rates; 3) there does not exist a special global risk factor which has identical influence on all countries. By designing an empirical test using exchange rates and equity returns of 28 countries from 1988 to August of 2014, I show that the moment conditions are rejected in the data. The failure of the exchange-rate-recovered SDFs to price countries' assets reflects the violation of my assumptions, and highlights the importance of the special global risk factor to price assets in different countries.

Thesis Supervisor: Leonid Kogan
Title: Nippon Telegraph and Telephone Professor of Management

Thesis Supervisor: Jonathan A. Parker
Title: Robert C. Merton (1970) Professor of Finance

Acknowledgments

Words cannot express my gratitude towards my advisors Hui Chen, Leonid Kogan and Jonathan Parker. I met with Leonid and Hui early on when I was still a Master of Finance student. At that time, I barely knew anything about finance. But Leonid and Hui spotted my potential and passion for academic research and took me into the prestigious finance PhD program which utterly changed my life trajectory. I met with Jonathan later on, after he joined Sloan in 2013. Shortly after his arrival, Jonathan became a popular choice of advisor among PhD students because of his brilliance, his encyclopedic knowledge and his strong willingness to work with and help students. Throughout my years in the PhD program, my advisors endowed me with their continuous support. They encouraged me to explore new ideas. They taught me how to deal with the job market. And their comments and critiques gradually shaped my research taste and helped me develop the distinctive MIT rigor. For that, I'll be eternally grateful. Unfortunately, I don't see any possibility that I can repay my debt to them in the future. The only thing that I can try is to mimic them and pass on their generosity and kindness to the people that I will encounter in my future career.

During my years at MIT Sloan, I also benefitted tremendously from my interactions with other distinguished scholars and professors, including: Randy Cohen, Dan Greenwald, Andy Lo, Debbie Lucas, Andrey Malenko, Paul Mende, Chris Palmer, Jun Pan, Steve Ross, Antoinette Schoar, Eric So, Adrien Verdelhan, Jiang Wang and Haoxiang Zhu. It has been a great privilege to learn from them.

I also thank my fellow graduate students, especially Victor Duarte, Hyungjune Kang, Seung Ki Kwak, Fangzhou Lu, Anton Petukhov, Yu Shi, Jian Sun and Yupeng Wang. We shared food and work, as well as lots of stories and laughters.

Last but not the least, I owe my utmost thanks to my parents and my grandparents for their unconditional support. Life is stochastic. Circumstances change. But they are always around by my side.

Contents

List of Tables

List of Figures

Chapter 1

Individual Stock-picking Skills in Active Mutual Funds

1.1 Introduction

It is well-known that skilled active mutual fund managers who can predictably out-perform passive benchmarks are difficult to identify.[1] Current evaluation methods measure a fund manager's skills by comparing the time average of his/her returns[2] to that of an appropriately chosen benchmark. For example, fund manager i's skills can be measured as:

$$\hat{\alpha}_i = \frac{1}{T}\sum_t \left(r_{i,t} - r_{i,t}^b\right),$$

where $\{r_{i,t}\}_{t=1}^T$ are fund i's realized returns, $\left\{r_{i,t}^b\right\}_{t=1}^T$ are the returns of a passive benchmark representing fund i's exposure to systematic factors.

While this in-sample alpha is straightforward to construct, it has poor predictive power of the manager's out-of-sample alpha within the active mutual fund industry. Table 1.1 replicates the Carhart (1997) regression during a recent sample period. According to the table, none of the post-ranking alphas of the five in-sample-alpha-ranked portfolios of funds is statistically significant at 5% confidence level. The predictive power of the in-sample Carhart four-factor alpha for future fund performance is weak and statistically insignificant.

[1]See, for example, Carhart (1997), Kosowski et al. (2006), Barras, Scaillet, and Wermers (2010), Fama and French (2010), etc.

[2]All performances are before fees unless specified otherwise.

Quintile	α (in %)	mkt	smb	hml	umd
1	−0.83	1.02	0.28	0.06	0.00
	[−1.22]	[61.56]	[11.00]	[2.51]	[0.17]
2	−0.02	0.98	0.15	0.06	−0.01
	[−0.03]	[71.07]	[7.02]	[2.86]	[−0.40]
3	−0.13	0.99	0.13	0.06	0.00
	[−0.26]	[79.50]	[5.76]	[2.75]	[0.27]
4	0.33	1.00	0.18	0.03	0.01
	[0.63]	[72.70]	[7.92]	[1.26]	[0.84]
5	1.13*	1.02	0.31	−0.08	0.03
	[1.74]	[59.76]	[11.89]	[−2.85]	[1.54]

Table 1.1: The Weak Persistence of Alpha

This table documents the before-fees out-of-sample performance of the trading strategy that sorts funds by their historical before-fees alphas. By the end of each quarter, the mutual funds in the cross section are sorted into five quintiles based on the four-factor alphas computed from their proceeding 24 months' before-fees returns. The trading strategy is rebalanced every three months. The post-ranking annualized before-fees alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. The alphas with statistical significance are marked with "*". The sample period is from January 1991 to December 2015.

The search for skilled managers based on in-sample alpha has poor out-of-sample performance because it has little power to distinguish skill from luck given the short fund performance histories.[3] The conventional positive alpha ($\alpha_i > 0$) condition requires the mean of the fund's return to be higher than the mean of the benchmark return, i.e. $\mathbb{E}\left(r_{i,t}\right) > \mathbb{E}\left(r^b_{i,t}\right)$. Thus, it suffers from the empirical problem that mean is difficult to estimate in finite sample, and it is relatively easy for unskilled managers to achieve high in-sample alphas with excessive risk-taking or unobservable factor exposures[4].

In this paper, I focus on a subset of skilled fund managers who generate positive alpha by making profitable bets on firm specific risks (stock-picking), and show that a new first-order stochastic dominance (FSD) condition can be imposed to identify such skilled stock-pickers. The new FSD condition states that the return distribution of a skilled stock-picker should first-order stochastically dominate that of a manager

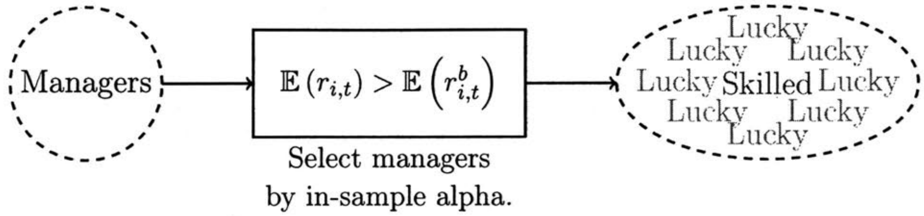
[3] See Kosowski et al. (2006), Fama and French (2010) and Barras, Scaillet, and Wermers (2010) for ex post analysis.

[4] See the empirical evidence documented by Chevalier and Ellison (1997) and a more recent structural estimation by Koijen (2014).

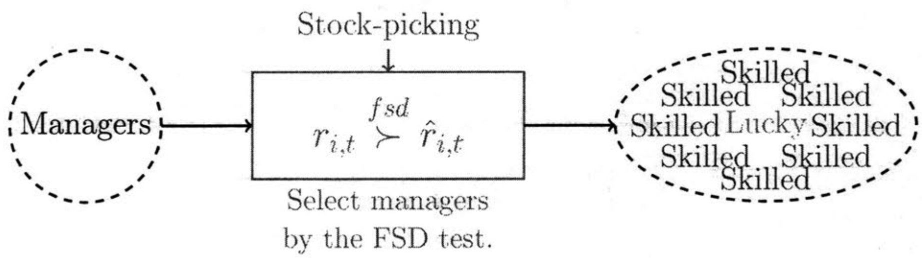
with similar investment style but no stock-picking skills, i.e. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$. The FSD condition is a more stringent requirement than conditions focused on the mean alone such as $\alpha_i > 0$, because it uses information from the entire distribution of returns and excludes funds with heavy left tails in their return distributions due to excessive risk-taking or unobservable factor exposures.

As a methodological contribution of this paper, I construct an FSD filter to select funds by testing the FSD condition. The test requires extending the benchmark from a **single return** to a **return distribution in each period**. Specifically, for each fund in each period, I construct a counterfactual return distribution from a bootstrap exercise by creating replica funds with random portfolios. The replica funds maintain the same portfolio weights as the original fund and invest in stocks with similar observable characteristics, so that the replica funds resemble the original fund in the degree of diversification and loadings on observable factors. However, the specific choices of stocks in a replica fund's portfolio is determined randomly. As a result, the replica funds emulate the return distribution of the original fund, meanwhile break the association between the portfolio weights and the stock choices, which reflects the stock-picking skills of the fund manager. The comparison between the original fund's return and the replica funds' return distribution then enables the econometrician to conduct a statistical test on the manager's stock-picking skills in each single period with only one observation. With repeated observations over time, the FSD filter selects funds by requiring the percentiles of the original fund's returns among the replica funds to first-order stochastically dominate a standard uniform distribution.

To understand the source of the additional statistical power that the new FSD condition is able to provide, I conduct simulations to compare the performance of a filter based on the conventional positive alpha condition with the performance of the FSD filter, as illustrated in Figure 1.1. Simulations show that the FSD filter outperforms the positive alpha filter in handling two statistical problems in finite sample – heteroscedasticity and benchmark mis-specification.The heteroscedasticity problem is defined as idiosyncratic volatility being time-varying and more volatile than the fund's true alpha. I show that the FSD filter has better performance than the positive alpha filter with heteroscedasticity calibrated to the real-world level because the positive alpha filter places equal weight on all observations regardless of idiosyncratic volatility; whereas the FSD filter places relatively higher(lower) weights on observations from high(low) signal-to-noise ratio periods. The benchmark



(a) Positive Alpha Filter



(b) FSD Filter

Figure 1.1: Positive Alpha Filter VS FSD Filter

This figure illustrates the comparison between the positive alpha filter and the FSD filter. Panel (a) illustrates the ineffectiveness of the positive alpha filter in identifying skilled mutual fund managers; whereas Panel (b) proposes the installation of the new FSD filter. The FSD filter selects managers whose return distributions first-order stochastically dominate managers with similar investment styles but no stock-picking skills, i.e. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$.

mis-specification problem is defined as the situation that some managers might take on factors that are not observable to the econometrician. The performance of the positive alpha filter suffers due to the additional noise from the unobservable factors. The positive alpha filter tends to erroneously select mis-specified managers who take on unobservable factors with high in-sample realizations rather than the truly skilled managers who are able to deliver positive out-of-sample alphas. The FSD filter, on the other hand, is unaffected by this problem thanks to a detection mechanism. The unobservable factors taken by the mis-specified managers induce heavier left tails in their return distributions compared to the replica funds thereby violating the FSD condition. The two-stage sort using both FSD and in-sample alpha yields similar patterns in simulation as the same sort with the actual data, suggesting that these two statistical problems might be at play in the real world.

In the empirical part of this paper, I show that the FSD filter is indeed effective in selecting skilled stock-pickers. From January 1991 to December 2015, the FSD filter identifies a time-varying group of active mutual funds that are able to, on average, outperform the Carhart four-factor benchmark by 204 bps ($t = 2.78$) per year out of sample before management fees (78 bps ($t = 1.07$) per year after fees). More interestingly, among the funds selected by the FSD filter, in-sample alpha is significantly predictive of out-of-sample alpha. The combination of the FSD filter and the standard $\hat{\alpha}$ sort is especially powerful in identifying skilled stock-pickers, with the top quintile of funds in the second-stage sort by in-sample alpha outperforming the Carhart benchmark by as high as 355 bps ($t = 3.24$) per year before management fees (224 bps ($t = 2.05$) per year after fees). By investigating the fund return gaps[5], I further verify that about 50% of the Carhart four-factor alphas of the outperforming funds are resulted from profitable unobserved within-quarter trades. The finding lends additional support to the empirical success of the evaluation method in identifying skilled stock-pickers and is consistent with the view that profitable information is usually short-lived in a stock market that is largely liquid and efficient.

The investigation into the observable characteristics of the outperforming funds also produces interesting findings. The identified outperforming funds manifest characteristics that are distinctive from an average fund in the industry in the following aspects:

[5]See Kacperczyk, Sialm, and Zheng (2008).

1. They have the same size as an average fund measured as asset under management, but they are able to charge higher fees.

2. They keep fewer stocks within their portfolios.

3. Controlling for realized in-sample alphas, funds that satisfy the FSD condition attract more flows.

The finding with the fee setting is partially consistent with the prediction by Berk and Green (2004) in the sense that more successful funds are able to extract higher rents, but inconsistent in the specific mechanism. In Berk and Green (2004), skilled managers demand compensation by growing the size of their funds meanwhile keeping the fees fixed. My finding, on the other hand, suggests that the outperforming managers are able to charge higher fees directly rather than growing the size of their funds. The finding with portfolio concentration is consistent with the theory by Van Nieuwerburgh and Veldkamp (2010). My finding verifies their prediction that informed investors could voluntarily keep under-diversified portfolios in order to become specialized when information acquisition is costly. The finding with the fund flows echoes the empirical work by Berk and van Binsbergen (2016) and Barber, Huang, and Odean (2016). Based on their arguments that fund flows reflect investors' evaluations of managers' skills, I show that investors infer the quality of funds from the properties of their return distributions that are beyond the mean or alpha. Yet, the positive out-of-sample alphas of the identified funds also suggest that the magnitudes of the fund flows are still insufficient to fully arbitrage away all the outperformances according to the logic proposed by Berk and Green (2004).

The remaining of this article is organized as the following. Section 2 offers a review of the related literature. Section 3 provides details about the bootstrap exercise to construct the counterfactual return distribution in each single period. Section 4 develops a factor model to clarify the definition of stock-picking skills and the sufficient conditions required to impose the FSD condition. Section 5 describes the construction of the FSD filter using the counterfactual return distribution. Section 6 includes theoretical proofs and simulation exercises to illustrate the advantageous econometric properties of the FSD condition over the positive alpha condition. Section 7 describes the data and documents the empirical findings. Section 8 concludes.

1.2 Related Literature

Systematic academic research on the active mutual fund industry dates back to, at least, Treynor and Mazuy (1966) and Sharpe (1966). Early empirical work such as Jensen (1968) and Malkiel (1995) establish that active mutual funds, on average, cannot outperform the market index before fees, and significantly under-perform the passive index after fees. The findings are largely consistent with the efficient market hypothesis proposed by Malkiel and Fama (1970). Despite the mediocre average performance, researchers also investigate whether historical fund performances can be used to select funds that are able to deliver superior returns in the future. Early work by Hendricks, Patel, and Zeckhauser (1993), Goetzmann and Ibbotson (1994) and Brown and Goetzmann (1995) document the "hot-hand" effect that funds with good performances in the past also tend to outperform their peers going forward. However, the classic paper by Carhart (1997) demonstrates that much of the "hot-hand" effect can be attributed to the momentum of stock prices discovered by Jegadeesh and Titman (1993), and fund performance does not seem to persist once momentum is adjusted for.

Since the seminal paper by Wermers (2000), researchers started to use survivorship-bias-free holdings data to better characterize fund styles and identify different types of investment skills. The influential paper by Daniel et al. (1997) divides the universe of stocks into size, value and momentum buckets, and characterizes funds' styles by their portfolio weights in different stock buckets. Moreover, they show that there is persistent stock-picking skills, but no significant market-timing skills in their sample period. Later papers then discover that various holdings characteristics can be used to infer the skills of fund managers and predict their future performances. For example, Cohen, Coval, and Pástor (2005) find that funds that have overlapping holdings with past successful funds tend to outperform others going forward; Kacperczyk, Sialm, and Zheng (2008) find that funds with profitable unobserved actions are also likely to generate trading profits in the future; Cremers and Petajisto (2009) show that funds with more active weights outperform the ones that are suspected to be closet indices. Other notable examples include: Grinblatt and Titman (1989), Grinblatt, Titman, and Wermers (1995), Chen, Jegadeesh, and Wermers (2000), Kacperczyk, Sialm, and Zheng (2005), Alexander, Cici, and Gibson (2006), Jiang, Yao, and Yu (2007), Kacperczyk and Seru (2007), Baker et al. (2010), Da, Gao, and Jagannathan (2010), Huang, Sialm, and Zhang (2011), Kacperczyk, van Nieuwerburgh, and Veldkamp (2014), Agarwal et al. (2015), etc. This paper contributes to this line of literature by proposing a general test on the manager's information advantage regarding the idiosyncratic risks of the securities he/she keeps in the portfolio. In the same spirit, Iskoz and Wang (2003) also propose a methodology to test whether a money manager incorporates private information in portfolio construction by investigating the connections between fund holdings and return distributions. The difference between this paper and their work is that they consider the relation between general types of private information and future stock return distributions; whereas this paper focuses on a particular type of private information on firm-specific risks and imposes a detailed restriction on fund return distributions – the FSD condition.

The relation between luck and skill in the context of the active mutual fund industry was first formally addressed by Kosowski et al. (2006). Their paper proposes a bootstrap exercise in the time-series to verify the existence of fund skills ex post. Fama and French (2010) employs a similar methodology and verifies the results of Kosowski et al. (2006) in a more recent sample period. Barras, Scaillet, and Wermers (2010) classifies funds into three categories: unskilled, zero-alpha and skilled, by implementing a novel statistical procedure to account for false discoveries. This paper is closely related to these three papers in that it also aims to account for luck in fund managers' performances. However, this paper contributes to this line of research in two important ways. First, existing papers can only test fund skills in the time-series with repeated observations; whereas this paper shows that a statistical test on stock-picking skills can be conducted with even only one observation by carrying out a bootstrap procedure in the cross section rather than in the time series. Second, existing papers can only ex post identify funds whose realized alphas are unlikely to be explained by luck; whereas this paper is able to ex ante identify fund skills so that profitable trading strategies can be formed.

My findings on the observable characteristics of the outperforming funds are related to a number of earlier findings in the literature. The finding that the outperforming funds have the same size as the industry average but are able to charge higher fees is partially consistent with Berk and Green (2004) that skilled managers are able to extract higher rents from fund investors, although the specific mechanism is different. The finding that a large portion of the alphas of the identified outperforming funds are due to their unobserved within-quarter trades is consistent with Kacperczyk, Sialm, and Zheng (2008), where they show that return gap is indicative of the skills of a fund manager. The finding that the outperforming funds tend to keep more concentrated portfolios verifies the theoretical prediction by Van Nieuwerburgh and Veldkamp (2010) that informed investors can voluntarily choose to become under-diversified when information acquisition is endogenous. The finding that among the funds selected by the FSD filter, the ones with larger alphas also have more trading is related to Pástor, Stambaugh, and Taylor (2017), where they show that skilled managers are able to make more profits when they trade more. Finally, the finding that controlling for realized alpha, funds that satisfy the FSD condition attract more flows than others echoes the work by Barber, Huang, and Odean (2016) and Berk and van Binsbergen (2016), where they argue that fund flows reflect investors' evaluations on fund managers' skills.

1.3 Benchmark Extension

Under current evaluation methods, the return of a fund $r_{i,t}$ is compared to a benchmark return $r^b_{i,t}$ in every period. The single-period fund outperformance is then computed as the difference between these two returns, $r_{i,t} - r^b_{i,t}$. This approach provides a point estimate of the fund's outperformance in this single period, yet offers no information about its statistical significance. The extension from the benchmark return $r^b_{i,t}$ to the counterfactual return distribution $\langle \hat{r}_{i,t} \rangle$ allows the econometrician to obtain both the point estimate and the statistical significance of the single-period fund outperformance due to stock-picking by comparing $r_{i,t}$ to $\langle \hat{r}_{i,t} \rangle$. The additional distributional information can then be used to implement the FSD condition as elaborated in Section 5.

The construction of the counterfactual return distribution is a bootstrap exercise that mimics the fund's portfolio by investing in stocks of similar characteristics with the same portfolio weights meanwhile randomizes the specific stock choices. Specifically, the procedure can be summarized with the following 4 steps:

1. Retrieve the most recent portfolio of the fund that is available.

2. Create a replica portfolio by replacing each stock in the original portfolio with a new stock[6] of similar characteristics that is randomly chosen, meanwhile keeping the portfolio weights unchanged.

3. Compute the hypothetical return of the replica portfolio by taking the inner product between the portfolio weights and the returns of the replaced stocks.

4. Repeat Step 2 and Step 3 to generate a distribution of hypothetical portfolio returns.

In order to find stocks of similar characteristics with a given stock as required in Step 2, I follow and extend the approach proposed by Daniel et al. (1997) and Wermers (2003). For US stocks that are traded on AMEX, NYSE and Nasdaq, I first sort them into 5 size buckets by their market capitalization.[7] Within each size bucket, I further divide the stocks into 5 value buckets by their book-to-market ratio. Then I repeat the same procedure and divide the stocks within each value bucket into 5 momentum buckets by their preceding one-year return. Lastly, I divide the stocks in each momentum bucket further into 5 volatility buckets by their return volatility. The procedure thus categorizes all stocks into $5 \times 5 \times 5 \times 5 = 625$ non-overlapping buckets and is repeated once in a year by the end of June. For each stock within the original portfolio, Step 2 is carried out by finding a random replacing stock within the same bucket as the original stock. The weights of the replica portfolio are kept unchanged as the original portfolio. Table 1.2 offers an example to illustrate the bootstrap procedure. Panel A is a snapshot of the portfolio of Longleaf Partners Fund by the end of 2012/12. Panel B is a simulated replica portfolio.

In Step 3, the hypothetical return of fund i's replica in period t is computed as:

$$\hat{r}_{i,t} = \sum_{j} w_{i,j,t-1} \cdot \tilde{r}_{\hat{j},t}$$

where $w_{i,j,t-1}$ denotes the portfolio weight of stock j within the portfolio of fund i by the end of period $t-1$; $\hat{j}$ denotes the random replacement of stock j in the replica portfolio; $\tilde{r}_{\hat{j},t}$ denotes the return of the replacing stock in period t.

A few comments are in order regarding the bootstrap procedure. First, by controlling for stock characteristics in the replica portfolios, I ignore the fact that the

[6]The new stock can be the same as the original stock.

[7]The breakpoints of the size buckets are defined by NYSE stocks only.

Panel A: Real Fund			Panel B: Replica Fund		
Stock Ticker	Bucket No.	Weight(%)	Stock Ticker	Bucket No.	Weight(%)
ABT	597	5.24	MO	597	5.24
BEN	552	4.68	DOV	552	4.68
BK	576	6.73	SYK	576	6.73
BRK	561	4.65	PG	561	4.65
CHK	603	8.07	C	603	8.07
CNX	428	7.16	BBY	428	7.16
DELL	529	5.55	A	529	5.55
DIS	598	5.70	DIS	598	5.70
DTV	502	8.19	WU	502	8.19
FDX	534	8.00	FDX	534	8.00
L	581	10.00	MDT	581	10.00
LVLT	405	6.15	LVLT	405	6.15
MDLZ	592	5.39	MDLZ	592	5.39
PHG	356	1.26	PHG	356	1.26
TRV	568	6.65	FISV	568	6.65
VMC	489	6.58	ARW	489	6.58

Table 1.2: Bootstrap Example

This table offers an example in order to illustrate the bootstrap procedure that constructs replica portfolios and the counterfactual return distribution. Panel A is a snapshot of the portfolio of Longleaf Partners Fund by the end of 2012/12. Panel B is a simulated replica portfolio. The replica portfolio is created by replacing each stock in the real portfolio with another stock that is randomly chosen in the same bucket as the original stock. The portfolio weights of the replica fund are identical as the real fund.

choice of stock characteristics might also reflect the fund manager's skills. In other words, by comparing the real fund with replica funds of similar factor loadings and degree of diversification, the bootstrap exercise only measures the stock-picking skills of the manager and is silent about potential factor-timing skills[8]. As discussed in Daniel et al. (1997), stock-picking skills seem to be more prevalent among successful mutual fund managers. Secondly, I extend the Daniel et al. (1997) stock classification to include volatility as an additional dimension. The matching of volatility serves two purposes. On the one hand, recent literature has documented that stock-level volatility might represent systematic risks that are priced in the cross section of stocks.[9] On the other hand, I match the stock-level idiosyncratic volatility in the replica funds to the real fund so that the portfolio-level idiosyncratic volatility of the replica funds would also be comparable to that of the real fund. The purpose of the matching of idiosyncratic volatility will be further discussed in Section 4. Thirdly, the holdings information of the real fund is only employed to extract the weight distribution and stock characteristics of the fund's investment. The specific choices of stocks in the real portfolio are not used. Therefore, even though the holdings information is only empirically available at quarterly frequency, the counterfactual return distribution can be constructed at much higher frequencies, such as monthly or daily frequencies, by interpolating portfolio characteristics. Finally, the extension from a single benchmark return to the counterfactual return distribution extracts additional information about the distribution from the data so that a statistical test on stock-picking skills can be formed in every period. Outperformance computed under current evaluation methods as the difference between a fund's return and the benchmark return can be regarded as a point estimator of the manager's skills in a single period; whereas the comparison between the fund's return and the counterfactual return distribution provides both the point estimate and the statistical significance of the manager's stock-picking skills in each period.

[8]See Grinblatt and Titman (1989) and Daniel et al. (1997) for the definition of the two types of investing skills

[9]See, for example, Ang et al. (2009), Fu (2009), etc.

1.4 The Factor Model

I propose a factor model in this section to formalize the definition of stock-picking skills. The model serves to clarify the specific assumptions required to establish the FSD condition. The analysis shows that one set of sufficient conditions to impose the FSD condition is to limit the search scope to fund managers who are: 1) skilled at stock-picking; 2) unbiased towards unobservable factors; 3) sufficiently diversified.

1.4.1 The Economy

The economy considered in this section is a frictionless financial market with a factor structure. There are J factors(denoted as $\{F_{j,t}\}_{j=1}^{J}$) that are observable to both fund managers and the econometrician; and L factors(denoted as $\{f_{l,t}\}_{l=1}^{L}$) that are only observable to fund managers but not to the econometrician. There are K stocks traded in the market. The excess return (relative to the risk-free rate) of any stock within this market can be decomposed into three parts: the exposure to the J observable factors, the exposure to the L unobservable factors, and the idiosyncratic component:

$$\tilde{r}_{k,t} = r_f + \sum_j \beta_{k,j} F_{j,t} + \sum_l \gamma_{k,l} f_{l,t} + \epsilon_{k,t}$$

where $\tilde{r}_{k,t}$ denotes the return of stock k at time t; $F_{j,t}(f_{l,t})$ is the realization of the observable(unobservable) factor $j(i)$ at time t; $\beta_{k,j}(\gamma_{k,l})$ denotes of the loading of stock k on factor $j(l)$; $\epsilon_{k,t}$ is the idiosyncratic shock in stock k's return.

The factors and the idiosyncratic shocks represent different sources of risks and are assumed to be mutually independent. For an economic interpretation, the factors $\{F_{j,t}\}_{j=1}^{J}$ and $\{f_{l,t}\}_{l=1}^{L}$ can be regarded as $J+L$ different types of market-wide risks; whereas the idiosyncratic shocks $\{\epsilon_{k,t}\}_{k=1}^{K}$ represent firm-specific risks.

Assumption 1. *The factors $\{F_{j,t}\}_{j=1}^{J}$ and $\{f_{l,t}\}_{l=1}^{L}$, and the idiosyncratic shocks $\{\epsilon_{k,t}\}_{k=1}^{K}$ are mutually independent. That is $\forall j,\ j'\ \ F_{j,t} \perp F_{j',t}$, $\forall l,\ l'\ \ f_{l,t} \perp f_{l',t}$, $\forall k,\ k'\ \ \epsilon_{k,t} \perp \epsilon_{k',t}$, $\forall j,\ l\ \ F_{j,t} \perp f_{l,t}$, $\forall j,\ k\ \ F_{j,t} \perp \epsilon_{k,t}$, $\forall l,\ k\ \ f_{l,t} \perp \epsilon_{k,t}$, where $\perp$ denotes that two random variables are independent to each other.*

Regarding the idiosyncratic shocks, they have zero expectation under the econometrician's information set so that there is no asymptotic arbitrage in this economy according to Ross (1976).

Assumption 2. *Idiosyncratic shocks have zero expectation under the econometrician's information set:* $\forall k, \quad \mathbb{E}(\epsilon_{k,t}) = 0$.

1.4.2 Connection between Real and Replica Fund Returns

For each real fund in each period, replica portfolios are constructed according to the procedure in Section 3. During the construction, each stock within the real fund's portfolio is replaced randomly with another stock in the same bucket. In this economy, the buckets are defined by the econometrician according to the J observable factors.

Assumption 3. *For each stock within the original portfolio, its replacement in the replica portfolio has the same exposure to observable factors, i.e.*

$$\forall k,\ j \quad \beta_{k,j} = \beta_{\hat{k},j}$$

where $\beta_{k,j}$ is stock k's exposure to observable factor j; $\hat{k}$ labels the replacing stock of stock k; $\beta_{\hat{k},j}$ is the exposure to observable factor j of stock k's replacing stock.

For expositional clarity, I adopt the following notation. I denote the return of fund i in period t as

$$\begin{aligned}
r_{i,t} &= \sum_k w_{i,k,t-1}\tilde{r}_{k,t} \\
&= \sum_k w_{i,k,t-1}\left(r_f + \sum_j \beta_{k,j}F_{j,t} + \sum_l \gamma_{k.l}f_{l,t} + \epsilon_{k,t}\right) \\
&= r_f + \sum_k w_{i,k,t-1}\left(\sum_j \beta_{k,j}F_{j,t}\right) + \sum_k w_{i,k,t-1}\left(\sum_l \gamma_{k.l}f_{l,t}\right) + \sum_k w_{i,k,t-1}\epsilon_{k,t} \\
&= r_f + \sum_j \left(\sum_k w_{i,k,t-1}\beta_{k,j}\right)F_{j,t} + \sum_l \left(\sum_k w_{i,k,t-1}\gamma_{k,l}\right)f_{l,t} + \sum_k w_{i,k,t-1}\epsilon_{k,t} \\
&\equiv r_f + \sum_j \beta_{i,j,t}F_{j,t} + \underbrace{\underbrace{\sum_l \gamma_{i,l,t}f_{l,t}}_{v_{i,t}} + \underbrace{\sum_k w_{i,k,t-1}\epsilon_{k,t}}_{e_{i,t}}}_{u_{i,t}} \\
&\equiv r_f + \sum_j \beta_{i,j,t}F_{j,t} + u_{i,t}.
\end{aligned}$$

Likewise, the return of fund i's replica is written as

$$
\begin{aligned}
\hat{r}_{i,t} &= \sum_k w_{i,k,t-1} \tilde{r}_{\hat{k},t} \\
&= \sum_k w_{i,k,t-1} \left(r_f + \sum_j \beta_{\hat{k},j} F_{j,t} + \sum_l \gamma_{\hat{k},l} f_{l,t} + \epsilon_{\hat{k},t} \right) \\
&= r_f + \sum_k w_{i,k,t-1} \left(\sum_j \beta_{\hat{k},j} F_{j,t} \right) + \sum_k w_{i,k,t-1} \left(\sum_l \gamma_{\hat{k},l} f_{l,t} \right) + \sum_k w_{i,k,t-1} \epsilon_{\hat{k},t} \\
&= r_f + \sum_j \left(\sum_k w_{i,k,t-1} \beta_{\hat{k},j} \right) F_{j,t} + \sum_l \left(\sum_k w_{i,k,t-1} \gamma_{\hat{k},l} \right) f_{l,t} + \sum_k w_{i,k,t-1} \epsilon_{\hat{k},t} \\
&\equiv r_f + \sum_j \hat{\beta}_{i,j,t} F_{j,t} + \underbrace{\underbrace{\sum_l \hat{\gamma}_{i,l,t} f_{l,t}}_{\hat{v}_{i,t}} + \underbrace{\sum_k w_{i,k,t-1} \epsilon_{\hat{k},t}}_{\hat{e}_{i,t}}}_{\hat{u}_{i,t}} \\
&\equiv r_f + \sum_j \hat{\beta}_{i,j,t} F_{j,t} + \hat{u}_{i,t}
\end{aligned}
$$

The excess return of a real(replica) fund can be decomposed into three parts: the exposure to the observable factors $\sum_j \beta_{i,j,t} F_{j,t} (\sum_j \hat{\beta}_{i,j,t} F_{j,t})$, the exposure to the unobservable factors $v_{i,t} \equiv \sum_l \gamma_{i,l,t} f_{l,t} (\hat{v}_{i,t} \equiv \sum_l \hat{\gamma}_{i,l,t} f_{l,t})$, and the exposure to the idiosyncratic shocks $e_{i,t} \equiv \sum_k w_{i,k,t-1} \epsilon_{k,t} (\hat{e}_{i,t} \equiv \sum_k w_{i,k,t-1} \epsilon_{\hat{k},t})$.

Note that fund i and its replica have the same portfolio weights by construction. A direct outcome of this and Assumption 3 is that the real fund and the replica have the same exposure to observable factors.

Proposition 1. *For observable factors, the original portfolio and the replica portfolio have the same loadings, i.e.* $\forall i,\ j,\ t, \quad \beta_{i,j,t} = \hat{\beta}_{i,j,t}$.

Proof. See Appendix. □

By construction, the econometrician ensures that the replica fund has the same loadings on observable factors as the original fund. As for the unobservable factors, the replica fund's loadings should be unbiased since the stocks within the replica fund are picked randomly once the observable factor loadings are match.

Proposition 2. *For well-diversified portfolio weight distribution* $\{w_{i,k,t-1}\}$*, the replica fund's loadings on the unobservable factors are unbiased. That is,* $\forall i,\ l, \quad \hat{\gamma}_{i,l,t} =$

$\sum_K w_{i,k,t-1}\gamma_{\hat{k},l} \approx \sum_K w_{i,k,t-1}\bar{\gamma}_{k,l}$, *where* $\bar{\gamma}_{k,l} \equiv \mathbb{E}\left(\gamma_{k.l} | \{\beta_{k,j}\}\right)$ *is the average loading on factor* $f_{l,t}$ *for stocks in the bucket identified by* $\{\beta_{k,j}\}$*.*

Proof. See Appendix. □

Proposition 2 shows that the replica fund's loadings on the unobservable factors are unbiased because the stocks within the replica fund's portfolio are chosen randomly. On the other hand, if the loadings on the observable factors can already sufficiently characterize the style of the original fund, then the original fund's loadings on the unobservable factors should also be unbiased.

Definition 1. Fund i's style is well-specified by the observable factors iff

$$\begin{aligned} &\sum_k w_{i,k,t-1}\left(\gamma_{k.l} - \mathbb{E}\left(\gamma_{k.l} | \{\beta_{k,j}\}\right)\right) \\ \equiv &\sum_k w_{i,k,t-1}\delta_{k.l} \\ \approx &0 \end{aligned}$$

Equivalently,

$$\gamma_{i,l,t} \approx \sum_k w_{i,k,t-1}\bar{\gamma}_{k,l}$$

As mentioned in Section 3, the matching of style between the original fund and the replica funds indicates that only stock-picking skills rather than factor-timing skills can be measured in this exercise. A fund manager manifests stock-picking skills in his ability to predict idiosyncratic stock returns $\{\epsilon_{k,t}\}_{k=1}^K$. The stock-picking skills of a fund manager can be defined more formally as follows.

Definition 2. Define skilled stock-pickers as the managers with superior information about $\{\epsilon_{k,t}\}_{k=1}^K$ and satisfy:

1. Better firm-specific information, not contingent on factor realizations:

$$\begin{aligned} &\mathbb{E}\left(e_{i,t} | \{w_{i,k,t-1}\}\right) = \mathbb{E}\left(\sum_k w_{i,k,t-1}\epsilon_{k,t} | \{w_{i,k,t-1}\}\right) \equiv \alpha_{i,t} \\ >&\mathbb{E}\left(\hat{e}_{i,t} | \{w_{i,k,t-1}\}\right) = \mathbb{E}\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} | \{w_{i,k,t-1}\}\right) = 0 \end{aligned}$$

and

$$\epsilon_{k,t} \perp F_{j,t} | \{w_{i,k,t-1}\}, \quad \forall k,\ j$$

$$\epsilon_{k,t} \perp f_{l,t} | \{w_{i,k,t-1}\}, \quad \forall k,\ l$$

$$\epsilon_{k,t} \perp \epsilon_{k',t} | \{w_{i,k,t-1}\}, \quad \forall k,\ k'$$

2. Sufficiently diversified, so that Proposition 2 holds and Central Limit Theorem applies:

$$\sum_k w_{i,k,t-1}\epsilon_{k,t} | \{w_{i,k,t-1}\} \sim N\left(\alpha_{i,t}, Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t} | \{w_{i,k,t-1}\}\right)\right)$$
$$\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} | \{w_{i,k,t-1}\} \sim N\left(0, Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} | \left\{w_{i,\hat{k},t-1}\right\}\right)\right)$$

3. Style well-specified by observable factors, no bias towards unobservable factors:

$$\begin{aligned} &\sum_k w_{i,k,t-1}\left(\gamma_{k.l} - \mathbb{E}\left(\gamma_{k.l} | \{\beta_{k,j}\}\right)\right) \\ \equiv &\sum_k w_{i,k,t-1}\delta_{k.l} \\ \approx &0 \end{aligned}$$

Equivalently,

$$\gamma_{i,l,t} \approx \sum_k w_{i,k,t-1}\bar{\gamma}_{k,l} \approx \hat{\gamma}_{i,l,t}$$

An important clarification is warranted. Definition 2 does not aim to exclusively define all types of stock-picking skills under common sense. Instead, the definition only serves to draw the boundary of the empirical search. In other words, a skilled stock-picker under common sense might be excluded by Definition 2 for being significantly under-diversified or biased towards certain unobservable factors to the econometrician. The purpose of this project is to identify a large enough subset of all skilled stock-pickers who are able to deliver positive alpha out of sample.

The rest of this subsection intends to show that, the return distribution of a skilled stock-picker defined above is a mean shift from the return distribution of a corresponding replica fund. That is

$$\begin{aligned}
& r_{i,t} - \alpha_{i,t} \sim \hat{r}_{i,t} \\
\iff & \sum_j \beta_{i,j,t} F_{j,t} + \sum_l \gamma_{i,l,t} f_{l,t} + \sum_k w_{i,k,t-1} \epsilon_{k,t} - \alpha_{i,t} \\
& \sim \sum_j \hat{\beta}_{i,j,t} F_{j,t} + \sum_l \hat{\gamma}_{i,l,t} f_{l,t} + \sum_k w_{i,k,t-1} \epsilon_{\hat{k},t} \\
\iff & \sum_j \beta_{i,j,t} F_{j,t} + v_{i,t} + e_{i,t} - \alpha_{i,t} \sim \sum_j \hat{\beta}_{i,j,t} F_{j,t} + \hat{v}_{i,t} + \hat{e}_{i,t}.
\end{aligned}$$

So far, the following have already been established:

1. $\sum_j \beta_{i,j,t} F_{j,t} = \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.
2. $v_{i,t} = \sum_l \gamma_{i,l,t} f_{l,t} \approx \hat{v}_{i,t} = \sum_l \hat{\gamma}_{i,l,t} f_{l,t}$.
3. $e_{i,t} \perp v_{i,t}$, $e_{i,t} \perp \sum_j \beta_{i,j,t} F_{j,t}$, $\hat{e}_{i,t} \perp \hat{v}_{i,t}$ and $\hat{e}_{i,t} \perp \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.

The next step is to prove that $e_{i,t} - \alpha_{i,t} \sim \hat{e}_{i,t}$ for a skilled stock-picker.

Proposition 3. *For a skilled stock-picker defined in Definition 2, the residual risk of the fund portfolio is a mean shift to that of the corresponding replica fund. That is,*

$$e_{i,t} - \alpha_{i,t} \sim \hat{e}_{i,t}.$$

Proof. See Appendix. □

Proposition 4. *The return of a skilled stock-picker defined in Definition 2 is approximately a mean shift to the return of the replica fund. That is,*

$$r_{i,t} - \alpha_{i,t} \sim \hat{r}_{i,t}.$$

Proof. Directly from the fact that:

1. $\sum_j \beta_{i,j,t} F_{j,t} = \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.
2. $v_{i,t} = \sum_l \gamma_{i,l,t} f_{l,t} \approx \hat{v}_{i,t} = \sum_l \hat{\gamma}_{i,l,t} f_{l,t}$.
3. $e_{i,t} \perp v_{i,t}$, $e_{i,t} \perp \sum_j \beta_{i,j,t} F_{j,t}$, $\hat{e}_{i,t} \perp \hat{v}_{i,t}$ and $\hat{e}_{i,l} \perp \sum_j \hat{\beta}_{i,j,t} F_{j,t}$.
4. $e_{i,t} - \alpha_{i,t} \sim \hat{e}_{i,t}$.

$\square$

The intuition of this result is straightforward. According to Definition 2, the real fund and the replica fund have the same exposures to all systematic risks. If the real fund's manager is skilled at picking stocks, then the idiosyncratic component of the real fund's return has higher mean compared to the replica fund's return. On the other hand, the distribution of the idiosyncratic component of the real fund has the same shape as the replica fund due to the Central Limit Theorem and the fact that the volatility of the stocks in both funds are matched by construction. The result then follows because of the independence between the systematic and idiosyncratic components in fund returns.

Proposition 5. *The return of a skilled stock-picker defined in Definition 2 first-order stochastically dominates the return of the replica fund. That is,*

$$r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}.$$

Proof. The result is directly from Proposition 4 and the fact that $\alpha_{i,t} > 0$ according to Definition 2. $\square$

1.5 FSD Implementation

This section demonstrates how to implement the FSD condition with the time-series of the counterfactual return distributions $\{\langle \hat{r}_{i,t} \rangle\}_{t=1}^{T}$. A test statistic for the FSD condition is proposed, and its finite-sample distribution is computed with bootstrap simulation.

1.5.1 Ranking FSD

The key step to implement the FSD condition is to transform the return FSD condition to a ranking FSD condition.

Proposition 6. *A real fund's return being first-order stochastically dominant to a replica fund's return is equivalent to the condition that the ranking of the real fund's return among the cohort of replica funds being first-order stochastically dominant to the ranking of the replica fund's return among the cohort of replica funds. Moreover,* *the ranking of a replica fund's return among the cohort of replica funds follows a standard uniform distribution. That is*

$$r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t} \iff Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right) \overset{fsd}{\succ} Pct\left(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle\right) \sim Unif\,(0,1)$$

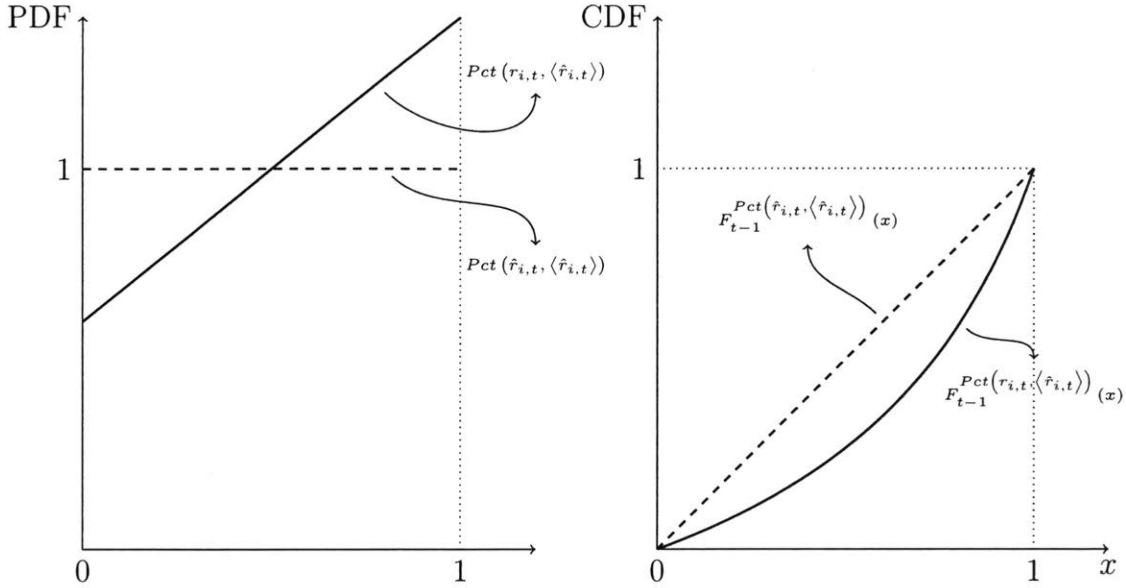
where $Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)\left(Pct\left(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle\right)\right)$ *denotes the percentile of the real(replica) fund's return in the counterfactual return distribution;* $Unif\,(0,1)$ *denotes the uniform distribution with support* $[0,1]$.

Proof. See Appendix. $\square$

Proposition 7. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t} \iff F_{t-1}^{Pct(r_{i,t},\langle \hat{r}_{i,t} \rangle)}(x) < F_{t-1}^{Pct(\hat{r}_{i,t},\langle \hat{r}_{i,t} \rangle)}(x) = x$, *where* F_{t-1} *denotes the conditional CDF.*

Proof. Immediate from Proposition 6. $\square$

Figure 1.2 offers a graphical illustration of the $\overset{fsd}{\succ}$ condition. Panel (a) is a demonstration regarding the relation between $Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$ and $Pct\left(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$. The dashed line plots the PDF of $Pct\left(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$, which is a flat horizontal line constant at 1 since $Pct\left(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$ follows a standard uniform distribution. The solid line is an example of $Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$. Since $Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right) \overset{fsd}{\succ} Pct\left(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$, the solid line has a smaller left tail compared to the dashed line, but a larger right tail in the PDF plot. Panel (b) offers the illustration on the same relation with CDF plots. The $Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right) \overset{fsd}{\succ} Pct\left(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$ condition is reflected in the plot as the solid curve strictly lies below the dashed line.



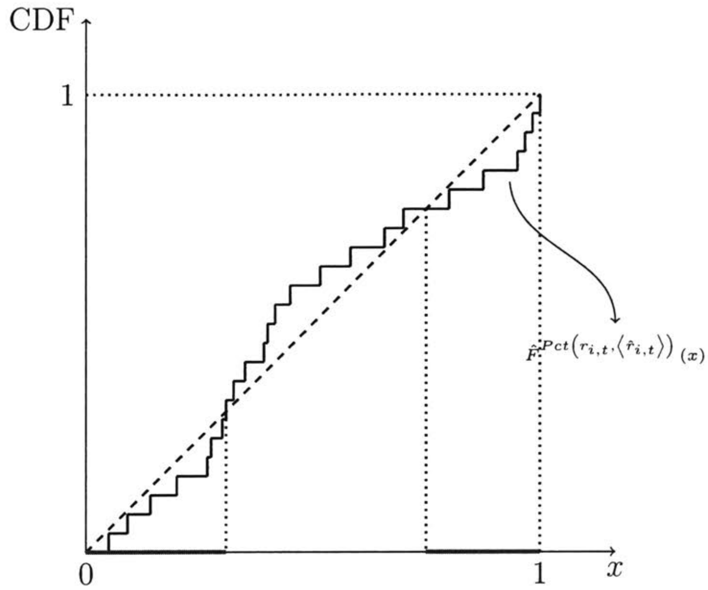
(a) (b)

Figure 1.2: Graphical Illustration of the FSD Condition
This figure offers a graphical illustration of the FSD condition. Panel (a) is a demonstration regarding the relation between $Pct\,(r_{i,t}, \langle \hat{r}_{i,t} \rangle)$ and $Pct\,(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)$. The dashed line plots the PDF of $Pct\,(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)$, which is a flat horizontal line constant at 1 since $Pct\,(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)$ follows a standard uniform distribution. The solid line is an example of $Pct\,(r_{i,t}, \langle \hat{r}_{i,t} \rangle)$. Since $Pct\,(r_{i,t}, \langle \hat{r}_{i,t} \rangle) \overset{fsd}{\succ} Pct\,(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)$, the solid line has a smaller left tail compared to the dashed line, but a larger right tail in the PDF plot. Panel (b) offers the illustration on the same relation with CDF plots. The $Pct\,(r_{i,t}, \langle \hat{r}_{i,t} \rangle) \overset{fsd}{\succ} Pct\,(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)$ condition is reflected in the plot as the solid curve strictly lies below the dashed line.

Proposition 7 establishes the relation between $Pct\,(r_{i,t}, \langle \hat{r}_{i,t} \rangle)$ and $Pct\,(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)$ conditional on the information set by the end of $t-1$. Conditional relations between distributions are not empirically observable. However, Proposition 7 also indicates that the conditional CDF of $Pct\,(r_{i,t}, \langle \hat{r}_{i,t} \rangle)$ always lies below the 45 degree line, which is an appealing feature to facilitate time aggregation.

Proposition 8. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t},\ \forall t \Rightarrow F^{Pct(r_{i,t}, \langle \hat{r}_{i,t} \rangle)}(x) < F^{Pct(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)}(x) = x$, *where F denotes the unconditional CDF.*

Proof. $F^{Pct(r_{i,t}, \langle \hat{r}_{i,t} \rangle)}(x) = \mathbb{E}\left[F_{t-1}^{Pct(r_{i,t}, \langle \hat{r}_{i,t} \rangle)}(x)\right] < \mathbb{E}\left[F_{t-1}^{Pct(\hat{r}_{i,t}, \langle \hat{r}_{i,t} \rangle)}(x)\right] = \mathbb{E}(x) = x.$ $\square$



Test Statistic: $\hat{\theta}$ = length of —

Figure 1.3: FSD Test Statistic $\hat{\theta}$

This figure illustrates the construction of the FSD test statistic $\hat{\theta}$. The solid step function is an illustration of the empirical CDF of $Pct(r_{i,t}, \langle\hat{r}_{i,t}\rangle)$. The test statistic $\hat{\theta}$ is constructed as the measure of the region where the empirical CDF falls below the dashed 45 degree line.

Note that the unconditional CDF of the ranking of a manager among replica funds is an empirically observable object. I therefore, employ the empirical counterpart of $F^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$, i.e. $\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$, to implement the FSD condition.

1.5.2 FSD Test Statistic

Figure 1.3 illustrates the construction of the FSD test statistic. Specifically, for each fund, I construct its empirical ranking CDF $\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$ with its historical returns.[10] The FSD test statistic $\hat{\theta}_i \in [0,1]$ is then defined as the measure of the region where $\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$ lies below the 45 degree line. The FSD condition is perfectly satisfied in sample if $\hat{\theta}_i = 1$, and a higher $\hat{\theta}_i$ indicates a better fit of the FSD condition.

[10]I use the 24 proceeding monthly returns in the empirical exercise.

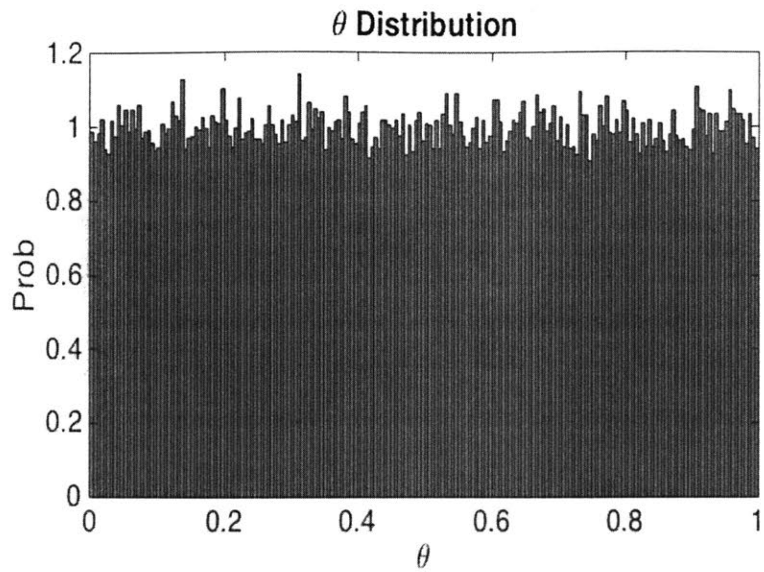


Figure 1.4: Simulated $\hat{\theta}$ Null Distribution
This figure plots distribution of the FSD test statistic $\hat{\theta}$ constructed from 24 observations under the null hypothesis that $Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right)$ follows a standard uniform distribution.

Figure 1.4 demonstrates the simulated finite-sample distribution of the FSD test statistic $\hat{\theta}$ constructed from 24 observations under the null hypothesis that $Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right) \sim Unif\left(0,1\right)$. The null distribution of $\hat{\theta}$ seems to follow a standard uniform distribution itself, although the proof of this result is beyond the scope of this paper. According to the simulation, a test size of 10%(5%) corresponds to the critical value of 0.90(0.95).

In general, the test statistic $\hat{\theta}$ can only be used to evaluate of the goodness of fit of the FSD condition, but it is unable to measure the magnitude of the fund's true α. Therefore, in practice, the FSD filter is better used in combination with the standard $\hat{\alpha}$ sort. The FSD filter serves to rule out potential false positives; whereas the $\hat{\alpha}$ sort measures the magnitude of the potential stock-picking skills.

1.6 Finite Sample Robustness

This section describes two specific mechanisms through which the FSD condition is able to improve the power of the conventional positive alpha condition. The two mechanisms correspond to two statistical problems that might be present in the data: heteroscedasticity and benchmark mis-specification. The heteroscedasticity problem is defined as the situation where idiosyncratic volatility is time-varying and more volatile than the fund's true alpha; whereas the benchmark mis-specification problem is defined as some managers taking on factors that are unobservable to the econometrician. The FSD condition possesses superior econometric properties compared to the positive alpha condition because:

1. Heteroscedasticity:
   (a) Positive alpha condition: Assigns equal weight to all observations regardless of the level of idiosyncratic volatility.
   (b) FSD condition: Weights observations differently according to the signal-to-noise ratio in different periods.
2. Benchmark Mis-specification:
   (a) Positive alpha condition: Tends to mistakenly identify mis-specified managers who take on unobservable factors with high in-sample realizations as being skilled.
   (b) FSD condition: Offers a detection mechanism to rule out mis-specified managers by checking the left tails of their return distributions.

I provide both theoretical arguments and simulation results to illustrate the superiority of the FSD condition to the positive alpha condition in these aspects.

1.6.1 Simulation Environment

In order to illustrate the arguments, I construct an artificial economy with the following features. I simulate 1000 fund managers, among whom 20 are skilled with α being 25 bps per month. The percentage of skilled managers and the magnitude of their α is determined according to the findings documented in Fama and French (2010). Except for the case studying benchmark mis-specification, I assume a one-factor structure, e.g. the market factor $r_{m,t}$. Without loss of generality, I assume that all funds have unit loading on the single factor, and the risk-free rate is zero. I also assume that the single factor follows normal distribution: $r_{m,t} \sim N\left(0, 0.06^2\right)$, i.e. the factor has 0 mean and $6\%(6\% \times \sqrt{12} = 21\%)$ monthly(annualized) volatility. Therefore, the return of an unskilled manager is

$$r_{i,t}^{unskilled} = r_{m,t} + u_{i,t}, \quad \mathbb{E}\left(u_{i,t}\right) = 0, \; r_{m,t} \sim N\left(0, 0.06^2\right)$$

and the return of a skilled manager is

$$r_{i,t}^{skilled} = 25bps + r_{m,t} + u_{i,t}, \quad \mathbb{E}\left(u_{i,t}\right) = 0, \; r_{m,t} \sim N\left(0, 0.06^2\right).$$

The performances of the replica funds in the construction of the FSD test statistic have the same properties as the unskilled managers. That is

$$\hat{r}_{i,t} = r_{m,t} + \hat{u}_{i,t}, \quad \hat{u}_{i,t} \sim u_{i,t}, \; r_{m,t} \sim N\left(0, 0.06^2\right).$$

I then select 20 best performing managers in the simulated data by using the test statistics $\hat{\alpha}$ and $\hat{\theta}$ of the positive alpha condition and the FSD condition respectively, and compare the accuracy of these two filters in identifying skilled fund managers.

1.6.2 Robustness to Heteroscedasticity

In order to study the influence of return heteroscedasticity on the performance measures, I assume that the volatility of idiosyncratic returns follows the following (almost) AR1 process:

$$\sigma_{i,t} = \max\left\{\sigma_{i,t-1} + \rho\left(\bar{\sigma} - \sigma_{i,t-1}\right) + \zeta\epsilon_t, \; 0\right\}.$$

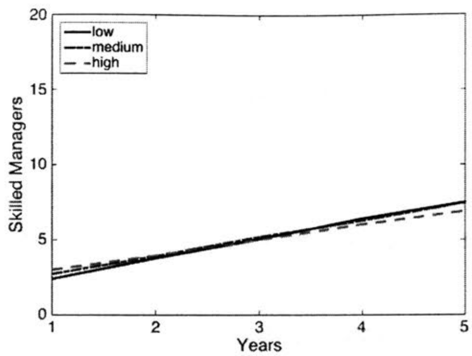
where ρ determines the speed of mean-reversion of the volatility process, $\bar{\sigma}$ is the long-run volatility, ζ is the volatility of the volatility process, and ϵ_t is a standard normal shock, i.e. $\epsilon_t \sim N(0,1)$.

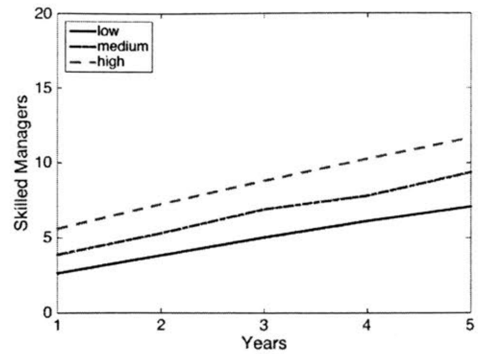
For simplicity, I assume that the idiosyncratic return components are also normal:

$$u_{i,t} \sim \hat{u}_{i,t} \sim N\left(0, \sigma_{i,t}^2\right).$$

In the simulations, I specify: $\rho = 0.9^*$, $\bar{\sigma} = 0.01^*$; and I consider three different degrees of heteroscedasticity: $\zeta_L = 0.0012$, $\zeta_M = 0.0018$ and $\zeta_H = 0.0024^*$.[11] The

[11]Values of the "*" sign are calibrated with the real data.





(a) The $\alpha_i > 0$ Filter

(b) The $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ Filter

Figure 1.5: The Heteroscedastic Case: Filter Accuracy

The plot compares the accuracy of the $\alpha_i > 0$ and the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filters when idiosyncratic volatility is time-varying. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of skilled managers that the corresponding performance measure is able to identify over 500 simulation paths. The black solid(blue dotted, red dashed) line represents the level of heteroscedasticity with $\zeta = 0.0012(\zeta = 0.0018,\ \zeta = 0.0024^{*})$. Panel (a) plots the effectiveness of the $\alpha_i > 0$ filter; whereas Panel (b) plots the effectiveness of the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filter.

return processes for the managers in this heteroscedastic economy are thus fully specified.

In this economy, the positive alpha filter should have poor performance when heteroscedasticity gets intensified because it assigns equal weight to all observations so that lucky shocks from high volatility periods are of large magnitudes and are difficult to be cancelled out by shocks in other periods.

The FSD filter, on the other hand, naturally does not suffer from this problem because idiosyncratic volatility is adjusted period by period when the ranking of the real fund return among replica funds($Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$) is taken.

Figure 1.5 compares the effectiveness of the positive alpha filter versus the FSD filter in the heteroscedastic economy. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of skilled managers that the corresponding filter is able to identify over 500 simulation paths. The black solid(blue dotted, red dashed) line represents the level of heteroscedasticity with $\zeta = 0.0012(\zeta = 0.0018,\ \zeta = 0.0024^*)$. Panel (a) plots the effectiveness of the positive alpha filter; whereas Panel (b) plots the effectiveness of the FSD filter.

From the figure, the real-world level of heteroscedasticity is mild enough so that the positive alpha filter is virtually unaffected. Interestingly, Panel (b) of the figure shows that the effectiveness of the FSD filter improves with heteroscedasticity. The intuition of this result is that idiosyncratic volatility gets adjusted in each period. High volatility periods are assigned with low weights and low volatility periods are assigned with high weights. Therefore, the FSD filter improves with heteroscedasticity because it is able to take advantage of the high signal-to-noise ratio in periods with low idiosyncratic volatility.

1.6.3 Benchmark Mis-specification Detection

Another difficult problem encountered by the positive alpha condition is that the benchmark index to which a fund's performances are compared might be inappropriately chosen so that the outperformances in each period as well as the overall α might be measured with error.

To fix ideas, I modified the aforementioned simulation environment to introduce a third type of fund managers – the mis-specified managers. The return processes of the three types of managers are as follows:

The unskilled managers:

$$r_{i,t}^{unskilled} = r_{m,t} + u_{i,t}, \quad u_{i,t} \sim N\left(0, 0.01^2\right),\ r_{m,t} \sim N\left(0, 0.06^2\right).$$

The skilled managers:

$$r_{i,t}^{skilled} = 25bps + r_{m,t} + u_{i,t}, \quad u_{i,t} \sim N\left(0, 0.01^2\right),\ r_{m,t} \sim N\left(0, 0.06^2\right).$$

The mis-specified managers:

$$r_{i,t}^{mis-spec} = r_{m,t} + f_{i,t} + u_{i,t}, \quad u_{i,t} \sim N\left(0, 0.01^2\right),\ r_{m,t} \sim N\left(0, 0.06^2\right),\ f_{i,t} \sim N\left(0, \sigma_f^2\right).$$

The returns of the replica funds have the same properties for all three types of managers. That is

$$\hat{r}_{i,t} = r_{m,t} + \hat{u}_{i,t}, \quad \hat{u}_{i,t} \sim u_{i,t},\ r_{m,t} \sim N\left(0, 0.06^2\right).$$

Notice that for the mis-specified managers, the replica funds have the same exposure to the observable factor $r_{m,t}$, but do not load on the unobservable factor $f_{i,t}$.

The mis-specified managers have no stock-picking skills so that they generate zero α. The difference between the mis-specified and the unskilled managers is that a mis-specified manager takes on factor risk $f_{i,t}$ that is not observable to the econometrician. Thus, $f_{i,t}$ is not controlled for in the replica funds. The volatility of the uncontrolled factor σ_f is a measure of the degree of mis-specification in this economy.

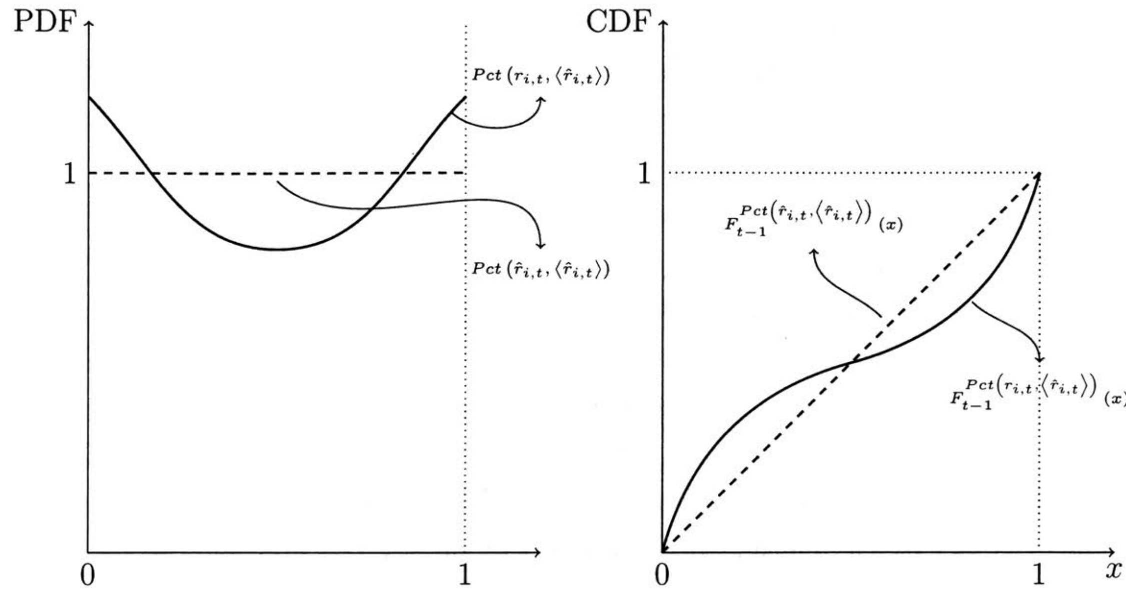
I assume idiosyncratic volatility to be constant in this case so that $u_{i,t} \sim \hat{u}_{i,t} \sim N\left(0, 0.01^2\right)$. I consider three levels of uncontrolled factor volatility with $\sigma_f = 0.01$ ($\sigma_f = 0.03$, $\sigma_f = 0.05$) representing the case of mild (moderate, severe) mis-specification. For simplicity, the factor is assumed to follow normal distribution: $f_{i,t} \sim N\left(0, \sigma_f^2\right)$, and the uncontrolled factors for two managers are uncorrelated. There are 1000 managers in total in the economy. Among them, 20 managers are skilled, 100 managers are mis-specified, and the remaining 880 managers are unskilled.

The existence of the mis-specified managers might severely compromise the effectiveness of the positive alpha filter. To see that, the $\hat{\alpha}$ of a mis-specified manager is

$$\begin{aligned}\hat{\alpha}_i &= \frac{1}{T}\sum_t \left(r_{i,t} - r_{m,t}\right) \\ &= \frac{1}{T}\sum_t f_{i,t} + \frac{1}{T}\sum_t u_{i,t}.\end{aligned}$$

Thus the existence of unobservable factors might obscure the measurement of skill because $\hat{\alpha}_i$ can be dominated by the realization of $\frac{1}{T}\sum_t f_{i,t}$ especially when sample is short (T is small) or mis-specification is severe (σ_f is large).

The FSD condition is able to alleviate this problem by offering a detention mechanism. Suppose the fund has non-trivial loading on a factor that is not controlled in the bootstrap process during the construction of the replica portfolios, then the first-order stochastic dominance condition is likely to be violated. Indeed, during periods in which the uncontrolled factor has large positive realizations, the manager would rank highly compared to the replica funds, and vice versa during periods when the uncontrolled factor has large negative realizations. As as result, the PDF of the ranking of the manager compared to the replica funds($Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right)$) shall have both large left and right tails, violating the first-order stochastic dominance condition. Figure 1.6 offers a graphical illustration of the detection mechanism.



(a) (b)

Figure 1.6: Detection for Uncontrolled Factors

This figure offers a graphical illustration of the detention mechanism of the FSD condition for mis-specified managers. Panel (a) shows that the PDF of the mis-specified manager's ranking($Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$) has both larger left and right tails compared to the standard uniform distribution; Panel (b) shows that the CDF of the mis-specified manager's ranking($Pct\left(r_{i,t}, \langle \hat{r}_{i,t} \rangle\right)$ goes above the 45 degree line for some region.

The following proposition provides a more general statement for this argument when the uncontrolled factor is allowed to have non-zero risk premium.

Proposition 9. *Consider a mis-specified manager's return process:*

$$\begin{aligned} r_{i,t} &= r_f + \sum_j \beta_{i,j} F_{j,t} + \sum_l \gamma_{i,l} f_{l,t} + e_{i,t} \\ &\equiv r_f + \sum_j \beta_{i,j} F_{j,t} + \tilde{f}_{i,t} + e_{i,t} \end{aligned}$$

$$\tilde{f}_{i,t} \sim N\left(\mu_f, \sigma_f^2\right), \quad e_{i,t} \sim N\left(0, \sigma_i^2\right), \quad \tilde{f}_{i,t} \perp e_{i,t}.$$

where $\{F_{j,t}\}_{j=1}^{J}$ *are the observable factors,* $\{f_{l,t}\}_{l=1}^{L}$ *are the unobservable factors,* $e_{i,t}$ *is the idiosyncratic component.*

A corresponding replica fund has the return process:

$$\hat{r}_{i,t} = r_f + \sum_j \beta_{i,j} F_{j,t} + \hat{e}_{i,t}$$
$$\hat{e}_{i,t} \sim N\left(0, \sigma_i^2\right).$$

The first-order stochastic dominance condition $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ *is violated as long as* $\sigma_f > 0$.

Proof. See Appendix. □

Of course, the proposition only holds under the special condition that both the uncontrolled factors and the idiosyncratic risk are normal, and are independent to each other. The conclusion of the proposition can be violated if one considers alternative distribution specifications of the uncontrolled factors. However, the proposition conveys the intuition that the FSD condition is able to detect the existence of the uncontrolled factors because the mis-specified managers are likely to have a larger left tail in their return distributions compared to the replica funds when they take on factors that are not controlled by the econometrician, as demonstrated in Figure 1.7.

Figure 1.8 and 1.9 compare the effectiveness of the positive alpha filter versus the FSD filter in simulation. Figure 1.8 demonstrates the two filters' ability to identify skilled managers; whereas Figure 1.9 illustrates their tendencies to select mis-specified managers. Specifically, the x-axis is the formation period from which the performance measures are constructed. For Figure 1.8, the y-axis is the average number of skilled managers that the corresponding filter is able to identify over 500 simulation paths; whereas for Figure 1.9, the y-axis is the average number of mis-specified managers that the corresponding filter erroneously selects over 500 simulation paths. The black solid(blue dotted, red dashed) line represents the situation where benchmark mis-specification is mild $\sigma_f = 0.01$(moderate $\sigma_f = 0.03$, severe $\sigma_f = 0.05$).

From the figures, the positive alpha filter identifies fewer skilled managers when benchmark mis-specification becomes more severe because it tends to erroneously select mis-specified managers whose uncontrolled factors have high in-sample realizations. The FSD filter, however, is unaffected by benchmark mis-specification because

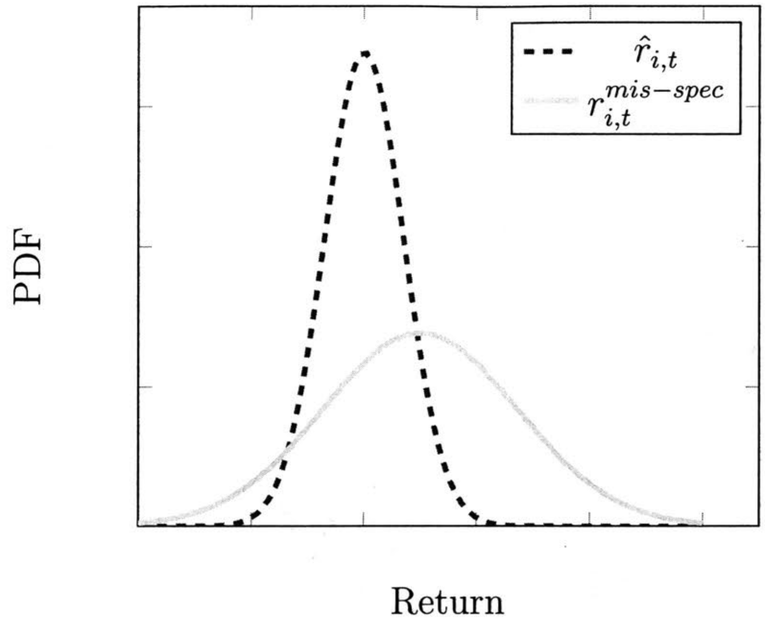


Figure 1.7: Mis-specified Return Distribution
This figure illustrates that the return distribution of a mis-specified manager carries a heavier left tail compared to the return distribution of the replica fund.

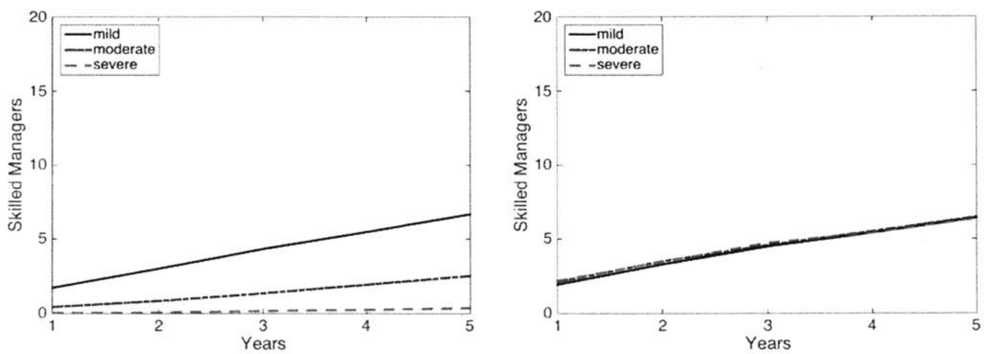
the first-order stochastic dominance condition excludes the mis-specified managers as their rankings relative to replica funds would have both larger left and right tails compared to a uniform distribution.

1.6.4 Two-stage Sort

In order to illustrate the combined effects of the FSD filter and the $\hat{\alpha}$ sort, I conduct a two-stage sort in this subsection and compare the results to the search with the $\hat{\alpha}$ sort alone in simulation. The economy features both heteroscedasticity and benchmark mis-specification. In addition, I also allow the magnitude of the true alphas to vary among skilled managers in the economy. The heterogeneous alphas suggest a role for the $\hat{\alpha}$ sort in the second stage.

There are three types managers in this economy skilled, unskilled and mis-specified. Among the skilled managers, there are five subgroups with different αs. The return processes of the managers are as follows.

The skilled managers:

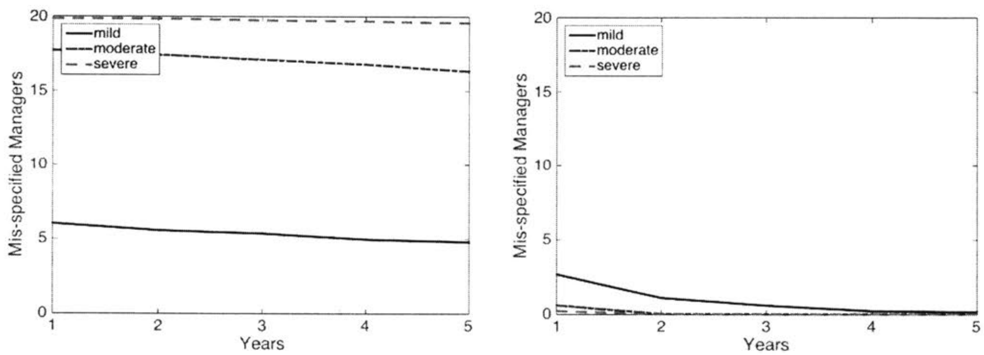


(a) The $\alpha_i > 0$ Filter

(b) The $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ Filter

Figure 1.8: The Mis-specification Case: Filter Accuracy

The plot compares the accuracy of the $\alpha_i > 0$ and the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filters in the presence of benchmark mis-specification. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of skilled managers that the corresponding performance measure is able to identify over 500 simulation paths. The black solid(blue dotted, red dashed) line represents the situation where benchmark mis-specification is mild $\sigma_f = 0.01$(moderate $\sigma_f = 0.03$, severe $\sigma_f = 0.05$). Panel (a) plots the effectiveness of the $\alpha_i > 0$ filter; whereas Panel (b) plots the effectiveness of the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filter.



(a) The $\alpha_i > 0$ Filter

(b) The $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ Filter

Figure 1.9: The Mis-specification Case: Filter Mistake

The plot compares the tendency to select mis-specified managers of the $\alpha_i > 0$ and the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filters in the presence of benchmark mis-specification. The x-axis is the formation period from which the performance measures are constructed. The y-axis is the average number of mis-specified managers that the corresponding performance measure erroneously selects over 500 simulation paths. The black solid(blue dotted, red dashed) line represents the situation where benchmark mis-specification is mild $\sigma_f = 0.01$(moderate $\sigma_f = 0.03$, severe $\sigma_f = 0.05$). Panel (a) plots the performance of the $\alpha_i > 0$ filter; whereas Panel (b) plots the performance of the $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}$ filter.

$$r_{i,t}^{skilled} = \alpha_i + r_{m,t} + u_{i,t}, \text{ 100 funds:}$$

$$\alpha_i = 5bps, \text{ 20 funds;}$$

$$\alpha_i = 10bps, \text{ 20 funds;}$$

$$\alpha_i = 15bps, \text{ 20 funds;}$$

$$\alpha_i = 20bps, \text{ 20 funds;}$$

$$\alpha_i = 25bps, \text{ 20 funds.}$$

The mis-specified managers:

$$r_{i,t}^{mis-spec} = r_{m,t} + f_{i,t} + u_{i,t}, \text{ 100 funds.}$$

The unskilled managers:

$$r_{i,t}^{unskilled} = r_{m,t} + u_{i,t}, \text{ 800 funds.}$$

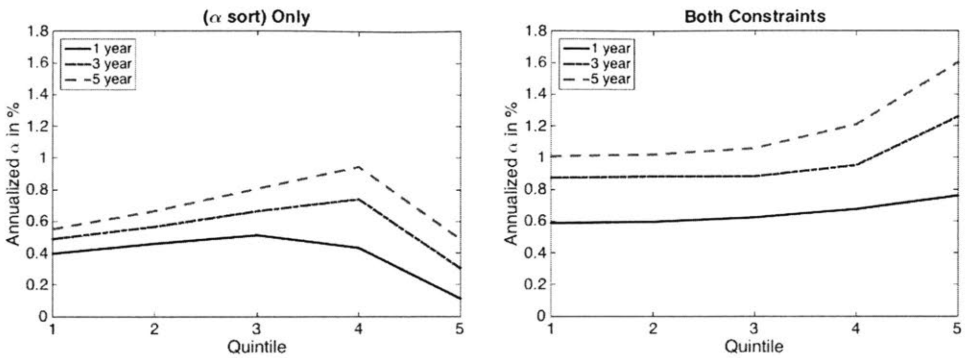
The replica funds:

$$\hat{r}_{i,t} = r_{m,t} + \hat{u}_{i,t}, \; \hat{u}_{i,t} \sim u_{i,t}.$$

The market factor still has the same properties as before: $r_{m,t} \sim N(0, 0.06^2)$. The economy features heteroscedasticity and the distribution of the idiosyncratic components follows: $N\left(0, \sigma_{i,t}^2\right)$, $\sigma_{i,t} = \max\left\{\sigma_{i,t-1} + \rho\left(\bar{\sigma} - \sigma_{i,t-1}\right) + \zeta\epsilon_t, \; 0\right\}$, $\rho = 0.9^*$, $\bar{\sigma} = 0.01^*$, $\zeta = 0.0024^*$. The economy also features benchmark mis-specification. I assume that the uncontrolled factors take moderate volatility as in the previous subsection: $f_{i,t} \sim N\left(0, \sigma_f^2\right)$, $\sigma_f = 0.03$.

The two-stage sort is conducted with the following procedure. The FSD filter first selects 100 managers with the highest in-sample FSD test statistics. The 100 managers are then sorted into 5 quintiles based on their in-sample realized $\hat{\alpha}$s. As a comparison, I also form 5 quintiles of funds with the $\hat{\alpha}$ sort alone, with each quintile containing 20 funds to match the sample size. Therefore, the top quintile of the single $\hat{\alpha}$ sort consists of the 20 funds with the highest in-sample $\hat{\alpha}$s, and likewise for other quintiles.

Figure 1.10 compares the performance of the two-stage sort versus the performance of the single $\hat{\alpha}$ sort in simulation. The black solid(blue dotted, red dashed) line



(a) ($\hat{\alpha}$ sort) Only　　　(b) $(r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}) + (\hat{\alpha}$ sort$)$

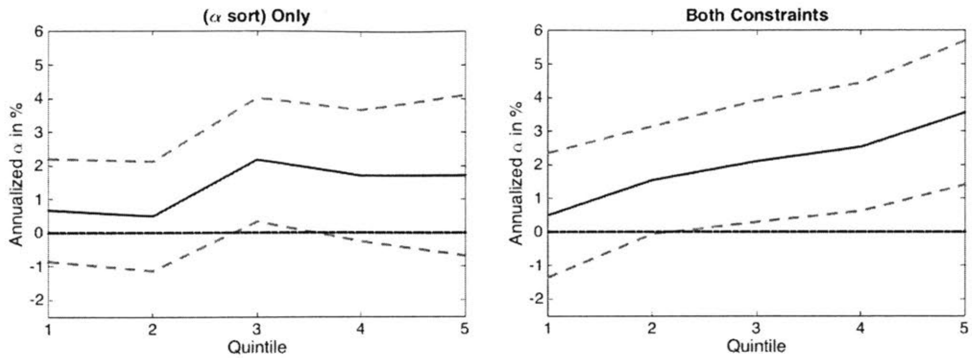
Figure 1.10: Out-of-sample α, in Simulation, Matched Sample Size
The plot compares the out-of-sample α in simulation of the $\hat{\alpha}$ sort only with the combination of the FSD filter and the $\hat{\alpha}$ sort. The sample sizes in the two panels are matched. Each quintile contains 2% of all funds in the cross section. The black solid(blue dotted, red dashed) line represents the sort using 1(3, 5) year(s) of historical data.

represents the sort using 1(3, 5) year(s) of historical data. The out-of-sample αs in the two-stage sort increase monotonically with the in-sample $\hat{\alpha}$ quintiles; whereas the out-of-sample αs in the single $\hat{\alpha}$ sort is non-monotonic, especially for the top quintile. The non-monotonicity is caused by the presence of the mis-specified managers obscuring the measurement of α in the single $\hat{\alpha}$ sort. Figure 1.11 plots the results of the same sorts conducted with the actual data. The out-of-sample αs of the sorts with the actual data follow similar patterns as the ones in simulation. Specifically, the out-of-sample αs are non-monotonic with the in-sample $\hat{\alpha}$ quintiles in a single $\hat{\alpha}$ sort, but a two-stage sort with the FSD filter in the first stage is able to fix this problem. The finding suggests that the mechanisms studied in this section might also be present in the actual data.

1.7 Empirical Findings

1.7.1 Data

I obtain monthly after-fees fund returns along with other fund characteristics such as fund size, age, name, expense ratio, etc. from CRSP Survivor-Bias-Free US Mutual Fund Database. I compute the before-fees returns by adding back the expense ratio



(a) ($\hat{\alpha}$ sort) Only (b) $(r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t}) + (\hat{\alpha}$ sort)

Figure 1.11: Out-of-sample α, in Data, Matched Sample Size
The plot compares the out-of-sample α in actual data of the $\hat{\alpha}$ sort only with the combination of the FSD filter and the $\hat{\alpha}$ sort. The sample sizes in the two panels are matched. Each quintile contains 2% of all funds in the cross section. The sort uses data from the most recent 24 months. Both point estimates and confidence intervals are plotted. The sample period is from January 1991 to December 2015.

to the after-fees fund returns. I obtain fund holdings from Thomson Reuters Mutual Fund Holdings (s12), formerly known as the CDA/Spectrum Mutual Fund Holdings Database. Both databases are standard in this line of research. Their popularity arose largely due to their efforts to eliminate survivorship bias by making an attempt to include all funds that have ever existed in the US market. In fact, Linnainmaa (2013) raised the concern of a potential reverse survivorship bias by using these databases as funds hit by a series of unlucky negative shocks tend to exit the market, leaving behind trajectories of poor performances without the chances to "clear their names". Therefore, my finding of superior out-of-sample performances is unlikely to be caused potential survivorship bias. I follow the standard approach to link these two databases with the MFLINKS database constructed by Prof. Russ Wermers, and I obtain stock prices and returns from the CRSP Monthly Stock File.

I limit my focus to domestic, open-end, actively managed, US equity funds. I employ the investment objectives code (crsp_obj_cd) that has been recently introduced by CRSP as my screening variable to identify such funds.[12] Doshi, Elkamhi, and Simutin (2015) shows that the funds identified with the crsp_obj_cd are almost

[12] I include funds with crsp_obj_cd that begins with "EDC" or "EDY"; exclude funds with crsp_obj_cd being "EDYH" or "EDYS"; and exclude option income funds with Strategic Insight Objectives code being "OPI". I then eliminate index funds by screening fund names.

identical to the funds identified with the investment objectives codes from other data vendors that have been used in earlier literature.[13] To reduce the impact from very small funds, I require the funds in my sample to have at least \$10 million under management and hold at least 10 stocks in their portfolios. I aggregate funds with multiple share classes into a single class as these different share classes have the same portfolio composition. In order to have enough funds for this project, I take the sample period from January 1991 to December 2015. I have 2693 distinctive funds in my sample and 399,631 fund-month observations. Table 1.3 documents the summary statistics of the funds that are included in my sample.

1.7.2 Out-of-sample Performances

Table 1.4 documents the out-of-sample performances of the funds identified by the FSD condition. Specifically, by the end of each quarter, I compute the empirical CDF of the percentile of each fund in the counterfactual return distribution ($\hat{F}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$). I then construct the FSD test statistic $\hat{\theta}$ for each fund. I install the FSD filter to select funds with $\hat{\theta} \geq 0.90$, which corresponds to a 10% test size. The FSD filter alone is able to identify a group of fund managers who are able to outperform the Carhart benchmark by 204 bps per year before fees (78 bps per year after fees) out of sample.

More interestingly, among the funds that have been identified by the FSD filter, I further sort them into 5 quintiles by their Carhart four-factor alphas during the proceeding 24 months. Table 1.4 shows that for those FSD satisfying funds, historical alphas do predict future performances. Specifically, the average out-of-sample alphas of the funds increase monotonically with historical realized alphas. The funds in the top quintile are able to, on average, outperform the Carhart four-factor benchmark by as much as 355 bps per year before fees (224 bps per year after fees). The finding of alpha persistence among the identified funds is consistent with the arguments that the FSD condition is able to identify a group of fund managers who are potentially skilled at stock-picking.

Figure 1.12 plots the time-series of the before-fees performances of the selected mutual funds. Panel (a) plots the time-series of the fund performances for all funds selected by the FSD filter; whereas Panel (b) plots the top quintile of the funds

[13] I thank the authors for sharing their SAS code online.

Year	# of Funds	Age	TNA (in 10^6 \$)	# of Stocks	Fees (in bps)
1991	168	19	360	77	125
1992	182	18	517	86	130
1993	241	16	601	92	130
1994	271	16	717	94	135
1995	275	17	779	96	137
1996	289	17	1226	101	134
1997	347	16	1386	108	132
1998	448	15	1530	115	128
1999	555	14	1720	105	130
2000	602	13	1942	104	133
2001	658	13	1654	113	135
2002	775	13	1368	105	140
2003	916	13	988	103	143
2004	1011	13	1257	110	149
2005	1067	13	1347	113	145
2006	1159	14	1380	114	143
2007	1207	14	1465	117	141
2008	1245	14	1561	118	137
2009	1260	15	888	123	137
2010	1224	15	1205	123	140
2011	1194	15	1389	121	135
2012	1258	16	1244	115	131
2013	1235	16	1376	119	129
2014	1181	17	1915	120	126
2015	1087	18	2032	120	122

Table 1.3: Fund Characteristics Summary

This table documents the characteristics of the funds in my sample. The documented characteristics include total number of funds in the sample, the average fund age, the average total net assets in million dollars, the average number of stocks that a fund holds, the average fees that a fund charges in basis points per year.

Quintile	Sample Share	α (in %)	IR	SR	mkt	smb	hml	umd
		Before Fees						
1	1.83%	0.50	0.11	0.59	1.03	0.29	0.03	0.04
		[0.53]			[46.26]	[8.82]	[0.87]	[2.08]
2	1.94%	1.54*	0.38	0.66	1.01	0.20	0.06	0.03
		[1.88]			[59.22]	[5.89]	[2.07]	[1.58]
3	1.95%	2.10**	0.47	0.71	1.00	0.23	0.10	0.05
		[2.28]			[43.55]	[7.12]	[2.35]	[2.48]
4	1.94%	2.53***	0.52	0.71	1.03	0.26	0.01	0.07
		[2.60]			[47.04]	[6.80]	[0.21]	[2.53]
5	1.88%	3.55***	0.67	0.72	1.07	0.45	−0.10	0.08
		[3.24]			[38.85]	[13.47]	[−2.87]	[3.26]
1st Stage	9.55%	2.04***	0.57	0.69	1.03	0.28	0.02	0.05
		[2.78]			[61.14]	[10.55]	[0.72]	[2.99]
All Funds	100%	0.03	0.02	0.56	1.00	0.21	0.02	0.01
		[0.07]			[81.56]	[10.74]	[1.19]	[0.94]
		After Fees						
1	1.83%	−0.74	−0.16	0.51	1.04	0.29	0.03	0.04
		[−0.79]			[46.34]	[8.78]	[0.87]	[2.10]
2	1.94%	0.33	0.08	0.58	1.01	0.20	0.06	0.03
		[0.40]			[59.59]	[5.89]	[2.05]	[1.55]
3	1.95%	0.89	0.20	0.64	1.00	0.23	0.10	0.05
		[0.95]			[43.51]	[7.09]	[2.32]	[2.50]
4	1.94%	1.25	0.26	0.64	1.03	0.26	0.01	0.07
		[1.29]			[47.08]	[6.80]	[0.22]	[2.54]
5	1.88%	2.24**	0.43	0.65	1.07	0.45	−0.10	0.08
		[2.05]			[39.23]	[13.50]	[−2.90]	[3.28]
1st Stage	9.55%	0.78	0.22	0.62	1.03	0.28	0.02	0.05
		[1.07]			[61.27]	[10.52]	[0.68]	[2.99]
All Funds	100%	−1.18**	−0.53	0.48	1.01	0.21	0.02	0.01
		[−2.39]			[81.97]	[10.70]	[1.20]	[0.99]

Table 1.4: Out-of-sample Performances of the Selected Funds

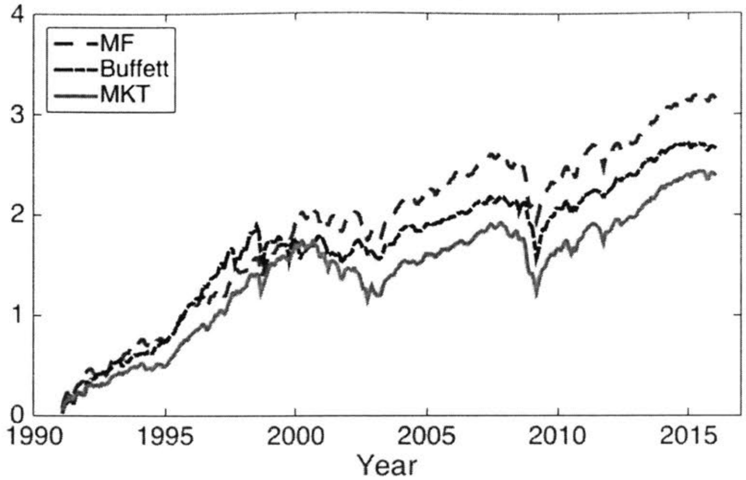
This table documents the out-of-sample performances of the funds whose returns first-order stochastically dominated the returns of the replica funds during the 24 months prior to portfolio formation. Specifically, by the end of each quarter, I compute the empirical CDF of the percentile of each fund in the counterfactual return distribution ($F^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}(x)$). I then construct the FSD test statistic $\hat{\theta}$ for each fund. I select the funds with $\hat{\theta} \geq 0.90$. I then sort the selected funds into 5 quintiles based on their proceeding 24 months' four-factor alpha. The trading strategy is rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. The alphas with statistical significance are marked with "*". "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

with the highest historical alpha within the funds selected by the FSD filter. The blue dashed(black dotted, red solid) line is the log cumulative before-fees return of the selected mutual funds(Berkshire Hathaway, the market). Figure 1.13 plots the time-series of the before-fees outperformances of the selected mutual funds. The outperformance is defined as the log cumulative return of longing the portfolio of identified funds and shorting the market. From the figure, fund outperformances seem to be most pronounced during the dot-com bubble periods, but the outperformances are in general consistent over the sample period.

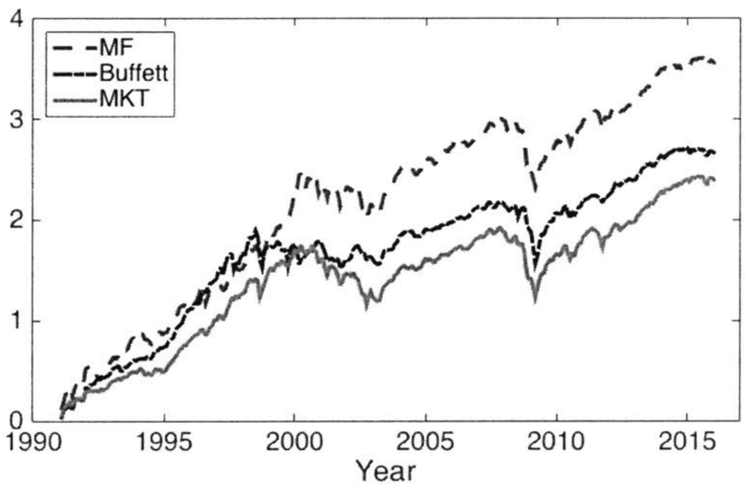
Figure 1.14 plots the histograms of the before-fees excess returns of the selected mutual funds. Panel (a) plots the histogram of the returns in excess of the Carhart four-factor benchmark for all funds selected by the FSD filter; whereas Panel (b) plots the histogram of the returns in excess of the Carhart four-factor benchmark for the top quintile of the funds with the highest historical alpha within the funds selected by the FSD filter. From the figure, it is obvious that the identified funds are more likely to realize positive excess returns than negative excess returns compared to the Carhart four-factor benchmark.

1.7.3 Fund Characteristics

Table 1.5 compares the observable characteristics of the identified funds with the cross-sectional average of all funds in the sample. From the table, the funds identified by the FSD filter do not differ in size compared to an average fund in the industry. However, the funds in higher second-stage quintiles charge more fees. The finding is partially consistent with Berk and Green (2004) in the sense that more skilled managers are able to extract higher rents in equilibrium, although the specific mechanism is different. In Berk and Green (2004), skilled managers receive compensation by growing the size of their funds, leaving the fees unchanged. My finding, on the other hand, suggests that they demand higher fees directly. Funds in higher second-stage quintiles also tend to keep fewer stocks in their portfolios, thereby being more concentrated. This finding is consistent with the theory proposed by Van Nieuwerburgh and Veldkamp (2010) that informed investors can choose to be specialized when information acquisition is costly. The finding that funds in higher quintiles also trade more is related to the finding by Pástor, Stambaugh, and Taylor (2017) that trades by active mutual fund managers tend to be profitable.

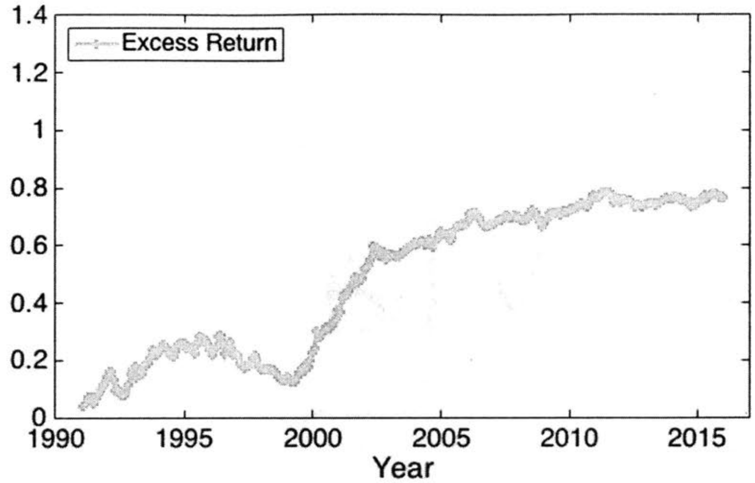


(a) 1st Stage

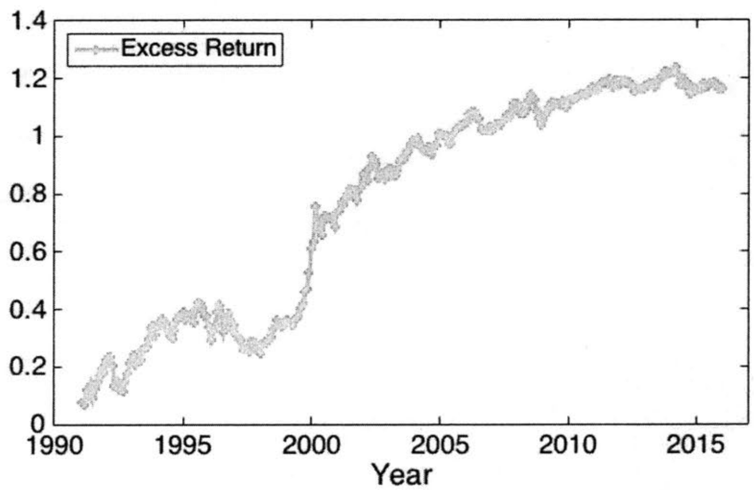


(b) Top $\hat{\alpha}$ Quintile, 2nd Stage

Figure 1.12: Out-of-sample Fund Performances: Time Series
This panel of plots documents the time-series of the before-fees performances of the selected mutual funds. Panel (a) plots the time-series of the fund performances for all funds selected by the FSD filter; whereas Panel (b) plots the top quintile of the funds with the highest historical alpha within the funds selected by the FSD filter. The blue dashed line is the log cumulative before-fees return of the selected mutual funds, the black dotted line is the log cumulative return of Berkshire Hathaway, the red solid line is the log cumulative return of the market.

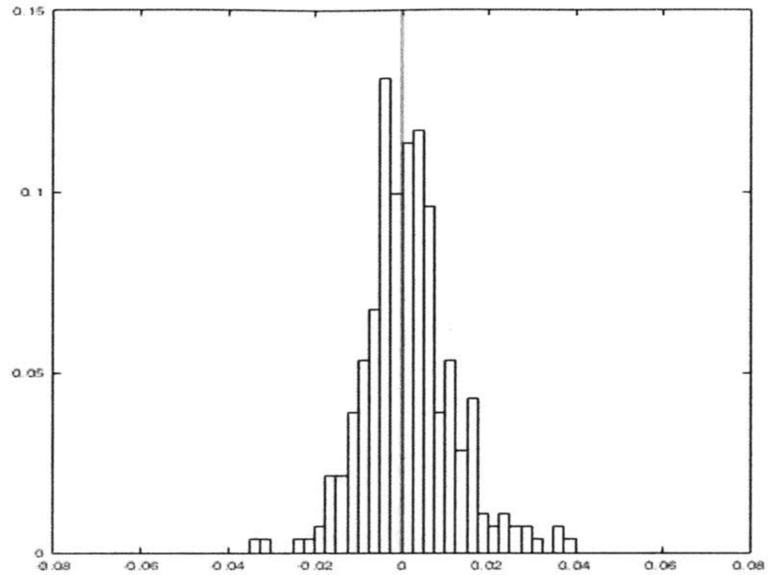


(a) 1st Stage

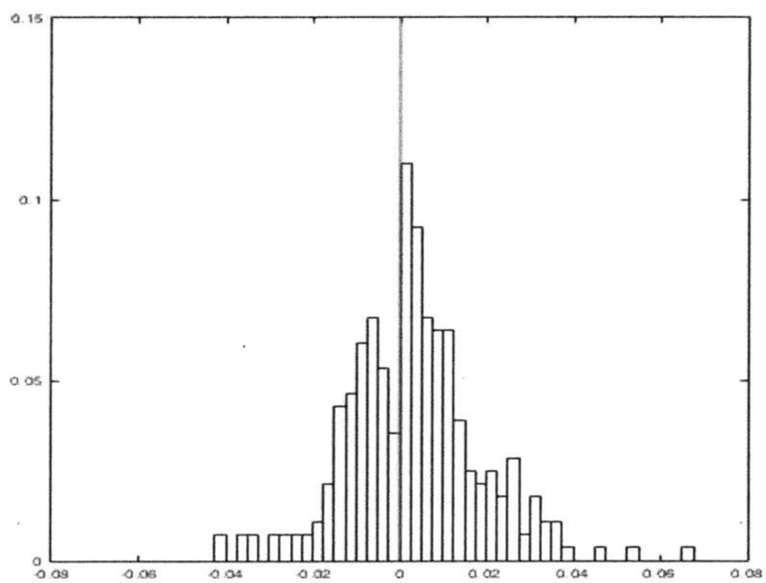


(b) Top $\hat{\alpha}$ Quintile, 2nd Stage

Figure 1.13: Out-of-sample Fund Outperformances: Time Series
This panel of plots documents the time-series of the before-fees outperformances of the selected mutual funds. Outperformance is defined as the log cumulative return of the trading strategy that longs the portfolio of the identified funds and shorts the market. Panel (a) plots the time-series of the fund outperformances for all funds selected by the FSD filter; whereas Panel (b) plots the top quintile of the funds with the highest historical alpha within the funds selected by the FSD filter.



(a) 1st Stage



(b) Top $\hat{\alpha}$ Quintile, 2nd Stage

Figure 1.14: Out-of-sample Fund Excess Returns: Histogram
This panel of plots documents the histograms of the before-fees Carhart excess returns of the selected mutual funds. Panel (a) plots the histogram for all funds selected by the FSD filter; whereas Panel (b) plots the histogram of the top quintile of the funds with the highest historical alpha within the funds selected by the FSD filter.

Quintile	Sample Share	Age	Age Norm.	TNA Norm.	# of Stocks Norm.	Fees (in bps)	Fees Norm.	Turnover Ratio	Turnover Norm.
1	1.83%	15.27	1.00	0.73	0.99	120.42	1.05	0.81	0.98
2	1.94%	16.41	1.08	1.11	1.10	117.23	1.02	0.69	0.84
3	1.95%	15.20	1.00	0.97	1.03	119.01	1.04	0.70	0.86
4	1.94%	15.92	1.05	1.06	0.97	125.43	1.09	0.73	0.88
5	1.88%	15.80	1.05	1.01	0.80	127.27	1.11	0.80	0.97
1st Stage	9.55%	15.73	1.04	0.98	0.99	121.94	1.06	0.75	0.92
All Funds	100%	15.18	1	1	1	114.78	1	0.82	1

Table 1.5: Characteristics of the Selected Funds

This table documents the characteristics of the selected funds and compare them with the cross-sectional average. "Norm." denotes the normalization procedure that takes the ratio between the corresponding variable and the cross-sectional average when all funds in the sample are included. "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

1.7.4 The Return Gap

One potential concern regarding the outperformances of the identified funds is that instead of possessing stock-picking skills, those funds might be loading on momentum factors that the Carhart benchmark does not perfectly control for. In order to rule out such possibility, I study the return gaps of the identified funds. The return gap is defined as the difference between a fund's actual return from the hypothetical return that the fund might have earned by keeping the portfolio weights at the beginning of the quarter unchanged throughout the entire quarter:

$$rgap_{i,t} \equiv r_{i,t} - \sum_j w_{i,j,\underline{t}} \tilde{r}_{j,t}$$

where $w_{i,j,\underline{t}}$ denotes the portfolio weight of stock j of fund i at the most recent quarter-end of month t; $\tilde{r}_{j,t}$ denotes the return of stock j during month t.

The return gap measures the profitability of the unobserved within-quarter actions conducted by a fund manager. I regress the return gaps of the identified funds against the Carhart four-factor benchmark:

$$rgap_{i,t} = \alpha_i^{rgap} + \beta_{m,i}^{rgap} (r_{m,t} - r_f) + \beta_{smb,i}^{rgap} smb_t + \beta_{smb,i}^{rgap} smb_t + \beta_{hml,i}^{rgap} hml_t + \beta_{umd,i}^{rgap} umd_t + \epsilon_{i,t}^{rgap}.$$

The results are documented in Table 1.6. The table shows that the out-of-sample alphas resulting from the return gaps also increase monotonically with historical alphas for funds selected by the FSD filter. Moreover, the return gaps account for about half of the total out-of-sample alphas for all quintiles. The finding is consistent with the results of Kacperczyk, Sialm, and Zheng (2008) that the return gap is indicative of manager skills. The profitability of the return gap offers strong support that the identified managers are skilled because they are able to make profitable within-quarter trades. It rules out the concern that the outperformances of the identified managers are entirely driven by their loadings on some uncontrolled momentum factors.

Quintile	Sample Share	α^{rgap} (in %)	mkt	smb	hml	umd
1	1.83%	0.74**	0.02	0.03	−0.01	0.02
		[2.31]	[2.31]	[1.73]	[−1.30]	[3.08]
2	1.94%	0.77***	0.02	0.02	−0.01	0.01
		[3.94]	[4.26]	[1.61]	[−0.91]	[2.83]
3	1.95%	0.91***	0.03	0.01	−0.00	0.01
		[3.64]	[4.76]	[0.67]	[−0.23]	[2.54]
4	1.94%	0.85***	0.02	0.01	−0.01	0.01
		[3.29]	[3.26]	[0.99]	[−0.50]	[2.01]
5	1.88%	1.60***	0.02	−0.02	−0.03	0.02
		[3.04]	[1.47]	[−0.63]	[−1.42]	[2.43]
1st Stage	9.55%	0.97***	0.02	0.01	-0.01	0.01
		[5.19]	[4.10]	[1.07]	[-1.11]	[3.49]
All Funds	100%	0.21	0.00	0.02	0.00	0.02
		[1.26]	[0.70]	[1.46]	[0.14]	[3.60]

Table 1.6: The Return Gaps

This table documents the return gaps of the selected mutual funds. The return gap is defined as the difference between a mutual fund's actual return versus the hypothetical return generated by keeping the holdings within the mutual fund's portfolio by the end of the proceeding quarter. The time-series of the return gaps of different funds are then averaged within the corresponding quintiles and regressed against the Carhart four factors. The portfolios are rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. The alphas with statistical significance are marked with "*". "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

1.7.5 The Copycat Strategy

The return gap analysis suggests that the identified fund managers are able to generate profits from their unobserved within-quarter actions. Therefore, it should also suggest that it would be difficult for an out-sider to free-ride on those managers stock-picking endeavors. Indeed, Table 1.7 documents the performance of the trading strategy that aims to mimic the performances of the selected outperforming mutual fund managers. To ensure implementability, by the end of each quarter, the stock holdings from the end of the previous quarter are retrieved for the managers who have been identified by the FSD filter. The trading strategy then invests in the stocks that the managers were holding as of the end of the previous quarter. The portfolios are rebalanced

Quintile	Sample Share	α (in %)	IR	SR	mkt	smb	hml	umd
1	1.83%	0.08	0.02	0.52	1.06	0.27	0.06	−0.05
		[0.07]			[38.95]	[6.39]	[1.47]	[−2.11]
2	1.94%	1.50*	0.32	0.60	1.07	0.15	0.05	−0.04
		[1.68]			[54.40]	[3.10]	[1.74]	[−2.48]
3	1.95%	1.65	0.32	0.64	1.06	0.23	0.13	−0.00
		[1.60]			[38.82]	[5.30]	[2.48]	[−0.13]
4	1.94%	1.52	0.29	0.61	1.09	0.28	0.01	0.01
		[1.46]			[45.45]	[6.00]	[0.21]	[0.32]
5	1.88%	2.02*	0.36	0.60	1.15	0.43	−0.10	−0.00
		[1.77]			[37.36]	[9.76]	[−2.74]	[−0.05]
1st Stage	9.55%	1.30	0.32	0.60	1.09	0.27	0.03	−0.02
		[1.62]			[53.87]	[6.74]	[0.93]	[-1.01]
All Funds	100%	0.20	0.07	0.51	1.09	0.18	0.00	−0.08
		[0.34]			[66.63]	[5.32]	[0.05]	[-6.70]

Table 1.7: Performance of the Copycat Strategy
This table documents the performance of the trading strategy that aims to mimic the performances of the selected outperforming mutual funds. To ensure implementability, by the end of each quarter, the stock holdings from the end of the previous quarter are retrieved for the managers who have been identified as skilled. The trading strategy then invests in the stocks that the managers were holding as of the end of the previous quarter. The portfolios are rebalanced every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics. The alphas with statistical significance are marked with "*". "Sample Share" is the number of funds in the portfolio as a percentage of the cross section. The sample period is from January 1991 to December 2015.

every three months. The post-ranking annualized alphas and factor loadings are documented along with their heteroscedasticity-robust t-statistics.

Consistent with the analysis on the return gap, the copycat strategy loses about half of the profitability compare to the managers' total before-fees performances. Interestingly, the profitability of the copycat strategy is comparable to that of the after-fees returns that investors are able to earn by investing in the funds. The finding suggests that the fees of the identified funds might be set rationally in equilibrium. This result verifies the findings documented by Frank et al. (2004) for a limited sample of high-expense funds.

1.7.6 Fund Flow Responses

According to Berk and van Binsbergen (2015) and Barber, Huang, and Odean (2016), fund flows contain information about fund investors' evaluations of managers' investment skills. In order to understand whether fund investors infer managers' skills using signals correlated with the FSD condition, I run the following Fama-Macbeth regression:

$$Flow_{i,t} = Const + \delta_0 \times FSD_{i,t} + (\beta + \delta_1 \times FSD_{i,t}) \times \hat{\alpha}_i^{[t-1-T,t-1]} + X_i + \epsilon_{i,t},$$

where X_i represents control variables including fund age, log fund size, fees, and the number of stocks in the portfolio; $\hat{\alpha}_i^{[t-1-T,t-1]}$ is the trailing in-sample realized Carhart four-factor alpha; $FSD_{i,t}$ is a dummy variable that equals to one for funds with a FSD test statistic $\hat{\theta}$ higher than 0.90, and zero otherwise.

The parameters of interest are δ_0 and δ_1. A positive δ_0 indicates that controlling for realized in-sample alphas, the funds that satisfy the FSD condition attract more flows than other funds. A positive δ_1 indicates that flows are more sensitive to realized in-sample alphas for funds that satisfy the FSD condition.

Table 1.8 presents the regression results. Both δ_0 and δ_1 are highly significantly positive. The finding suggests that fund investors reveal preferences towards fund return distribution properties beyond the first moment(mean/alpha). They appreciate funds with return distributions satisfying the FSD condition more than other funds controlling for realized alpha. On the other hand, the still positive out-of-sample alphas of the FSD identified funds suggest the presence of certain informational frictions in the fund market so that the fund outperformances are not fully arbitraged away according to the logic of Berk and Green (2004).

1.8 Conclusion

Due to the strong influence of luck in fund managers' performances, the search for skilled managers with predictable outperformances is a challenging task. Existing alpha based evaluation methods have poor out-of-sample performances because alpha is related to the mean of the return distribution and mean is difficult to estimate in short samples. I show that, by limiting the search scope to a specific subset of skilled

	$Flow_{i,t}$	$Flow_{i,t}$	$Flow_{i,t}$	$Flow_{i,t}$
$\hat{\alpha}_i^{[t-1-T,t-1]}$	2.75***	2.68***	2.66***	2.64***
	[41.26]	[40.59]	[40.58]	[40.45]
$FSD_{i,t}$		0.0063***		0.0031***
		[11.42]		[5.22]
$FSD_{i,t} \times \hat{\alpha}_i^{[t-1-T,t-1]}$			1.29***	0.84***
			[10.06]	[5.74]

Table 1.8: Flow Responses

This table documents the Fama-Macbeth regression results of $Flow_{i,t} = Const + \delta_0 \times FSD_{i,t} + (\beta + \delta_1 \times FSD_{i,t}) \times \hat{\alpha}_i^{[t-1-T,t-1]} + X_i + \epsilon_{i,t}$. The dependent variable is the flow of each fund in every month, and the independent variables include the trailing in-sample alpha, the dummy variable corresponding to the FSD condition, and the interaction between the two. The control variables include fund age, log fund size, fees, and the number of stocks in the portfolio. The sample period is from January 1991 to December 2015, with 399,631 observations.

managers – the skilled stock-pickers, a new first-order stochastic dominance condition can be imposed to improve the effectiveness of the search. The new FSD filter complements the conventional $\hat{\alpha}$ sort because it is robust to finite-sample problems such as heteroscedasticity and benchmark mis-specification. The empirical part of this project demonstrates the superior performance of the combination of the new FSD filter and the standard $\hat{\alpha}$ sort in identifying outperforming stock-pickers. My findings confirm various theoretical and empirical results discussed earlier in the literature and are also able to shed new light on our understanding about the active mutual fund industry.

References

Agarwal, Vikas, Kevin A Mullally, Yuehua Tang, and Baozhong Yang. 2015. "Mandatory portfolio disclosure, stock liquidity, and mutual fund performance." *The Journal of Finance* 70 (6):2733–2776.

Alexander, Gordon J, Gjergji Cici, and Scott Gibson. 2006. "Does motivation matter when assessing trade performance? An analysis of mutual funds." *The Review of Financial Studies* 20 (1):125–150.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang. 2009. "High idiosyncratic volatility and low returns: International and further US evidence." *Journal of Financial Economics* 91 (1):1–23.

Baker, Malcolm, Lubomir Litov, Jessica A Wachter, and Jeffrey Wurgler. 2010. "Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements." *Journal of Financial and Quantitative Analysis* 45 (5):1111–1131.

Barber, Brad M, Xing Huang, and Terrance Odean. 2016. "Which factors matter to investors? Evidence from mutual fund flows." *The Review of Financial Studies* 29 (10):2600–2642.

Barras, Laurent, Olivier Scaillet, and Russ Wermers. 2010. "False discoveries in mutual fund performance: Measuring luck in estimated alphas." *The journal of finance* 65 (1):179–216.

Berk, Jonathan B and Richard C Green. 2004. "Mutual Fund Flows and Performance in Rational Markets." *Journal of Political Economy* 112 (6):1269–1295.

Berk, Jonathan B and Jules H van Binsbergen. 2015. "Measuring skill in the mutual fund industry." *Journal of Financial Economics* 118 (1):1–20.

———. 2016. "Assessing asset pricing models using revealed preference." *Journal of Financial Economics* 119 (1):1–23.

Brown, Stephen J and William N Goetzmann. 1995. "Performance persistence." *The Journal of finance* 50 (2):679–698.

Carhart, Mark M. 1997. "On persistence in mutual fund performance." *The Journal of finance* 52 (1):57–82.

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers. 2000. "The value of active mutual fund management: An examination of the stockholdings and trades of fund managers." *Journal of Financial and quantitative Analysis* 35 (3):343–368.

Chevalier, Judith and Glenn Ellison. 1997. "Risk taking by mutual funds as a response to incentives." *Journal of Political Economy* 105 (6):1167–1200.

Cohen, Randolph B, Joshua D Coval, and L'uboš Pástor. 2005. "Judging fund managers by the company they keep." *The Journal of Finance* 60 (3):1057–1096.

Cremers, KJ Martijn and Antti Petajisto. 2009. "How active is your fund manager? A new measure that predicts performance." *Review of Financial Studies* 22 (9):3329–3365.

Da, Zhi, Pengjie Gao, and Ravi Jagannathan. 2010. "Impatient trading, liquidity provision, and stock selection by mutual funds." *The Review of Financial Studies* 24 (3):675–720.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers. 1997. "Measuring mutual fund performance with characteristic-based benchmarks." *The Journal of finance* 52 (3):1035–1058.

Doshi, Hitesh, Redouane Elkamhi, and Mikhail Simutin. 2015. "Managerial activeness and mutual fund performance." *Review of Asset Pricing Studies* 5 (2):156–184.

———. 2010. "Luck versus skill in the cross-section of mutual fund returns." *The journal of finance* 65 (5):1915–1947.

Frank, Mary Margaret, James M Poterba, Douglas A Shackelford, and John B Shoven. 2004. "Copycat funds: Information disclosure regulation and the returns to active management in the mutual fund industry." *The Journal of Law and Economics* 47 (2):515–541.

Fu, Fangjian. 2009. "Idiosyncratic risk and the cross-section of expected stock returns." *Journal of Financial Economics* 91 (1):24–37.

Goetzmann, William N and Roger G Ibbotson. 1994. "Do winners repeat?" *The Journal of Portfolio Management* 20 (2):9–18.

Grinblatt, Mark and Sheridan Titman. 1989. "Mutual fund performance: An analysis of quarterly portfolio holdings." *Journal of business* :393–416.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers. 1995. "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior." *The American economic review* :1088–1105.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser. 1993. "Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988." *The Journal of finance* 48 (1):93–130.

Huang, Jennifer, Clemens Sialm, and Hanjiang Zhang. 2011. "Risk shifting and mutual fund performance." *The Review of Financial Studies* 24 (8):2575–2616.

Iskoz, Sergey and Jiang Wang. 2003. "How to Tell If a Money Manager Knows More?" Tech. rep., National Bureau of Economic Research.

Jegadeesh, Narasimhan and Sheridan Titman. 1993. "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of finance* 48 (1):65–91.

Jensen, Michael C. 1968. "The performance of mutual funds in the period 1945–1964." *The Journal of finance* 23 (2):389–416.

Jiang, George J, Tong Yao, and Tong Yu. 2007. "Do mutual funds time the market? Evidence from portfolio holdings." *Journal of Financial Economics* 86 (3):724–758.

Kacperczyk, Marcin and Amit Seru. 2007. "Fund manager use of public information: New evidence on managerial skills." *The Journal of Finance* 62 (2):485–528.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. 2005. "On the industry concentration of actively managed equity mutual funds." *The Journal of Finance* 60 (4):1983–2011.

———. 2008. "Unobserved actions of mutual funds." *Review of Financial Studies* 21 (6):2379–2416.

Kacperczyk, Marcin, Stijn van Nieuwerburgh, and Laura Veldkamp. 2014. "Time-varying fund manager skill." *The Journal of Finance* 69 (4):1455–1484.

Koijen, Ralph SJ. 2014. "The Cross-Section of Managerial Ability, Incentives, and Risk Preferences." *The Journal of Finance* 69 (3):1051–1098.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White. 2006. "Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis." *The Journal of finance* 61 (6):2551–2595.

Linnainmaa, Juhani T. 2013. "Reverse survivorship bias." *The Journal of Finance* 68 (3):789–813.

Malkiel, Burton G. 1995. "Returns from investing in equity mutual funds 1971 to 1991." *The Journal of finance* 50 (2):549–572.

Malkiel, Burton G and Eugene F Fama. 1970. "Efficient capital markets: A review of theory and empirical work." *The journal of Finance* 25 (2):383–417.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor. 2017. "Do funds make more when they trade more?" *The Journal of Finance* 72 (4):1483–1528.

Ross, Stephen A. 1976. "The arbitrage theory of capital asset pricing." *Journal of economic theory* 13 (3):341–360.

Sharpe, William F. 1966. "Mutual fund performance." *The Journal of business* 39 (1):119–138.

Treynor, Jack and Kay Mazuy. 1966. "Can mutual funds outguess the market." *Harvard business review* 44 (4):131–136.

———. 2010. "Information acquisition and under-diversification." *The Review of Economic Studies* 77 (2):779–805.

———. 2000. "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses." *The Journal of Finance* 55 (4):1655–1695.

———. 2003. "Is money really'smart'? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence." .

Appendix

1.A Proposition Proofs

Proposition 1.

Proof. According to Assumption 3, $\forall k,\ j \quad \beta_{k,j} = \beta_{\hat{k},j}$. Therefore, $\beta_{i,j,t} = \sum_K w_{i,k,t-1}\beta_{k,j}$ $= \sum_K w_{i,k,t-1}\beta_{\hat{k},j} = \hat{\beta}_{i,j,t}$, i.e. the original portfolio and the replica portfolio have the same loadings on observable factors. □

Proposition 2.

Proof. Denote

$$\begin{aligned}
\delta_{\hat{k},l} &\equiv \gamma_{\hat{k}.l} - \mathbb{E}\left(\gamma_{\hat{k}.l} \middle| \left\{\beta_{\hat{k},j}\right\}\right) \\
&= \gamma_{\hat{k}.l} - \mathbb{E}\left(\gamma_{k.l} \middle| \left\{\beta_{k,j}\right\}\right) \\
&= \gamma_{\hat{k}.l} - \bar{\gamma}_{k,l}
\end{aligned}$$

$\delta_{\hat{k},l}$ is the deviation of stock $\hat{k}$'s loading on factor l from the average loading on factor l of the stocks in the same bucket as $\hat{k}$. Since stock $\hat{k}$ is chosen randomly within the bucket, $\mathbb{E}\left(\delta_{\hat{k},l} \middle| \left\{\beta_{k,j}\right\}\right)$[14].

[14] Here, the expectation is taken before the fund is constructed.

The second line comes from the fact that $\beta_{\hat{k},j} = \beta_{k,j}$.

$$\begin{aligned}
\hat{\gamma}_{i,l,t} &= \sum_k w_{i,k,t-1}\gamma_{\hat{k},l} \\
&= \sum_k w_{i,k,t-1}\left(\bar{\gamma}_{k,l} + \delta_{\hat{k},l}\right) \\
&= \sum_k w_{i,k,t-1}\bar{\gamma}_{k,l} + \sum_k w_{i,k,t-1}\delta_{\hat{k},l} \\
&\overset{LLN}{\approx} \sum_k w_{i,k,t-1}\bar{\gamma}_{k,l}
\end{aligned}$$

The last line is given by Law of Large Number. It holds when the portfolio is sufficiently diversified. □

Proposition 3.

Proof. According to Central Limit Theorem,

$$\begin{aligned}
e_{i,t}|\left\{w_{i,k,t-1}\right\} &\sim N\left(\alpha_{i,t}, Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t}|\left\{w_{i,k,t-1}\right\}\right)\right) \\
\hat{e}_{i,t}|\left\{w_{i,k,t-1}\right\} &\sim N\left(0, Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}|\left\{w_{i,\hat{k},t-1}\right\}\right)\right)
\end{aligned}$$

All we need to show is that $Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t}|\left\{w_{i,k,t-1}\right\}\right)$ $\approx Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t}|\left\{w_{i,\hat{k},t-1}\right\}\right)$. Notice that the real fund and the replica fund are picking stocks from the same stock volatility buckets, so that

$$\begin{aligned}
&Var\left(\tilde{r}_{k,t}\right) = Var\left(\tilde{r}_{\hat{k},t}\right) \\
\iff &Var\left(\sum_l \gamma_{k,l}f_{l,t} + \epsilon_{k,t}\right) = Var\left(\sum_l \gamma_{\hat{k},l}f_{l,t} + \epsilon_{\hat{k},t}\right)
\end{aligned}$$

On the other hand $\gamma_{k,l} = \bar{\gamma}_{k,l} + \delta_{k,l}$ and $\gamma_{\hat{k},l} = \bar{\gamma}_{k,l} + \delta_{\hat{k},l}$, so that the difference between $\gamma_{k,l}$ and $\gamma_{\hat{k},l}$ could be diversified away. Therefore,

$$
\begin{aligned}
& Var\left(\sum_k w_{i,k,t-1}\epsilon_{k,t} \middle| \{w_{i,k,t-1}\}\right) \\
=& \sum_k w_{i,k,t-1}^2 Var\left(\epsilon_{k,t}\right) \\
=& \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{k,l} f_{l,t} + \epsilon_{k,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right) \\
=& \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t} + \epsilon_{\hat{k},t}\right) - Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right)\right) \\
&+ \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right) \\
=& Var\left(\sum_k w_{i,k,t-1}\epsilon_{\hat{k},t} \middle| \left\{w_{i,\hat{k},t-1}\right\}\right) \\
&+ \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
& \sum_k w_{i,k,t-1}^2 \left(Var\left(\sum_l \gamma_{\hat{k},l} f_{l,t}\right) - Var\left(\sum_l \gamma_{k,l} f_{l,t}\right)\right) \\
=& \sum_k w_{i,k,t-1}^2 \left(\sum_l \left(\gamma_{\hat{k},l}^2 - \gamma_{\hat{k},l}^2\right) Var\left(f_{l,t}\right)\right) \\
\approx& \sum_k w_{i,k,t-1}^2 \left(\sum_l 2\bar{\gamma}_{k,l} \left(\delta_{\hat{k},l} - \delta_{k,l}\right) Var\left(f_{l,t}\right)\right) \\
=& 2\sum_l \left(\sum_k w_{i,k,t-1}^2 \bar{\gamma}_{k,l} \left(\delta_{\hat{k},l} - \delta_{k,l}\right)\right) Var\left(f_{l,t}\right) \\
\approx& \frac{2}{K}\sum_l \left(\sum_k w_{i,k,t-1} \bar{\gamma}_{k,l} \left(\delta_{\hat{k},l} - \delta_{k,l}\right)\right) Var\left(f_{l,t}\right) \\
\overset{LLN}{\approx}& 0
\end{aligned}
$$

$\square$

Proposition 6.

Proof. Since all replica funds are constructed randomly in the bootstrap procedure, $Pct\left(\hat{r}_{i,t}, \langle\hat{r}_{i,t}\rangle\right) \sim Unif\left(0,1\right)$ is obvious.

The equivalence condition is immediate from $Pct\left(\cdot, \langle\hat{r}_{i,t}\rangle\right)$ being monotonically increasing.

Denote $F_{t-1}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}\left(x\right)(F_{t-1}^{Pct(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle)}\left(x\right))$ as the conditional CDF of the ranking $r_{i,t}(\hat{r}_{i,t})$; $\langle\hat{r}_{i,t}\rangle\left[x\right]$ as the value at the x percentile of $\langle\hat{r}_{i,t}\rangle$.

$$\begin{aligned}
F_{t-1}^{Pct(r_{i,t},\langle\hat{r}_{i,t}\rangle)}\left(x\right) &\equiv Prob_{t-1}\left(Pct\left(r_{i,t}, \langle\hat{r}_{i,t}\rangle\right) \leq x\right) \\
&= Prob_{t-1}\left(r_{i,t} \leq \langle\hat{r}_{i,t}\rangle\left[x\right]\right) \\
&\equiv F_{t-1}^{r_{i,t}}\left(\langle\hat{r}_{i,t}\rangle\left[x\right]\right) \\
&< F_{t-1}^{\hat{r}_{i,t}}\left(\langle\hat{r}_{i,t}\rangle\left[x\right]\right) \\
&= Prob_{t-1}\left(\hat{r}_{i,t} \leq \langle\hat{r}_{i,t}\rangle\left[x\right]\right) \\
&= Prob_{t-1}\left(Pct\left(\hat{r}_{i,t}, \langle\hat{r}_{i,t}\rangle\right) \leq x\right) \\
&\equiv F_{t-1}^{Pct(\hat{r}_{i,t},\langle\hat{r}_{i,t}\rangle)}\left(x\right)
\end{aligned}$$

□

Proposition 9.

Proof. $r_{i,t} \overset{fsd}{\succ} \hat{r}_{i,t} \iff \tilde{f}_{i,t} + e_{i,t} \overset{fsd}{\succ} \hat{e}_{i,t}$.

Since $\tilde{f}_{i,t} \perp e_{i,t}$, $\tilde{f}_{i,t} + e_{i,t} \sim N\left(\mu_f, \sigma_f^2 + \sigma_i^2\right)$.

Denote the PDF of $\tilde{f}_{i,t} + e_{i,t}$ as $\phi\left(x\right) \equiv \frac{1}{\sqrt{2\pi\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[-\frac{\left(x-\mu_f\right)^2}{2\left(\sigma_f^2+\sigma_i^2\right)}\right]$; and the PDF of $\hat{e}_{i,t}$ as $\hat{\phi}\left(x\right) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{x^2}{2\sigma_i^2}\right]$.

Define

$$
\begin{aligned}
L(x) &\equiv \frac{\phi(x)}{\hat{\phi}(x)} \\
&= \sqrt{\frac{\sigma_i^2}{\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[\frac{x^2}{2\sigma_i^2} - \frac{(x-\mu_f)^2}{2\left(\sigma_f^2+\sigma_i^2\right)}\right] \\
&= \sqrt{\frac{\sigma_i^2}{\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[\frac{\left(\sigma_f^2+\sigma_i^2\right)x^2 - \sigma_i^2 (x-\mu_f)^2}{2\sigma_i^2\left(\sigma_f^2+\sigma_i^2\right)}\right] \\
&= \sqrt{\frac{\sigma_i^2}{\left(\sigma_f^2+\sigma_i^2\right)}} \exp\left[\frac{\sigma_f^2 x^2 + 2\sigma_i^2\mu_f x - \sigma_i^2\mu_f^2}{2\sigma_i^2\left(\sigma_f^2+\sigma_i^2\right)}\right].
\end{aligned}
$$

If $\sigma_f > 0$, then $\lim_{x\to-\infty} L(x) = +\infty$. Therefore, $\tilde{f}_{i,t} + e_{i,t}$ has a larger left tail compared to $\hat{e}_{i,t}$, and the first-order stochastic dominance condition is violated. □

1.B Additional Sorts

I document the results of several additional sorts in this section. The following tables report the out-of-sample performances of the double sort with the FSD test statistic $\hat{\theta}$ as the first variable, and $\hat{\alpha}$, $t-stat$ of $\hat{\alpha}$, information ratio, historical average of fund rankings $(Pct\,(r_{i,t}, \langle \hat{r}_{i,t} \rangle))$ as the second variable, respectively.

Decile(1st)	Quintile(2nd)	1	2	3	4	5
1	$\hat{\alpha}$	−2.30**	0.02	0.08	0.14	0.50
1	$t-stat$	[−2.29]	[0.03]	[0.10]	[0.19]	[0.64]
1	IR	−0.48	0.01	0.02	0.04	0.13
2	$\hat{\alpha}$	−1.65*	0.65	−0.86	−0.16	0.17
2	$t-stat$	[−1.84]	[0.94]	[−1.26]	[−0.22]	[0.21]
2	IR	−0.37	0.20	−0.26	−0.04	0.04
3	$\hat{\alpha}$	−0.43	−0.51	−0.56	0.06	0.15
3	$t-stat$	[−0.54]	[−0.65]	[−0.84]	[0.09]	[0.20]
3	IR	−0.11	−0.14	−0.17	0.02	0.04
4	$\hat{\alpha}$	−1.39	−0.54	0.08	0.00	0.06
4	$t-stat$	[−1.64]	[−0.87]	[0.12]	[−0.01]	[0.08]
4	IR	−0.35	−0.17	0.03	0.00	0.02
5	$\hat{\alpha}$	−1.97**	−1.09	−0.50	−0.70	−0.06
5	$t-stat$	[−2.23]	[−1.61]	[−0.76]	[−1.18]	[−0.07]
5	IR	−0.50	−0.32	−0.17	−0.24	−0.01
6	$\hat{\alpha}$	−1.26	0.68	−1.14*	0.67	−0.47
6	$t-stat$	[−1.41]	[1.05]	[−1.83]	[1.04]	[−0.62]
6	IR	−0.32	0.22	−0.38	0.23	−0.12
7	$\hat{\alpha}$	0.07	−0.08	−0.95*	0.04	0.64
7	$t-stat$	[0.09]	[−0.11]	[−1.71]	[0.06]	[0.75]
7	IR	0.02	−0.02	−0.35	0.01	0.15
8	$\hat{\alpha}$	0.60	−0.48	−0.18	0.60	0.15
8	$t-stat$	[0.69]	[−0.63]	[−0.28]	[0.80]	[0.16]
8	IR	0.14	−0.13	−0.06	0.17	0.03
9	$\hat{\alpha}$	0.21	0.18	−0.23	1.10	0.80
9	$t-stat$	[0.26]	[0.24]	[−0.31]	[1.59]	[0.83]
9	IR	0.06	0.05	−0.07	0.32	0.17
10	$\hat{\alpha}$	0.68	1.98**	2.17**	3.44***	2.97***
10	$t-stat$	[0.74]	[2.34]	[2.41]	[3.64]	[2.84]
10	IR	0.15	0.48	0.50	0.71	0.59

Table 1.9: Double Sort: $\hat{\theta}$ and $\hat{\alpha}$

This table documents the out-of-sample before-fees performance of the double sort by the FSD test statistic $\hat{\theta}$ and the historical in-sample $\hat{\alpha}$. Active funds in the cross section are first sorted into 10 deciles by the FSD test statistic $\hat{\theta}$ constructed from the 24 monthly observations prior to portfolio formation. The funds in each decile are then further sorted into 5 quintiles by the realized Carhart four-factor alpha during the proceeding 24 months. The portfolios of funds are rebalanced every 3 months. Out-of-sample alphas of the portfolios along with the t-statistics and information ratios are reported. Out-of-sample alphas with statistical significance are marked with "*". The sample period is from January 1991 to December 2015.

Decile(1st)	Quintile(2nd)	1	2	3	4	5
1	$\hat{\alpha}$	−1.61**	−0.23	−0.18	−0.06	0.44
1	$t-stat$	[−1.98]	[−0.28]	[−0.22]	[−0.07]	[0.54]
1	IR	−0.40	−0.06	−0.05	−0.02	0.11
2	$\hat{\alpha}$	−1.54**	0.11	−0.03	−0.79	0.47
2	$t-stat$	[−2.03]	[0.14]	[−0.04]	[−1.01]	[0.65]
2	IR	−0.40	0.03	−0.01	−0.21	0.14
3	$\hat{\alpha}$	−0.75	−0.87	−0.09	−0.41	0.78
3	$t-stat$	[−1.08]	[−1.17]	[−0.12]	[−0.53]	[1.12]
3	IR	−0.22	−0.25	−0.03	−0.12	0.23
4	$\hat{\alpha}$	−0.95	−0.70	−0.62	0.13	0.42
4	$t-stat$	[−1.38]	[−0.99]	[−0.92]	[0.16]	[0.61]
4	IR	−0.29	−0.21	−0.20	0.04	0.12
5	$\hat{\alpha}$	−1.41*	−1.33*	−1.27**	0.16	−0.30
5	$t-stat$	[−1.89]	[−1.77]	[−1.97]	[0.27]	[−0.42]
5	IR	−0.40	−0.37	−0.41	0.06	−0.09
6	$\hat{\alpha}$	−1.19	0.26	−1.08	0.29	0.12
6	$t-stat$	[−1.41]	[0.38]	[−1.46]	[0.45]	[0.18]
6	IR	−0.31	0.08	−0.32	0.10	0.04
7	$\hat{\alpha}$	−0.34	−0.22	−0.88	0.84	0.13
7	$t-stat$	[−0.48]	[−0.26]	[−1.13]	[1.25]	[0.18]
7	IR	−0.10	−0.06	−0.23	0.26	0.04
8	$\hat{\alpha}$	−0.21	−0.44	−0.12	0.93	0.52
8	$t-stat$	[−0.26]	[−0.58]	[−0.16]	[1.29]	[0.67]
8	IR	−0.05	−0.12	−0.03	0.26	0.13
9	$\hat{\alpha}$	0.37	−0.27	0.24	0.53	1.17*
9	$t-stat$	[0.44]	[−0.34]	[0.29]	[0.67]	[1.66]
9	IR	0.09	−0.07	0.06	0.14	0.33
10	$\hat{\alpha}$	1.12	1.74**	2.81***	2.24**	3.41***
10	$t-stat$	[1.17]	[2.00]	[3.04]	[2.54]	[4.37]
10	IR	0.25	0.41	0.60	0.51	0.87

Table 1.10: Double Sort: $\hat{\theta}$ and $t-stat$

This table documents the out-of-sample before-fees performance of the double sort by the FSD test statistic $\hat{\theta}$ and the t-statistic of historical in-sample $\hat{\alpha}$. Active funds in the cross section are first sorted into 10 deciles by the FSD test statistic $\hat{\theta}$ constructed from the 24 monthly observations prior to portfolio formation. The funds in each decile are then further sorted into 5 quintiles by the t-statistic of the realized Carhart four-factor alpha during the proceeding 24 months. The portfolios of funds are rebalanced every 3 months. Out-of-sample alphas of the portfolios along with the t-statistics and information ratios are reported. Out-of-sample alphas with statistical significance are marked with "*". The sample period is from January 1991 to December 2015.

Decile(1st)	Quintile(2nd)	1	2	3	4	5
1	$\hat{\alpha}$	−1.64**	−0.11	−0.34	0.07	0.40
1	$t-stat$	[−1.98]	[−0.13]	[−0.41]	[0.08]	[0.49]
1	IR	−0.40	−0.03	−0.09	0.02	0.10
2	$\hat{\alpha}$	−1.56**	0.10	−0.03	−0.78	0.44
2	$t-stat$	[−2.05]	[0.13]	[−0.04]	[−1.01]	[0.61]
2	IR	−0.40	0.03	−0.01	−0.21	0.13
3	$\hat{\alpha}$	−0.76	−0.83	−0.05	−0.39	0.76
3	$t-stat$	[−1.08]	[−1.12]	[−0.08]	[−0.51]	[1.11]
3	IR	−0.22	−0.24	−0.02	−0.11	0.23
4	$\hat{\alpha}$	−1.05	−0.60	−0.66	0.09	0.42
4	$t-stat$	[−1.54]	[−0.85]	[−0.98]	[0.11]	[0.61]
4	IR	−0.33	−0.18	−0.21	0.03	0.13
5	$\hat{\alpha}$	−1.35*	−1.41*	−1.22*	0.08	−0.30
5	$t-stat$	[−1.81]	[−1.88]	[−1.91]	[0.12]	[−0.41]
5	IR	−0.39	−0.39	−0.40	0.03	−0.09
6	$\hat{\alpha}$	−1.14	0.24	−1.16	0.31	0.15
6	$t-stat$	[−1.36]	[0.35]	[−1.54]	[0.49]	[0.22]
6	IR	−0.30	0.07	−0.34	0.11	0.05
7	$\hat{\alpha}$	−0.30	−0.27	−0.81	0.87	0.14
7	$t-stat$	[−0.43]	[−0.33]	[−1.04]	[1.29]	[0.20]
7	IR	−0.09	−0.07	−0.22	0.27	0.04
8	$\hat{\alpha}$	−0.19	−0.47	0.00	0.74	0.61
8	$t-stat$	[−0.23]	[−0.64]	[0.00]	[1.05]	[0.77]
8	IR	−0.05	−0.13	0.00	0.22	0.15
9	$\hat{\alpha}$	0.34	−0.20	0.19	0.56	1.08
9	$t-stat$	[0.41]	[−0.26]	[0.23]	[0.71]	[1.49]
9	IR	0.09	−0.06	0.05	0.15	0.29
10	$\hat{\alpha}$	1.06	1.77**	2.83***	2.25**	3.38***
10	$t-stat$	[1.12]	[2.04]	[3.07]	[2.53]	[4.33]
10	IR	0.23	0.41	0.61	0.51	0.86

Table 1.11: Double Sort: $\hat{\theta}$ and IR

This table documents the out-of-sample before-fees performance of the double sort by the FSD test statistic $\hat{\theta}$ and the historical in-sample information ratio (the ratio between $\hat{\alpha}$ and idiosyncratic volatility). Active funds in the cross section are first sorted into 10 deciles by the FSD test statistic $\hat{\theta}$ constructed from the 24 monthly observations prior to portfolio formation. The funds in each decile are then further sorted into 5 quintiles by the information ratio of the realized Carhart four-factor alpha during the proceeding 24 months. The portfolios of funds are rebalanced every 3 months. Out-of-sample alphas of the portfolios along with the t-statistics and information ratios are reported. Out-of-sample alphas with statistical significance are marked with "*". The sample period is from January 1991 to December 2015.

Decile(1st)	Quintile(2nd)	1	2	3	4	5
1	$\hat{\alpha}$	−0.20	−0.40	0.08	0.19	−1.06
1	$t-stat$	[−0.22]	[−0.45]	[0.10]	[0.26]	[−1.49]
1	IR	−0.05	−0.10	0.02	0.05	−0.32
2	$\hat{\alpha}$	−1.51*	−0.36	−0.27	−0.02	0.41
2	$t-stat$	[−1.76]	[−0.50]	[−0.45]	[−0.02]	[0.48]
2	IR	−0.37	−0.10	−0.09	0.00	0.10
3	$\hat{\alpha}$	0.23	−0.54	−0.18	−0.14	−0.82
3	$t-stat$	[0.27]	[−0.82]	[−0.29]	[−0.22]	[−1.14]
3	IR	0.06	−0.17	−0.06	−0.05	−0.25
4	$\hat{\alpha}$	−1.70**	−0.22	0.03	−0.33	0.43
4	$t-stat$	[−2.00]	[−0.33]	[0.04]	[−0.51]	[0.61]
4	IR	−0.44	−0.07	0.01	−0.10	0.12
5	$\hat{\alpha}$	−1.37*	−0.38	−0.73	−1.11*	−0.57
5	$t-stat$	[−1.72]	[−0.51]	[−1.09]	[−1.82]	[−0.81]
5	IR	−0.35	−0.11	−0.23	−0.36	−0.17
6	$\hat{\alpha}$	−0.56	0.41	0.19	−0.93	−0.64
6	$t-stat$	[−0.64]	[0.63]	[0.31]	[−1.38]	[−0.91]
6	IR	−0.14	0.14	0.06	−0.28	−0.20
7	$\hat{\alpha}$	0.02	−0.63	−0.14	1.21*	−0.87
7	$t-stat$	[0.02]	[−0.97]	[−0.24]	[1.89]	[−1.10]
7	IR	0.00	−0.20	−0.05	0.39	−0.21
8	$\hat{\alpha}$	0.74	−0.07	0.06	0.37	−0.43
8	$t-stat$	[0.86]	[−0.10]	[0.08]	[0.53]	[−0.51]
8	IR	0.18	−0.02	0.02	0.11	−0.10
9	$\hat{\alpha}$	1.16	−0.11	−0.26	1.41**	0.01
9	$t-stat$	[1.33]	[−0.13]	[−0.36]	[1.99]	[0.01]
9	IR	0.28	−0.03	−0.08	0.40	0.00
10	$\hat{\alpha}$	2.16**	1.52*	3.03***	2.46***	1.94**
10	$t-stat$	[2.51]	[1.65]	[3.51]	[2.90]	[2.25]
10	IR	0.52	0.33	0.70	0.57	0.44

Table 1.12: Double Sort: $\hat{\theta}$ and Average Ranking

This table documents the out-of-sample before-fees performance of the double sort by the FSD test statistic $\hat{\theta}$ and the historical average ranking of the fund among replica funds $(Pct\,(r_{i,t}, \langle\hat{r}_{i,t}\rangle))$. Active funds in the cross section are first sorted into 10 deciles by the FSD test statistic $\hat{\theta}$ constructed from the 24 monthly observations prior to portfolio formation. The funds in each decile are then further sorted into 5 quintiles by the average of the rankings $(Pct\,(r_{i,t}, \langle\hat{r}_{i,t}\rangle))$ during the proceeding 24 months. The portfolios of funds are rebalanced every 3 months. Out-of-sample alphas of the portfolios along with the t-statistics and information ratios are reported. Out-of-sample alphas with statistical significance are marked with "*". The sample period is from January 1991 to December 2015.

Table 1.9 shows that out-of-sample alphas are only significant for funds in the top $\hat{\theta}$ decile. The result is consistent with the argument that funds failing the FSD test are unskilled and are unable to generate either positive or negative out-of-sample alphas. Table 1.10 and 1.11 show that using t-statistic or information ratio as the second variable in the double sort produce similar patterns as using in-sample alpha. Table 1.12 shows that using average historical fund rankings as a second variable does not generate additional spread in funds out-of-sample alphas. The finding suggests the sort based on in-sample alpha is more efficient at identifying skilled managers once problems such as heteroscedasticity and benchmark mis-specification are alleviated by the FSD test.

1.C FSD Persistence

If the FSD test is able to identify managers with persistent skills, then there should also be persistence in managers' FSD test statistics ($\hat{\theta}$). The analysis so far is based on monthly observations and uses 24 months of performances to predict the next 3 months. The overlap in the formation periods makes the construction of a transition matrix difficult. In order to have non-overlapping formation periods with sufficient observations, I turn to daily fund returns to construct the FSD test statistic $\hat{\theta}$. The high frequency daily data enables me to construct non-overlapping quarterly FSD test statistics, although the sample period is shorter due to data availability.

Table 1.13 replicates the double sort with the FSD test statistic $\hat{\theta}$ as the first variable and the in-sample $\hat{\alpha}$ as the second variable. $\hat{\alpha}$ is still estimated using the proceeding 24 monthly observations as before, but the FSD test statistic $\hat{\theta}$ is constructed from daily observations in the proceeding 3 month only. The table shows that the non-overlapping quarterly FSD test statistic is still effective in identifying outperforming managers, although the magnitudes of the out-of-sample alphas are smaller in this subsample.

Decile(1st)	Quintile(2nd)	1	2	3	4	5
1	$\hat{\alpha}$	−0.22	0.53	0.63	0.45	0.54
1	$t-stat$	[−0.19]	[0.49]	[0.69]	[0.43]	[0.50]
1	IR	−0.05	0.12	0.17	0.12	0.12
2	$\hat{\alpha}$	−0.15	0.53	0.03	0.34	0.83
2	$t-stat$	[−0.15]	[0.65]	[0.05]	[0.42]	[0.89]
2	IR	−0.04	0.17	0.01	0.11	0.22
3	$\hat{\alpha}$	−0.93	0.33	0.34	0.53	−0.20
3	$t-stat$	[−0.93]	[0.43]	[0.46]	[0.56]	[−0.21]
3	IR	−0.25	0.11	0.12	0.14	−0.05
4	$\hat{\alpha}$	−1.40	0.37	0.07	0.99	−0.05
4	$t-stat$	[−1.43]	[0.46]	[0.08]	[1.38]	[−0.05]
4	IR	−0.37	0.11	0.02	0.36	−0.01
5	$\hat{\alpha}$	−0.98	0.22	1.37**	0.24	1.04
5	$t-stat$	[−1.02]	[0.29]	[2.20]	[0.28]	[0.99]
5	IR	−0.27	0.08	0.52	0.07	0.25
6	$\hat{\alpha}$	0.25	−0.31	−0.19	0.18	0.68
6	$t-stat$	[0.27]	[−0.40]	[−0.28]	[0.21]	[0.67]
6	IR	0.06	−0.10	−0.07	0.05	0.16
7	$\hat{\alpha}$	0.53	−0.37	−0.33	0.33	0.06
7	$t-stat$	[0.55]	[−0.43]	[−0.51]	[0.38]	[0.05]
7	IR	0.14	−0.11	−0.12	0.10	0.01
8	$\hat{\alpha}$	1.04	0.14	−0.84	0.58	−0.72
8	$t-stat$	[1.19]	[0.20]	[−1.17]	[0.77]	[−0.60]
8	IR	0.29	0.05	−0.30	0.18	−0.14
9	$\hat{\alpha}$	2.02*	1.03	1.39*	1.65*	1.19
9	$t-stat$	[1.76]	[1.33]	[1.75]	[1.94]	[1.06]
9	IR	0.41	0.32	0.42	0.43	0.26
10	$\hat{\alpha}$	1.49	1.29	1.29*	1.42*	2.65**
10	$t-stat$	[1.58]	[1.53]	[1.65]	[1.68]	[2.39]
10	IR	0.37	0.36	0.38	0.39	0.51

Table 1.13: Double Sort: Daily $\hat{\theta}$ and $\hat{\alpha}$

This table documents the out-of-sample before-fees performance of the double sort by the FSD test statistic $\hat{\theta}$ and the historical in-sample $\hat{\alpha}$. Active funds in the cross section are first sorted into 10 deciles by the FSD test statistic $\hat{\theta}$ constructed from the daily fund returns during the 3 months prior to portfolio formation. The funds in each decile are then further sorted into 5 quintiles by the realized Carhart four-factor alpha during the proceeding 24 months. The portfolios of funds are rebalanced every 3 months. Out-of-sample alphas of the portfolios along with the t-statistics and information ratios are reported. Out-of-sample alphas with statistical significance are marked with "*". The sample period is from January 1999 to December 2015.

Table 1.15 presents the transition matrix of the quarterly non-overlapping FSD test statistics $\hat{\theta}$. Active funds in the cross section are sorted into 10 deciles by the FSD test statistic $\hat{\theta}$ constructed from the daily returns in each quarter. The table reports the estimated probabilities (in %) to transition from one decile (row decile) to another decile (column decile) during the next quarter.[15] Suppose there is no persistence in the FSD test statistics, then each cell in the table should be close to 10%. The estimation reveals persistence in both the low FSD decile (cell 1-1) and high FSD decile (cell 10-10). Interestingly, funds in the low FSD decile are also more likely to enter the high FSD decile in the subsequent quarter, i.e. cell 1-10 is high. The pattern might explain the double sort's failure to identify persistent underperforming funds in Table 1.13. I leave further investigations on this issue to future research.

[15]The sum of each row is slightly lower than 100% because some of the funds don't have enough daily observations in the subsequent quarter to construct the FSD test statistic.

Decile (pre)	(post) in %	1	2	3	4	5	6	7	8	9	10
1		13.12	11.25	10.32	8.84	7.79	7.83	8.03	8.31	10.10	10.88
2		12.18	11.29	10.69	9.23	8.57	9.10	7.88	9.09	9.69	9.12
3		10.87	10.39	11.44	9.77	9.16	9.36	8.93	8.77	9.30	8.53
4		9.06	10.10	10.63	10.88	10.67	9.98	9.40	8.86	8.70	8.46
5		8.53	9.07	9.91	10.60	10.93	10.55	10.63	9.28	8.97	8.55
6		8.93	8.67	9.81	9.72	11.55	10.53	10.52	10.13	9.17	8.22
7		7.36	8.99	9.09	10.59	10.57	10.76	10.63	10.49	9.36	8.81
8		8.73	8.65	9.35	9.76	10.36	9.79	11.04	10.16	9.77	10.04
9		8.46	9.49	8.71	8.02	8.90	8.96	10.74	11.35	11.08	11.05
10		9.10	9.01	8.79	8.49	8.88	8.99	10.34	10.10	11.15	12.26

Table 1.15: FSD Transition Matrix

The table documents the probability transition matrix of the fund deciles constructed by the FSD test statistic $\hat{\theta}$. By the end of each quarter, active funds in the cross section are sorted into 10 deciles by the FSD test statistic $\hat{\theta}$ constructed from the daily fund returns during the 3 months prior to portfolio formation. The estimated probabilities (in %) to transition from one decile (row decile) to another decile (column decile) during the next quarter are reported. The sample period is from January 1999 to December 2015.

Chapter 2

Mutual Fund Managers from the Buffett School

2.1 Introduction

Whether active professional money managers can outperform the stock market is a topic that has held an enduring fascination for researchers in academia and industry alike. Early documentations in the finance literature such as Hendricks, Patel, and Zeckhauser (1993), Brown and Goetzmann (1995) and Wermers (1997) suggested that the recent performance of a mutual fund manager can be used to predict his future performance, also known as the "Hot Hand" effect. However, Jensen (1969) failed to find the aforementioned "Hot Hand" effect, and Carhart (1997) suggested that the "Hot Hand" effect can be, to a large extent, explained by the stock-level momentum effect as documented in Jegadeesh and Titman (1993). The current consensus shared by most researchers in academia is that, on average, active mutual fund managers cannot outperform the aggregate market before fees after adjusted for popular risk factors, and they under-perform the market after fees (see Fama and French (2010)), echoing Malkiel and Fama (1970) that a liquid capital market should be largely efficient.

Of course, people's fascination about the money management industry does not stop at the "average manager". Even if on average managers cannot outperform the market, can some of the managers possess the skills to beat the market in a consistent fashion, and more importantly, is there a way to identify such managers ex ante? Kosowski et al. (2006), Fama and French (2010) and Linnainmaa (2013) studied the ex post realized fund performances and argued that some of the managers enjoyed superior performances that were unlikely to be explained by pure statistical coincidences. On the other hand, researchers have also made progress on the agenda of identifying skilled managers ex ante. Cohen, Coval, and Pástor (2005) showed that managers who hold portfolios that resemble other "good" managers also tend to be "good" managers; Kacperczyk, Sialm, and Zheng (2005) pointed out that managers with concentrated industry holdings tend to outperform the diversifiers; Kacperczyk, Sialm, and Zheng (2008) argued that managers with more unobserved actions seem to possess skills; and Cremers and Petajisto (2009) suggested that managers with more active shares tend to outperform their benchmarks. This paper aims to further enrich this line of literature.

In this paper, I identify two trading styles that can be used to predict mutual fund managers' future performances - Herding and Trading Intensity. The Herding measure is defined as a manager's tendency to contemporaneously trade in the same directions as the majority of other active mutual fund managers. My measure draws several key distinctions from the herding measure constructed in Jiang and Verardo (2013). Their measure captures the average sensitivity of portfolio weight changes to lagged changes in institutional ownership. My measure differs from theirs in that I intend to capture the contemporaneous herding behavior without a lag and my measure is explicitly about the "extensive margin" (many people trade in the same direction) whereas the institutional ownership variable employed in Jiang and Verardo (2013) is ambiguous in that respect (the increase in the institutional ownership can be caused by either the scenario that many funds are buying into the same stock, or the scenario that one fund is buying a lot into that stock). The Trading Intensity measure is defined as a manager's tendency to adjust portfolio weights. The Trading Intensity measure is a close cousin to the conventional turnover measure. The difference between Trading Intensity and turnover is that Trading Intensity is based on portfolio weights alone, whereas turnover is constructed from the dollar value of the securities being bought and sold. Therefore, the Trading Intensity measure captures a manager's tendency to change portfolio compositions and is less sensitive to fund flows than the turnover measure.

I document that, during 1995-2015, Herding can predict future fund performances with significant magnitudes both statistically and economically. Specifically, I show that the most anti-herding funds can outperform the benchmarks, whereas the funds with the strongest herding tendencies only earn average returns (before fees). The magnitude of outperformance can be further amplified when the funds are categorized by Trading Intensity in a second-stage sort. The group of funds with the least Trading Intensity within the group of the most anti-herding funds can earn an average Carhart alpha of 3.06% before fees and 1.94% after fees. Compared with an average fund, these funds tend to be older, larger and manage fewer stocks within their portfolios.

So who are those funds? On their websites, they all claim to be long-term, fundamental, value investors. In order words, these are the funds who embrace the investment philosophy exemplified by Warren Buffett. I verify their claims by showing that these managers do have longer portfolio ages compared with the average in the population and they load positively on the QMJ factor constructed in Asness, Frazzini, and Pedersen (2014) (also see Frazzini, Kabiller, and Pedersen (2013)). Warren Buffett, as an individual investor, has been regarded as a guru on long-term, fundamental investment, and has enjoyed widespread worship and extensive media coverage in the financial world. However, rigorous empirical research on Buffett's phenomenal success has been limited to case studies, as informally picking out funds who claim to share his investment approach ex post suffers from selection biases.[1] This paper formally establishes a quantitative system to ex ante identify Buffett-like managers, with Herding and Trading Intensity serving as a screening device. And I provide statistical support that the investors from the "Graham-and-Doddsville"(Buffett (1984)) do seem to possess the skills to outperform the market as advertised.

Another surprising finding from the data is that not only the group of Buffett-like managers can outperform the market by a large margin, but they also charge lower fees compared with an average manager in the industry. Specifically, they only charge 112 bps per year, leaving investors an after-fees alpha as high as 1.94%. This is at odds with the equilibrium described in Berk and Green (2004). In the Berk and Green world, mutual fund managers hold perfect bargaining power against their investors so that they can charge fees as high as their before-fees alpha, leaving investors only breaking even with the market after fees. Moreover, I show that these managers' long-term investment strategy can be easily replicated. The mechanical strategy of investing in their lagged portfolio compositions when they become available

[1] See Frazzini, Kabiller, and Pedersen (2013), Chirkova (2012), Martin and Puthenpurackal (2008), Statman and Scheid (2002), etc.

can almost perfectly recover the before-fees alpha earned by these managers. So potentially, if the managers were to post high fees, then their investors could just construct the trades on their own and enjoy all the benefits rather than investing with these managers. In other words, the very nature of the long-term investment philosophy, though profitable, also makes these managers vulnerable to free-riders.

The rest of the paper is arranged as the following. I describe the data in Section 2. I present the definition of the Herding and Trading Intensity measures in Section 3 and Section 4, respectively. Section 5 documents the identities of the selected mutual funds and relates their characteristics to the Herding and Trading Intensity measure. Section 6 performs robustness checks on the influence of internet stocks. Section 7 documents the low fees and the replicability of the managers' investment strategies. Section 8 concludes.

2.2 Data

I obtain monthly after-fees fund returns along with other fund characteristics such as fund size, age, name, expense ratio, etc. from CRSP Survivor-Bias-Free US Mutual Fund Database. I compute the before-fees returns by adding back the expense ratio to the after-fees fund returns. I obtain fund holdings from Thomson Reuters Mutual Fund Holdings (s12), formerly known as the CDA/Spectrum Mutual Fund Holdings Database. Both databases are standard in this line of research. Their popularity arose largely due to their efforts to eliminate survivorship bias by making an attempt to include all the funds that have ever existed in the US market. In fact, Linnainmaa (2013) raised the concern of a potential reverse survivorship bias by using these databases as funds hit by a series of unlucky negative shocks tend to exit the market, leaving behind trajectories of poor performances without the chances to "clear their names". Therefore, my finding of superior performances is unlikely to be caused potential survivorship bias. I follow the standard approach to link these two databases with the MFLINKS database constructed by Prof. Russ Wermers, and I obtain stock prices and returns from the CRSP Monthly Stock File.

I limit my focus on domestic, diversified, actively managed, US equity funds. I employ the investment objectives code (crsp_obj_cd) that has been recently introduced by CRSP as my screening variable to identify such funds.[2] Doshi, Elkamhi, and Simutin (2015) shows that the funds identified with the crsp_obj_cd are almost identical to the funds identified with the investment objectives codes from other data vendors that have been used in earlier literature.[3] To reduce the impact from very small funds, I require the funds in my sample to have at least $5 million under management and hold at least 20 stocks in their portfolios. I aggregate funds with multiple share classes into a single class as these different share classes share the same portfolio composition. Due to the limitation that my Herding measure requires a certain number of funds trading in the market, I pick my sample period from January 1995 to December 2015. I have 2693 distinctive funds in my sample and 338,180 fund-month observations. Table 2.1 documents the summary statistics of the funds that are included in my sample.

Table 2.1: Funds Summary Statistics

This table documents the summary statistics of the funds that are included in my sample. The sample period is 1995-2015. I identify domestic, diversified, actively managed, US equity funds by including funds with crsp_obj_cd beginning with "EDC" or "EDY"; excluding funds with crsp_obj_cd being "EDYH" or "EDYS"; and excluding option income funds with Strategic Insight Objectives code being "OPI". I further eliminate index funds by screening fund names. I require funds to have at least $5 million under management and hold at least 20 stocks in their portfolios.

Number of Funds is the number of identified actively managed funds in the cross-section of each month; TNA is the total net asset under management; Fund Age is computed as the time difference between the current month and the month of fund initiation; Expense Ratio (annualized) and Turnover Ratio are both directly from CRSP.

	Mean	Max	Min	p25	p75
Number of Funds	1342	1718	681	1186	1581
TNA (in million $)	1094	202306	5	58	728
Fund Age (in years)	12	85	0.08	5	15
Number of Holding Stocks	113	2336	20	50	117
Expense Ratio (%)	1.24	13.5	0	0.93	1.44
Turnover Ratio	0.89	42.63	0	0.36	1.12

[2] I include funds with crsp_obj_cd that begins with "EDC" or "EDY"; exclude funds with crsp_obj_cd being "EDYH" or "EDYS"; and exclude option income funds with Strategic Insight Objectives code being "OPI". I then eliminate index funds by screening fund names.

[3] I thank the authors for sharing their SAS code online.

2.3 Herding as a Predictor of Fund Performances

2.3.1 Stock-level Herding Measure

The herding measure for stock i during quarter t is defined as:

$$h_{i,t} = \frac{b_{i,t}}{b_{i,t} + s_{i,t}} - \frac{\sum_i b_{i,t}}{\sum_i b_{i,t} + \sum_i s_{i,t}}$$

where $b_{i,t}(s_{i,t})$ is the number of funds that have increased(decreased) position on stock i during quarter t; $\frac{\sum_i b_{i,t}}{\sum_i b_{i,t}+\sum_i s_{i,t}}$, i.e. the fraction of buying orders out of total transactions during quarter t, is a normalizing factor close to 0.5.

My stock-level herding measure is a close variant to the popular measure employed in Lakonishok, Shleifer, and Vishny (1992), Wermers (1999), Grinblatt, Titman, and Wermers (1995), etc. The difference between my measure and theirs is that my measure can take both positive and negative signs. A positive herding measure indicates that the majority of investors are buying into stock i during quarter t; and a negative value suggests that the majority of investors are selling out of the stock during quarter t. The herding measure is defined based on the extensive margin (number of funds), so that it is not influenced by the size of the funds that are participating in the trading of the stock.

The measure of herding is only meaningful when there are enough trades. In my empirical implementation, I require the herding measure to be only defined when the stock is traded by at least 100 funds for a given quarter, i.e. $b_{i,t} + s_{i,t} \geq 100$. Alas, such a requirement imposes a non-trivial constraint on my sample that I would need sufficient number of funds to be trading in the market. As a result, I choose my sample period to be from January 1995 to December 2015 as there are too few funds during the early episodes of the data.

2.3.2 Fund-level Herding Measure

To capture of the tendency of a fund to contemporaneously herd with the majority of other investors, I assign a score to each fund j for a given quarter t:

$$s_t^j = \sum_{i \in \mathbb{N}_t^j} \left(w_{i,t}^j - w_{i,t-1}^j\right) \cdot h_{i,t}$$

where $\mathbb{N}_t^j$ is the set of stocks within fund j's portfolio by the end of quarter t; $w_{i,t}^j(w_{i,t-1}^j)$ is the portfolio weight of stock i of fund j by the end of quarter $t(t-1)$[4].

The score s_t^j correlates the change of portfolio weights with the stock-level herding measure so that a fund would be assigned a high score if it tends to increase the weight of the stocks that the majority of other funds are buying into or decrease the weight of the stocks that the majority of other funds are selling out of.

The fund-level herding measure is then defined as the time-average of the scores:

$$H_t^j = \frac{1}{T} \sum_{u=t-T+1}^{t} s_u^j$$

In my empirical implementation, I take T to be 12 quarters.

2.3.3 Persistence of the Fund-level Herding Measure

Whether the fund-level herding measure is a relevant construction depends on whether it is able to capture a persistent fund trading style. In other words, the degree of herding is only a meaningful dimension of fund characteristics if the funds who herded in the past also tend to herd in the future.

To verify the persistence of herding as a style, I use the fund-level herding measure (H_t^j) to predict the one-period ahead herding score (s_{t+1}^j). [5] Table 2.2 shows that herding tendency is a persistent fund characteristic. Panel A of Table 2.2 displays a transition matrix. Each row of the table corresponds to a decile sorted by the fund-level herding measure at time t (H_t^j). Each column corresponds to a decile sorted by the one-period ahead herding score (s_{t+1}^j). The numbers in the table are the probabilities for a fund in a given H_t^j decile to land on a given s_{t+1}^j decile. The table shows that funds with low fund-level herding measure this quarter (H_t^j) also tend to acquire low one-period herding score the next quarter (s_{t+1}^j), and vice versa. Panel B forms 10 portfolios based on the fund-level herding measure each quarter (H_t^j), and computes the average next-quarter herding score (s_{t+1}^j) for each portfolio. Panel B

[4]If portfolio weights of the last quarter are not available, I take the portfolio weights as of two quarters ago. I ignore the observation if both the portfolio weights of last quarter and two quarters ago are not available.

[5]The persistence of the fund-level herding measure (fh_t^j) itself is not meaningful, since it is persistent by construction as a moving time average.

shows that, indeed, portfolios with low fund-level herding measure this quarter (H_t^j) tend to acquire low herding score next quarter (s_{t+1}^j) on average, and vice versa.

Table 2.2: Persistence of Fund-level Herding

This table documents the persistence of mutual funds' herding tendencies. Panel A displays a transition matrix. Each row of the table corresponds to a decile sorted by the fund-level herding measure at time t (H_t^j). Each column corresponds to a decile sorted by the one-period ahead herding score (s_{t+1}^j). The numbers in the table are the probabilities for a fund in a given H_t^j decile to land on a given s_{t+1}^j decile. The table shows that funds with low fund-level herding measure this quarter (H_t^j) also tend to acquire low one-period herding score the next quarter (s_{t+1}^j), and vice versa. Panel B forms 10 portfolios based on the fund-level herding measure each quarter (H_t^j), and computes the average next-quarter herding score (s_{t+1}^j) for each portfolio. Panel B shows that, indeed, portfolios with low fund-level herding measure this quarter (H_t^j) tend to acquire low herding score next quarter (s_{t+1}^j) on average, and vice versa.

Panel A: Transition Matrix

rank(H_t^j)\rank(s_{t+1}^j)	1	2	3	4	5	6	7	8	9	10
1	0.17	0.22	0.19	0.11	0.08	0.06	0.04	0.02	0.02	0.03
2	0.11	0.21	0.24	0.16	0.09	0.05	0.04	0.03	0.01	0.01
3	0.12	0.15	0.17	0.16	0.12	0.09	0.06	0.03	0.03	0.02
4	0.11	0.11	0.12	0.14	0.13	0.10	0.09	0.07	0.04	0.04
5	0.10	0.09	0.08	0.11	0.13	0.13	0.10	0.09	0.07	0.05
6	0.09	0.07	0.07	0.09	0.11	0.13	0.13	0.11	0.09	0.07
7	0.08	0.04	0.05	0.08	0.10	0.13	0.13	0.14	0.12	0.09
8	0.07	0.04	0.03	0.06	0.08	0.11	0.13	0.15	0.16	0.13
9	0.05	0.03	0.03	0.04	0.06	0.09	0.13	0.15	0.19	0.18
10	0.05	0.02	0.02	0.02	0.04	0.06	0.09	0.14	0.20	0.31

Panel B: Average Score

rank(H_t^j)	Avg(s_{t+1}^j)
1	0.0016
2	0.0017
3	0.0024
4	0.0036
5	0.0044
6	0.0055
7	0.0067
8	0.0081
9	0.0099
10	0.0133

2.3.4 Herding Tendency as a Predictor of Future Fund Performances

Now that I've demonstrated that fund-level herding as a persistent fund trading style, I turn to show that a fund's tendency to herd can be used to predict its future performance.

To investigate the predictive power of the fund-level herding tendency, I form 10 portfolios of active mutual funds at each quarter t based on the fund-level herding measure constructed with data until quarter $t-1$ (H^j_{t-1}).[6] The design of the portfolio rebalancing strategy ensures its implementability as the data employed at the time of portfolio construction is at least 3 months old. I then hold the portfolios for three months, record their performances, and rebalance again when the next quarter arrives.

Table 2.3 and Table 2.4 document the performances of the portfolios. The sample period is from January 1995 to December 2015. The tables show the annualized Carhart 4-factor alphas, regression R-squares, information ratios as well as the factor loadings of the portfolios. Table 2.3 takes the before-fees fund returns as the left-hand-side variable in the regression, whereas Table 2.4 documents the after-fees performances. Both tables show that the 4-factor alpha decreases almost monotonically with the fund-level herding measure. The portfolio with the lowest herding tendency outperforms the portfolio with the highest herding tendency by 209 bps (197 bps) before fees (after fees), adjusted for the Carhart factors. The magnitude of the outperformance is both economically and statistically significant. As for the before-fees performances, the decile with the lowest herding tendency is able to outperform the Carhart benchmark by 189 bps per year, whereas the decile with the highest herding tendency does not significantly outperform or under-perform the Carhart benchmark. After fees, the 4-factor alpha of the most anti-herding portfolio drops to 52 bps per year and is no longer statistically significant, whereas the portfolio with the highest herding tendency significantly under-performs the Carhart benchmark. In terms of the factor loadings, the anti-herding funds tend to hold small, high book-to-market stocks with low market betas, whereas the herding funds tend to hold large, low book-to-market stocks with high market betas. The pattern is consistent with the conjecture that the anti-herding funds are able to outperform the

[6] For funds whose herding measure are missing for a given quarter, their portfolio ranks are inherited from the last available values.

market thanks to their stock-picking skills. And they have more advantage at picking small, value firms as there might be more information asymmetry among such firms.

2.4 Second-stage Sort with Trading Intensity

So why are the anti-herding funds able to outperform the market? Is it because they make distinctive trades compared with other investors, or is it because they don't trade much at all. I define the measure of Trading Intensity to differentiate between these two hypotheses.

2.4.1 The Trading Intensity Measure

The Trading Intensity measure is defined as the following:

$$TI_t^j = \frac{1}{T} \sum_{u=t-T+1}^{t} \left(\sum_i \left| w_{i,u}^j - w_{i,u-1}^j \right| \right)$$

where $w_{i,u}^j(w_{i,u-1}^j)$ is stock i's weight in fund j's portfolio by the end of quarter $u(u-1)$.

The TI measure is an intuitive construction trying to capture of the intensity of a fund to rebalance its portfolio. It is closely related to the conventional turnover measure, but with some subtle yet important differences. The definition of turnover given by CRSP is $\frac{\min(buy_t, sell_t)}{avg(TNA_t)}$, where $buy_t(sell_t)$ is the dollar value of the securities bought(sold) by a fund during month t. My Trading Intensity measure differs from the turnover measure in that the Trading Intensity measure is based on portfolio weights alone, whereas the turnover measure is based on the dollar value of the securities being transacted. Therefore, the Trading Intensity measure is less sensitive to the influence of fund flows compared with the turnover measure. Consider, for example, an index fund passively tracking a fixed target portfolio. The Trading Intensity of this fund would always be zero, whereas the turnover measure would be non-zero as capital flows in and out of the fund. The Trading Intensity measure is thus more relevant to my purpose, which is to determine the source of profitability of the anti-herding funds.

Table 2.3: Herding as a Predictor of Future Fund Performance (Before-fees)

This table documents the ability of the fund-level herding measure to predict future fund performance. The sample period is from January 1995 to December 2015. The panels in the table show the annualized (before-fees) Carhart 4-factor alphas, regression R-squares, information ratios as well as the factor loadings of the portfolios. The series of the Carhart factors are from Prof. Ken French's website.

			Before-fees Performances				
Decile	α (%)	Market	Value	Size	Momentum	R-square (%)	IR
1	1.89***	0.96***	0.21***	0.41***	−0.02*	97.1	0.65
	[2.69]	[69.4]	[10.8]	[22.8]	[−1.92]		
2	1.39*	1.00***	0.18***	0.40***	0.00	97.1	0.47
	[1.94]	[70.4]	[9.19]	[21.9]	[−0.21]		
3	0.58	1.00***	0.11***	0.32***	0.01	97.3	0.21
	[0.85]	[74.9]	[6.13]	[18.3]	[1.33]		
4	0.38	1.00***	0.05***	0.23***	0.00	97.6	0.15
	[0.60]	[81.5]	[2.92]	[14.5]	[−0.05]		
5	−0.20	1.01***	0.03**	0.17***	0.00	98.3	−0.10
	[−0.40]	[99.4]	[2.20]	[12.7]	[0.21]		
6	0.00	1.01***	0.00	0.07***	0.00	98.7	0.00
	[0.00]	[116.2]	[−0.05]	[6.42]	[−0.38]		
7	−0.08	1.01***	−0.01	0.03***	0.02**	98.7	−0.04
	[−0.18]	[116.3]	[−1.23]	[2.93]	[2.51]		
8	−0.03	1.02***	−0.04***	0.01	0.03***	98.4	−0.02
	[−0.07]	[106.9]	[−3.11]	[1.17]	[4.18]		
9	−0.09	1.04***	−0.09***	0.00	0.04***	97.7	−0.04
	[−0.16]	[87.3]	[−5.45]	[−0.02]	[4.47]		
10	−0.20	1.06***	−0.18***	0.03	0.08***	96.5	−0.06
	[−0.26]	[69.3]	[−8.66]	[1.31]	[6.10]		
10-1	−2.09**	0.10***	−0.39***	−0.38***	0.10***	70.4	−0.58
	[−2.38]	[5.72]	[−16.3]	[−17.1]	[6.93]		

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

Table 2.4: Herding as a Predictor of Future Fund Performance (After-fees)

This table documents the ability of the fund-level herding measure to predict future fund performance. The sample period is from January 1995 to December 2015. The panels in the table show the annualized (after-fees) Carhart 4-factor alphas, regression R-squares, information ratios as well as the factor loadings of the portfolios. The series of the Carhart factors are from Prof. Ken French's website.

After-fees Performances							
Decile	α (%)	Market	Value	Size	Momentum	R-square (%)	IR
1	0.52	0.97***	0.21***	0.41***	−0.02*	97.1	0.18
	[0.73]	[69.8]	[10.8]	[22.6]	[−1.79]		
2	0.07	1.01***	0.18***	0.41***	0.00	97.1	0.02
	[0.09]	[70.6]	[9.12]	[22.0]	[−0.10]		
3	−0.70	1.01***	0.12***	0.32***	0.01	97.4	−0.25
	[−1.04]	[75.8]	[6.31]	[18.4]	[1.28]		
4	−0.82	1.01***	0.05***	0.23***	0.00	97.6	−0.31
	[−1.30]	[81.8]	[3.10]	[14.4]	[−0.04]		
5	−1.34***	1.01***	0.03**	0.16***	0.00	98.4	−0.63
	[−2.63]	[100.6]	[2.04]	[12.6]	[0.11]		
6	−1.16***	1.01***	0.00	0.08***	0.00	98.7	−0.63
	[−2.61]	[116.0]	[−0.21]	[6.75]	[−0.25]		
7	−1.31***	1.02***	−0.01	0.03***	0.02**	98.7	−0.70
	[−2.91]	[115.0]	[−1.08]	[2.98]	[2.46]		
8	−1.17**	1.03***	−0.04***	0.02	0.03***	98.5	−0.58
	[−2.42]	[107.7]	[−2.99]	[1.26]	[4.24]		
9	−1.32**	1.04***	−0.09***	−0.01	0.04***	97.7	−0.53
	[−2.19]	[87.9]	[−5.42]	[−0.37]	[4.51]		
10	−1.45*	1.07***	−0.18***	0.03	0.08***	96.5	−0.45
	[−1.87]	[69.5]	[−8.66]	[1.40]	[6.12]		
10-1	−1.97**	0.10***	−0.39***	−0.38***	0.10***	69.8	−0.54
	[−2.23]	[5.56]	[−16.2]	[−16.8]	[6.82]		

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

2.4.2 Second-stage Sort with Trading Intensity

In order to determine the source of profitability of the anti-herding funds, I first form 8 portfolios of active mutual funds sorted by their fund-level Herding measure. Then within each group of the mutual funds, I further form 8 portfolios sorted by the Trading Intensity of the funds. I end up with $8 \times 8 = 64$ portfolios of mutual funds as a result. Again, in order to ensure the implementability of the trading strategy, I make sure that the information used to construct the portfolios is at least 3 months old at the time of portfolio construction. I then hold the portfolios for 3 months, record their performances, and rebalance again by the end of the 3 months.

Table 2.5 and Table 2.4 document the annualized Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. Each column of the tables corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity.

From the tables, it is obvious that among the group of anti-herding funds, it is the group of funds with the least Trading Intensity (Cell 1-1) that achieves the highest performance. The null hypothesis that the group of funds within Cell 1-1 underperforms the average of all the remaining funds within the first column is rejected with a p-value being 0.03. Interestingly, the Trading Intensity measure alone cannot predict future fund performances, although it can be used to refine the first-stage sort by the fund-level herding measure.[7]

Table 2.7 documents the characteristics of the 8 Trading Intensity portfolios among the most Anti-herding funds (column 1 in Table 2.5). Age is the average age of the funds within the portfolio. Size is the aggregate TNA of the portfolio normalized by all the aggregate TNA of all the portfolios in the cross-section. If all the portfolios are of equal total TNA, the Size would be $1/64 = 0.01525$. NStocks is the average number of stocks held by the funds within the portfolio.

From the table, it is striking that the group of the funds with the best performances (anti-herding funds with the least Trading Intensity) tend to be older, larger, and hold fewer stocks within their portfolios, compared with an average fund in the cross section. The fact that these funds are larger is at odds with the assumption made in Berk and Green (2004) that mutual funds employ the technology that has a decreasing return to scale feature. On the other hand, the fact that these funds are larger but also

[7]Results are not tabulated.

Table 2.5: Double-sort Alphas (Before-fees)

This table documents the (before-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

Before-fees Alphas								
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	3.06***	1.22	0.55	0.67	0.16	−0.20	0.88	0.39
	[3.58]	[1.38]	[0.97]	[1.47]	[0.39]	[−0.40]	[1.27]	[0.49]
2	1.87**	1.24	1.22*	−0.11	−0.26	0.89	−0.38	0.42
	[2.36]	[1.52]	[1.74]	[−0.20]	[−0.49]	[1.55]	[−0.60]	[0.58]
3	2.28***	1.38	1.49**	0.27	−0.10	0.02	0.67	−0.17
	[2.81]	[1.55]	[2.53]	[0.48]	[−0.19]	[0.04]	[1.09]	[−0.23]
4	2.12***	0.13	0.19	0.36	−1.18**	0.30	0.54	0.07
	[2.64]	[0.14]	[0.24]	[0.60]	[−2.06]	[0.59]	[0.96]	[0.08]
5	1.51*	0.87	−0.01	0.46	−0.05	0.05	−0.33	−0.37
	[1.66]	[1.04]	[−0.01	[0.64]	[−0.09]	[0.09]	[−0.44]	[−0.43]
6	1.75**	0.57	0.26	0.15	−0.16	0.29	−0.23	−0.68
	[2.04]	[0.65]	[0.27]	[0.18]	[−0.26]	[0.42]	[−0.34]	[−0.71]
7	1.62*	−0.30	0.18	−0.90	0.22	−0.90	−0.76	−0.33
	[1.95]	[−0.35]	[0.21]	[−0.98]	[0.29]	[−1.26]	[−1.01]	[−0.34]
8	1.47	0.20	−0.06	−0.85	−0.27	−0.05	−0.22	−0.61
	[1.57]	[0.22]	[−0.05]	[−0.84]	[−0.28]	[−0.05]	[−0.24]	[−0.46]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

Table 2.6: Double-sort Alphas (After-fees)

This table documents the (after-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

After-fees Alphas								
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	1.94**	0.17	−0.31	−0.35	−0.92**	−1.19**	−0.38	−0.84
	[2.35]	[0.19]	[−0.55]	[−0.80]	[−2.35]	[−2.42]	[−0.54]	[−1.07]
2	0.62	0.01	0.10	−1.19**	−1.41***	−0.13	−1.47**	−0.82
	[0.79]	[0.01]	[0.15]	[−2.24]	[−2.64]	[−0.22]	[−2.39]	[−1.15]
3	0.90	0.06	0.28	−0.84	−1.18**	−1.20**	−0.49	−1.53**
	[1.12]	[0.07]	[0.48]	[−1.48]	[−2.26]	[−2.49]	[−0.80]	[−2.09]
4	0.70	−1.21	−0.95	−0.67	−2.41***	−0.82	−0.61	−1.15
	[0.87]	[−1.30]	[−1.17]	[−1.11]	[−4.19]	[−1.58]	[−1.10]	[−1.35]
5	−0.05	−0.32	−1.38	−0.60	−1.12**	−1.12**	−1.47**	−1.61*
	[−0.05]	[−0.38]	[−1.46]	[−0.83]	[−2.04]	[−2.12]	[−1.98]	[−1.91]
6	0.27	−0.75	−1.09	−0.98	−1.24*	−0.93	−1.40**	−2.06**
	[0.33]	[−0.83]	[−1.15]	[−1.18]	[−1.94]	[−1.38]	[−2.14]	[−2.14]
7	0.23	−1.58	−1.37	−2.28**	−1.02	−2.32***	−1.93***	−1.62
	[0.28]	[−1.85]	[−1.53]	[−2.49]	[−1.33]	[−3.17]	[−2.58]	[−1.64]
8	0.13	−1.16	−1.52	−2.30**	−1.66*	−1.47*	−1.71*	−1.90
	[0.14]	[−1.24]	[−1.26]	[−2.30]	[−1.65]	[−1.67]	[−1.88]	[−1.47]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

manage fewer stocks within their portfolios is similar to the finding in Kacperczyk, Sialm, and Zheng (2005) that concentrated funds tend to have better performances. If I further install a filter on the group of the such funds (Cell 1-1 in Table 2.5) requiring them to be at least 10 years old, have at least $200M under management, and hold no more than 200 stocks by the time of portfolio construction, the 4-factor alpha of the portfolio would further jump to as high as 3.67%(2.71%) per year, with a t-stat being 3.31(2.46), information ratio being 0.80(0.59), and Shape ratio being 0.66(0.59) before fees(after fees). In other words, the performance of the portfolio with the most anti-herding and anti-trading funds can be further significantly improved if young, small, and diversified funds are excluded.

Table 2.7: Characteristics of the Anti-herding Funds

This table documents the characteristics of the 8 Trading Intensity portfolios among the most Anti-herding funds (column 1 in Table 2.5). Age is the average age of the funds within the portfolio. Size is the aggregate TNA of the portfolio normalized by all the aggregate TNA of all the portfolios in the cross-section. If all the portfolios are of equal total TNA, the Size would be $1/64 = 0.01525$. NStocks is the average number of stocks held by the funds within the portfolio.

TI Rank	Age	Size	NStocks
1	14.12	0.0197	95.96
2	12.92	0.0126	133.18
3	12.19	0.0103	104.96
4	11.19	0.0091	125.20
5	10.59	0.0088	107.36
6	10.70	0.0073	110.08
7	9.29	0.0056	136.45
8	9.22	0.0033	119.33

2.5 Identities of the Ultra-performance Funds

So who are these funds(Cell 1-1 funds) really? Table 2.8 and Table 2.9 list the top 10 funds ranked by their number of times being picked up by the portfolio strategy, as well as the investment philosophies that they disclose on their websites. From the tables, it is obvious that the selected funds all share similar traits - they all claim to be long-term, fundamental, value investors who embrace the investment philosophy exemplified by Warren Buffett (See Buffett (1984)).

Table 2.8: Fund Identities

This table and the next table list the top 10 funds ranked by their number of times being picked up by the portfolio strategy. Occurrences is the total number of times that a fund being included in the filtered Cell 1-1 portfolio. Investment Philosophy is quoted from the website of the corresponding fund.

Rank	Fund Name	Occurrences	Investment Philosophy
1	Jensen Quality Growth Fund	37	The strength of our investment philosophy is based on an unwavering commitment to investing in quality businesses.
2	Franklin Managed Trust: Franklin Rising Dividends Fund	36	We employ a unique, disciplined approach to stock selection. Companies must meet the following criteria before stocks are considered for purchase: Consistent Dividend Increases, Substantial Dividend Increases, Strong Balance Sheets, Reinvested Earnings for Future Long-Term Growth, Attractive Price.
3	Royce Premier Fund	36	Each of our portfolio managers uses an active, bottom-up, risk-conscious, and fundamental investment approach...
4	Fenimore Asset Management Trust: FAM Value Fund	34	We are value investors and see stocks as economic interests in actual companies — and quality businesses garner our attention. We concentrate on small- to mid-cap companies that we think can grow over time and seek to purchase them at a discount to what we estimate they are worth. This is the "value" part of our philosophy.
5	Mairs & Power Growth Fund	25	Commits to long-term investing in consistently growing companies with minimal turnover. Seeks companies with long-term, durable competitive advantages at reasonable prices.

Table 2.9: Fund Identities Cont.

Rank	Fund Name	Occurrences	Investment Philosophy
6	Sequoia Fund	24	The Fund's investment objective is long-term growth of capital. A guiding principle is the consideration of equity securities, such as common stock, as units of ownership of a business and the purchase of them when the price appears low in relation to the value of the total enterprise.
7	Longleaf Partners Small-Cap Fund	23	We believe the key to our decades-long success has been high-conviction investing with a long-term time horizon in strong businesses with good people at deeply discounted prices.
8	Allianz Funds: AllianzGI NFJ Small-Cap Value Fund	22	For approximately 20 years, AllianzGI NFJ Small-Cap Value Fund has concentrated on dividend paying, small capitalization U.S. companies with long-term potential that has gone unrecognized by the market.
9	Ariel Appreciation Fund	21	Ariel's flagship value approach is built on the basic principle of targeting undervalued companies that show a strong potential for growth. We take advantage of the market's short-term thinking to optimize long-term results for our clients.
10	John Hancock Trust: Small Cap Value Trust	21	The manager employs a value-oriented investment approach in selecting stocks, using proprietary fundamental research to identify stocks the manager believes have distinct value characteristics based on industry-specific valuation criteria. The manager focuses on high-quality companies with a proven record of above-average rates of profitability that sell at a discount relative to the overall small-cap market.

I verify the funds' self-disclosed investment philosophies along two dimensions. I compare their portfolio ages with the average portfolio age across all managers in the industry to verify their claim of being long-term investors; I compute their loadings on AQR's Quality Minus Junk(QMJ) factor to verify their claim of being value investors who invest in quality business.

The age of a portfolio at any given point of time is defined as the average time that the constituent stocks that have been included in the portfolio. I then compute the average portfolio age of a fund by taking the time average of the portfolio age of the fund throughout the fund history. In my sample, the average portfolio age of the filtered Cell 1-1 portfolio is 3.30 years, whereas the average portfolio age for all funds in the cross section that are at least 5 years old is 1.86 years.[8] Therefore, apparently, the algorithm selected funds do hold stocks for much longer periods compared with an average fund in the industry. Thus, their claim of taking the long-term investment approach is verify.

Asness, Frazzini, and Pedersen (2014) defined a measure of firm quality from firm fundamentals, and constructed a Quality Minus Junk(QMJ) factor by longing the high quality firms and shorting the low quality firms. They showed that the QMJ factor demanded positive and significant risk premium in their sample. Frazzini, Kabiller, and Pedersen (2013) further showed that Warren Buffett's portfolio had high exposure to the QMJ factor, and the QMJ premium accounted for a non-trivial portion of Buffett's superior outperformance. Table 2.10 documents the results when the performance of the filtered Cell 1-1 portfolio is regressed against the benchmark including the QMJ factor in addition to the Carhart 4 factors. From the table, the QMJ loading is positive and statistically significant. The benchmark adjusted annualized alpha of the portfolio drops from 3.06%(1.94%) to 1.64%(0.56%) before(after) fees, compared with the regression on the Carhart 4 factors alone. Therefore, the selected funds load significantly on the QMJ factor and the QMJ premium accounted for a large part of their benchmark adjusted outperformance, verifying their claim that they aim to invest in quality business through careful fundamental research.

Warren Buffett, as a successful individual investor, has been a prominent figure in the financial world for decades. Alas, rigorous academic research on his success has

[8] I require the funds to be at least 5 years old in the comparison to eliminate the mechanical downward bias on portfolio age from the very young funds. But the effect of such young funds is small. The average portfolio age in the cross section is 1.67 years when all funds are included.

Table 2.10: QMJ Exposure

This table documents the results when the performance of the Cell 1-1 portfolio is regressed against the benchmark including the QMJ factor in addition to the Carhart 4 factors. α is annualized and in percentage. The Carhart 4 factors are from Prof. Ken French's website. The QMJ factor is from AQR's website. SR stands for annualized Sharpe ratio of the portfolio performance. IR stands for the annualized information ratio of the portfolio relative to the benchmark.

Panel A: Before-fees Performance

α (%)	Market	Value	Size	Momentum	QMJ	R-square (%)	SR	IR
1.64*	0.89***	0.25***	0.25***	−0.05***	0.19***	94	0.59	0.49
[1.94]	[42.70]	[10.22]	[11.60]	[−3.53]	[5.32]			

Panel B: After-fees Performance

α (%)	Market	Value	Size	Momentum	QMJ	R-square (%)	SR	IR
0.56	0.90***	0.25***	0.26***	−0.05***	0.18***	95	0.51	0.17
[0.68]	[44.76]	[10.65]	[12.56]	[−3.69]	[5.36]			

been limited to case studies (see Frazzini, Kabiller, and Pedersen (2013), Chirkova (2012), Martin and Puthenpurackal (2008), etc.), as informally identifying investors who claim to embrace the Buffett approach suffers from selection biases. This paper, on the other hand, develops a quantitative algorithm to systematically identify Buffett-like investors ex ante, and provides statistical support that they do seem to possess the skills to outperform the market as advertised.

2.5.1 Timing of the Outperformance

During my sample period, the algorithm selected managers were able to outperform the Carhart benchmark by a very large margin. But is the outperformance evenly distributed over the years or rather concentrated over certain periods? To answer this question, Figure 2.1 compares the performances between the selected managers, Berkshire Hathaway, and the market during the sample period. Specifically, the blue solid line plots the log cumulative (before-fees) return of the constructed fund portfolio, the green broken line and the black dotted line plot the log cumulative returns of Berkshire Hathaway and the market, respectively. From the figure, it is obvious that the outperformance of the fund portfolio and Berkshire Hathaway over the market is concentrated during the 1998-2003 period, which coincides with the dot-com bubble period. In other words, it seems that the Buffett-like mutual fund managers and Buffett himself were able to beat the market during the sample period because they were able to avoid the internet bubble.

Figure 2.3 plots the cross-sectional standard deviation of the residuals in the Carhart regressions of all the mutual funds in the cross section. In other words, this figure plots the dispersion of idiosyncratic returns of all the mutual funds in the cross section adjusted for the Carhart benchmark. The dispersion of idiosyncratic returns can be regarded as an indicator of profitable investment opportunities in the market. Also, a manager's performance during high dispersion periods might be more indicative of his or her skills. From the figure, the dispersion of idiosyncratic returns peaked around 2001. Interestingly, the magnitude of the peak during the internet bubble period dwarfs the peak during the 2008 financial crisis.

Comparing Figure 2.1 and Figure 2.3, it is reasonable to make the conjecture that the Buffett-like managers and Warren Buffett himself were able to beat the market during periods when lucrative investment opportunities abound and money managers' skills were most influential on their performances. However, whether the superior performances of the Buffett school during the 1998-2003 period is a single incident, or is a pattern that is going to repeat itself remains an open question that only time can tell.

Figure 2.1: Performance Comparison

This figure compares the performances between the selected managers, Berkshire Hathaway, and the market during the sample period. Specifically, the blue solid line plots the log cumulative (before-fees) return of the constructed fund portfolio, the black broken line and the red dotted line plot the log cumulative returns of Berkshire Hathaway and the market, respectively.

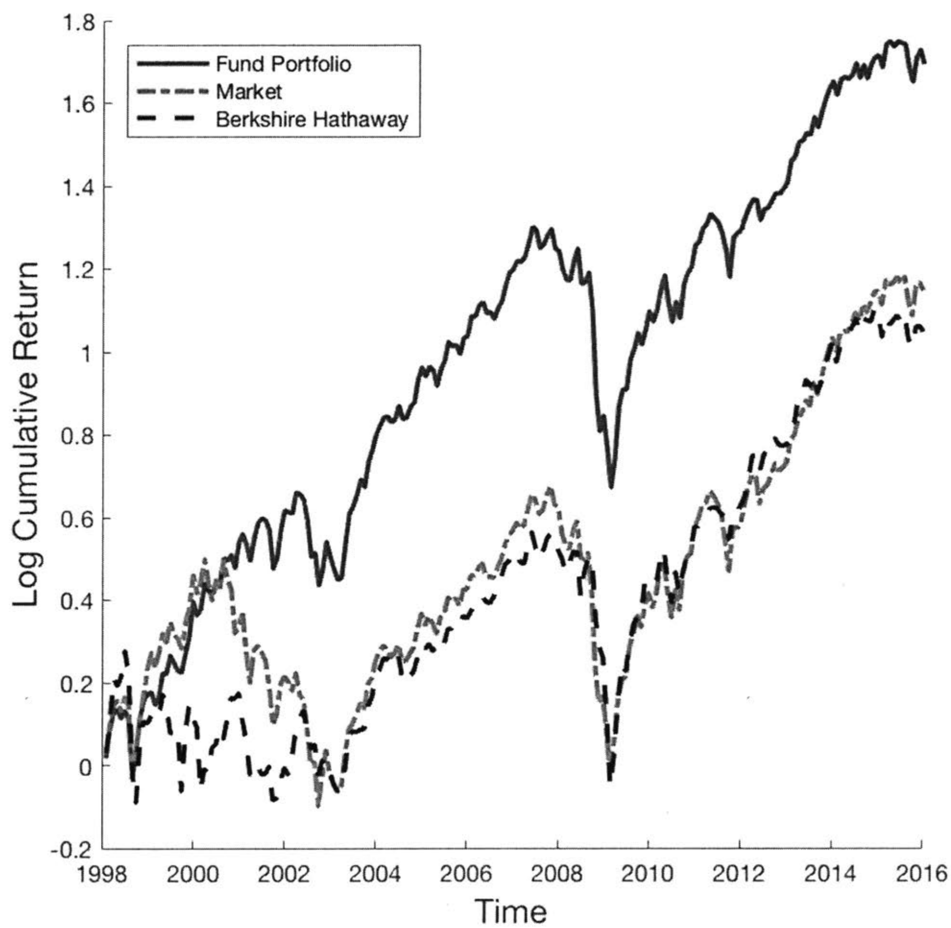


Figure 2.2: Log Cumulative Excess Return

This figure plots the time series of the log cumulative excess return of the selected Cell 1-1 funds. The log cumulative excess return is defined as the difference between the log cumulative before-fees return of the selected funds and the log cumulative return of the market.

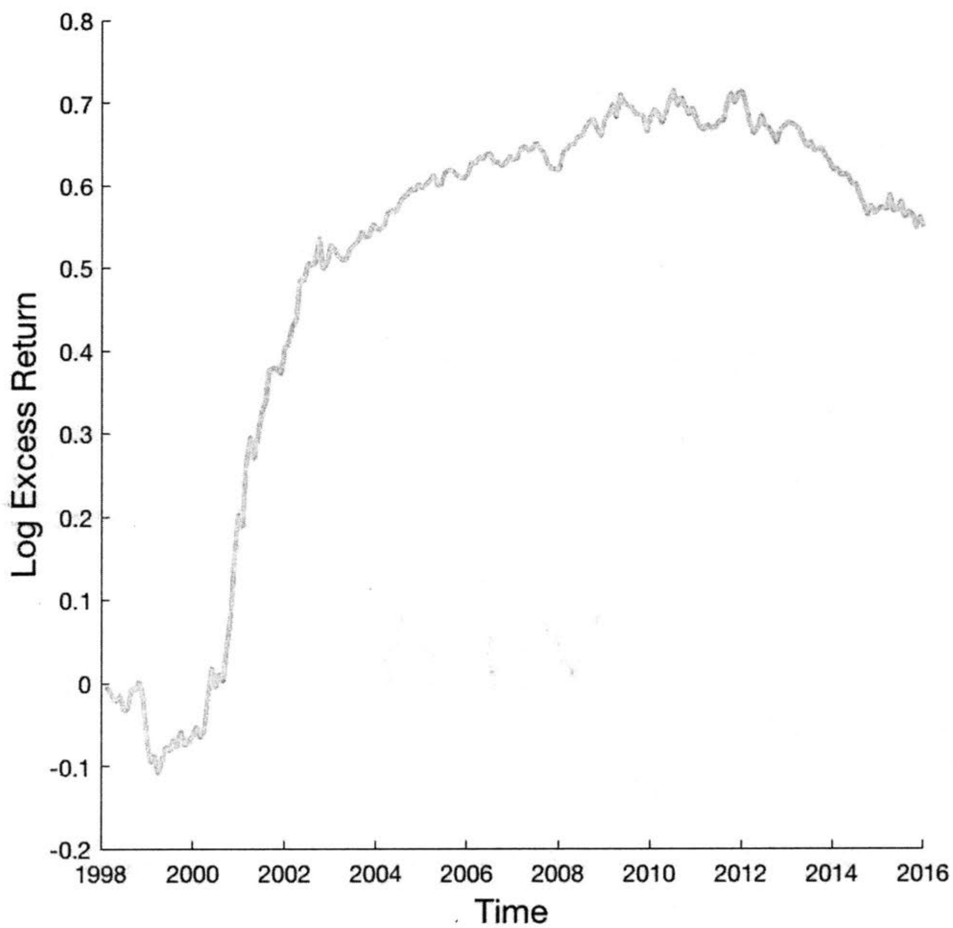
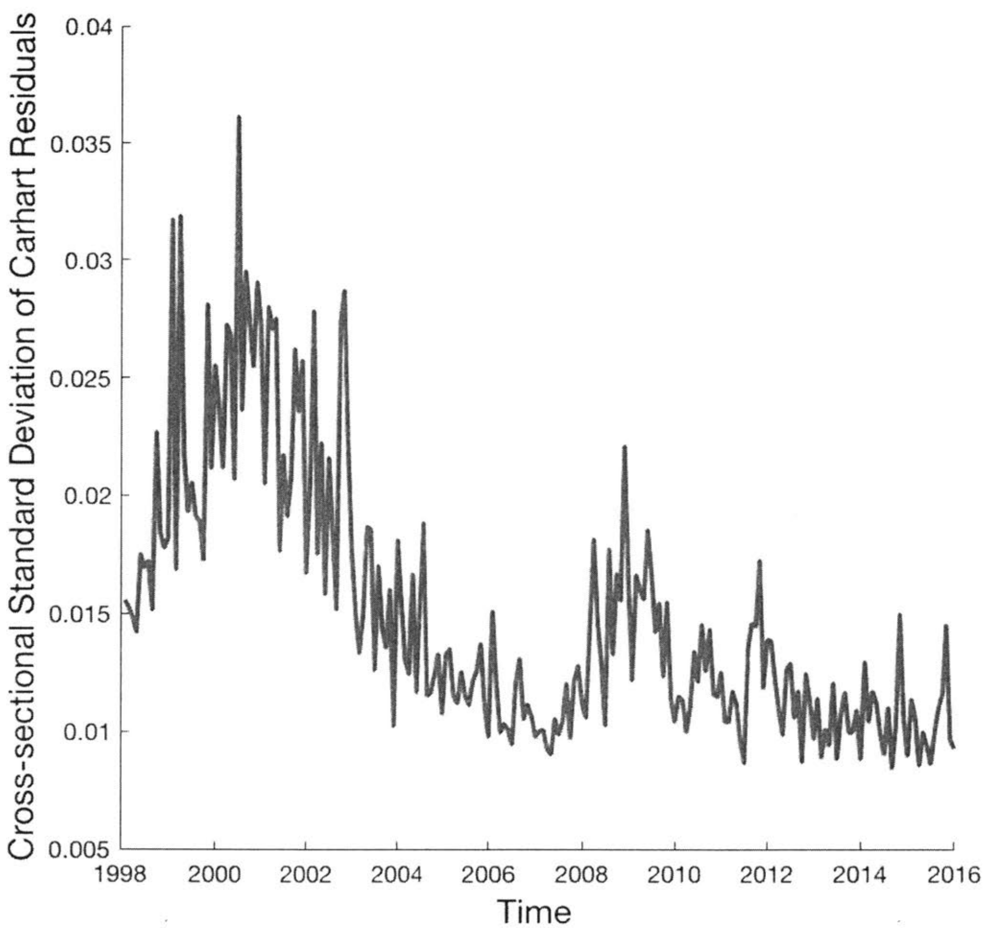


Figure 2.3: Residual Dispersion
This figure plots the cross-sectional standard deviation of the residuals in the Carhart regressions of all the mutual funds in the cross section. The Carhart regression is conducted on a rolling window that consists of 48 consecutive months. Only funds with available returns for at least 24 out of the 48 months are included in the regressions.



2.6 Robustness to Internet Stocks

One concern regarding the robustness of the outperformances of the identified Buffett-like managers is that they were simply lucky to sit out the internet stocks when the internet bubble bursted in early 2000. I perform a number of robustness checks to rule out this possibility.

Table 2.11, Table 2.12 and Table 2.13 redo the exercises in Table 2.5, Table 2.6 and Table 2.7, with the difference being that the Herding and Trading Intensity measures are now defined based on non-internet stocks only.[9] The tables show that the definition of Herding and Trading Intensity are not sensitive to the exclusion of the internet stocks.

Table 2.13: Characteristics of the Internet-stock-robust Anti-herding Funds

This table documents the characteristics of the 8 Trading Intensity portfolios among the most Anti-herding funds (column 1 in Table 2.11). Both measures are constructed without internet stocks. Age is the average age of the funds within the portfolio. Size is the aggregate TNA of the portfolio normalized by all the aggregate TNA of all the portfolios in the cross-section. If all the portfolios are of equal total TNA, the Size would be $1/64 = 0.01525$. NStocks is the average number of stocks held by the funds within the portfolio.

TI Rank	Age	Size	NStocks
1	14.86	0.0210	100.34
2	12.38	0.0125	137.89
3	11.76	0.0097	107.26
4	10.96	0.0083	114.59
5	10.59	0.0086	104.39
6	10.81	0.0064	109.98
7	9.76	0.0063	136.03
8	8.99	0.0028	115.30

Tables 2.14 - 2.17 conduct a different type of exercise. I investigate whether the identified fund managers are able to outperform a hypothetical investor who simply avoids internet stocks. In order to answer this question, in Table 2.14 and Table 2.15, I replace the Fama French market factor with the market factor constructed as the value-weighted portfolio of all NYSE, AMEX and NASDAQ stocks excluding the internet sector. To be even more conservative, Table 2.16 and Table 2.17 remove all NASDAQ stocks and define the market factor as the value-weighted portfolio of all

[9] I define internet stocks following Loughran and Ritter (2004) and Ljungqvist and Wilhelm (2003). High-tech stocks are defined as stocks with the following SIC codes: 3571, 3572, 3575, 3577, 3578 (computer hardware), 3661, 3663, 3669 (communications equipment), 3674 (electronics), 3812 (navigation equipment), 3823, 3825, 3826, 3827, 3829 (measuring and controlling devices), 4899 (communication services), and 7370, 7371, 7372, 7373, 7374, 7375, 7378, and 7379 (software).

Table 2.11: Double-sort Alphas w/o Internet Stocks (Before-fees)

This table documents the (before-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. Both measures are constructed without internet stocks. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

Before-fees Alphas								
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	2.56***	1.49*	1.16*	0.79*	0.15	0.36	0.39	0.30
	[2.92]	[1.76]	[1.85]	[1.86]	[0.38]	[0.79]	[0.65]	[0.45]
2	2.02**	1.26	0.70	−0.96*	0.02	−0.00	0.56	0.21
	[2.64]	[1.40]	[0.94]	[−1.89]	[0.05]	[−0.01]	[0.96]	[0.32]
3	1.41*	1.46	1.50	0.05	0.03	0.02	−0.06	0.80
	[1.71]	[1.58]	[2.16]	[0.09]	[0.05]	[0.04]	[−0.1]	[1.23]
4	1.73**	0.83	0.63	−0.09	−0.21	0.48	0.62	−0.07
	[2.21]	[0.87]	[0.74]	[−0.15]	[−0.35]	[0.97]	[1.22]	[−0.09]
5	1.06	0.47	0.34	0.11	−0.04	0.70	−1.08	0.09
	[1.14]	[0.49]	[0.40]	[0.15]	[−0.08]	[1.12]	[−1.77]	[0.11]
6	1.41*	0.03	−0.00	0.43	−0.82	0.04	−1.07	−1.21
	[1.77]	[0.03]	[−0.00]	[0.48]	[−1.07]	[0.07]	[−1.50]	[−1.38]
7	1.72**	0.42	0.78	−0.69	−0.11	−0.31	−0.36	−0.87
	[2.04]	[0.42]	[0.86]	[−0.82]	[−0.13]	[−0.45]	[−0.45]	[−0.88]
8	1.38	1.10	−0.30	−0.51	−0.34	0.54	0.07	−0.60
	[1.37]	[1.07]	[−0.28]	[−0.49]	[−0.34]	[0.62]	[0.06]	[−0.54]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

Table 2.12: Double-sort Alphas w/o Internet Stocks (After-fees)

This table documents the (after-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. Both measures are constructed without internet stocks. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

After-fees Alphas								
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	1.50*	0.36	0.04	−0.28	−0.90**	−0.69	−0.55	−0.98
	[1.78]	[0.43]	[0.06]	[−0.70]	[−2.26]	[−1.51]	[−0.90]	[−1.50]
2	0.82	−0.03	−0.18	−2.01***	−1.16**	−1.08**	−0.45	−0.87
	[1.08]	[−0.04]	[−0.25]	[−3.92]	[−2.22]	[−2.29]	[−0.78]	[−1.32]
3	0.02	0.16	0.32	−1.13*	−1.06**	−1.17**	−1.25**	−0.49
	[0.02]	[0.17]	[0.45]	[−1.94]	[−2.03]	[−2.19]	[−2.08]	[−0.75]
4	0.30	−0.70	−0.54	−1.25*	−1.38**	−0.59	−0.63	−1.38*
	[0.38]	[−0.74]	[−0.64]	[−1.94]	[−2.41]	[−1.18]	[−1.27]	[−1.74]
5	−0.38	−0.70	−0.95	−0.92	−1.11**	−0.53	−2.27***	−1.18
	[−0.41]	[−0.71]	[−1.13]	[−1.25]	[−2.16]	[−0.82]	[−3.67]	[−1.56]
6	0.03	−1.20	−1.46*	−0.75	−1.88**	−1.15*	−2.28***	−2.39***
	[0.03]	[−1.36]	[−1.73]	[−0.84]	[−2.34]	[−1.88]	[−3.29]	[−2.73]
7	0.39	−0.95	−0.84	−1.89**	−1.40*	−1.59**	−1.79**	−2.19**
	[0.46]	[−0.96]	[−0.93]	[−2.27]	[−1.74]	[−2.34]	[−2.27]	[−2.19]
8	0.02	−0.31	−1.72	−2.05**	−1.74*	−0.82	−1.27	−1.82*
	[0.02]	[−0.30]	[−1.58]	[−1.96]	[−1.70]	[−0.95]	[−1.14]	[−1.67]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

NYSE and AMEX stocks only. The numbers in these tables are similar to the results of the standard Fama French regression, showing that simply avoiding the internet stocks cannot explain these Buffett-like managers' superior performances.

2.7 Additional Findings

One surprising finding from the data is that the group of the ultra-performing Buffett-like managers charges very low fees. On average, they only charge 112 bps per year, which is even lower than the industry average of 124 bps per year during the sample period. The amount of fees is also shockingly low when compared with managers' before-fees performances. They were able to beat the Carhart benchmark by 3.06% per year before fees. So they leave investors the after-fees alpha as high as 1.94% per year. This is at odds with the equilibrium described in Berk and Green (2004). In the Berk and Green world, mutual fund managers hold perfect bargaining power against their investors so that they can charge fees as high as their before-fees alpha, leaving investors only breaking even with the market after fees. It is beyond the scope of this paper to rigorously determine the cause of the low fees.

Another interesting finding is that investment strategies of the identified managers can be easily replicated. For an investor who mechanically implemented the strategy to invest in the 3-month old lagged portfolio compositions of the group of Buffett-like managers, he or she would be able to earn a Carhart 4-factor alpha of 3.52% per year, which is comparable to the before-fees alpha that the managers were able to achieve themselves.[10] Moreover, the correlation between the Carhart residuals of the managers' raw performances and the implementable replicating strategy is 90%. Essentially, one can easily free-ride on the Buffett-like managers' research results by simply investing in their portfolio holdings when they are required to disclose them by regulation. Therefore, the very nature of the long-term investment philosophy, though profitable, also leaves the managers vulnerable to free-riders.

[10]The portfolio compositions have to be 3-month old to ensure that they are publicly available at the time of portfolio construction.

Table 2.14: Double-sort Alphas with Alternative Market 1 (Before-fees)

This table documents the (before-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. The market factor in the benchmark is constructed as the value-weighted portfolio of all NYSE, AMEX and NASDAQ stocks excluding the internet stocks. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

	Before-fees Alphas							
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	3.14***	1.32	0.69	0.82	0.35	−0.03	1.04	0.55
	[4.01]	[1.62]	[1.18]	[1.59]	[0.59]	[−0.06]	[1.43]	[0.67]
2	2.00***	1.41	1.38	0.07	−0.10	1.08	−0.20	0.63
	[2.63]	[1.61]	[1.83]	[0.10]	[−0.16]	[1.55]	[−0.28]	[0.73]
3	2.44***	1.57	1.64**	0.44	0.08	0.20	0.84	0.04
	[2.90]	[1.60]	[2.72]	[0.69]	[0.13]	[0.33]	[1.24]	[0.05]
4	2.28***	0.27	0.35	0.53	−1.02*	0.49	0.72	0.30
	[2.72]	[0.29]	[0.42]	[0.79]	[−1.73]	[0.76]	[1.08]	[0.30]
5	1.67*	1.01	0.16	0.63	0.11	0.24	−0.14	−0.14
	[1.78]	[1.22]	[0.16]	[0.82]	[0.19]	[0.36]	[−0.17]	[−0.14]
6	1.91**	0.74	0.41	0.32	0.01	0.50	−0.02	−0.47
	[2.14]	[0.81]	[0.42]	[0.36]	[0.02]	[0.60]	[−0.03]	[−0.45]
7	1.79**	−0.13	0.35	−0.71	0.40	−0.72	−0.58	−0.12
	[2.05]	[−0.15]	[0.38]	[−0.72]	[0.49]	[−0.92]	[−0.71]	[−0.11]
8	1.64*	0.39	0.10	−0.64	−0.09	0.11	−0.04	−0.38
	[1.67]	[0.40]	[0.08]	[−0.59]	[−0.09]	[0.13]	[−0.04]	[−0.26]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

Table 2.15: Double-sort Alphas with Alternative Market 1 (After-fees)

This table documents the (after-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. The market factor in the benchmark is constructed as the value-weighted portfolio of all NYSE, AMEX and NASDAQ stocks excluding the internet stocks. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

After-fees Alphas								
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	2.03***	0.26	−0.17	−0.19	−0.73	−1.03*	−0.21	−0.68
	[2.70]	[0.32]	[−0.30]	[−0.38]	[−1.25]	[−1.81]	[−0.30]	[−0.83]
2	0.74	0.18	0.27	−1.02	−1.24**	0.06	−1.29*	−0.61
	[1.00]	[0.20]	[0.36]	[−1.59]	[−2.11]	[0.08]	[−1.87]	[−0.70]
3	1.05	0.25	0.43	−0.67	−0.99	−1.02*	−0.31	−1.32
	[1.27]	[0.26]	[0.71]	[−1.07]	[−1.56]	[−1.68]	[−0.45]	[−1.53]
4	0.86	−1.07	−0.79	−0.50	−2.25***	−0.64	−0.43	−0.92
	[1.03]	[−1.16]	[−0.95]	[−0.74]	[−3.75]	[−1.01]	[−0.64]	[−0.92]
5	0.12	−0.17	−1.21	−0.43	−0.96	−0.92	−1.29	−1.38
	[0.12]	[−0.20]	[−1.25]	[−0.57]	[−1.60]	[−1.36]	[−1.55]	[−1.40]
6	0.43	−0.58	−0.94	−0.80	−1.06	−0.72	−1.20	−1.85*
	[0.50]	[−0.62]	[−1.00]	[−0.91]	[−1.51]	[−0.87]	[−1.53]	[−1.74]
7	0.40	−1.41	−1.20	−2.09**	−0.85	−2.14**	−1.75**	−1.41
	[0.46]	[−1.59]	[−1.31]	[−2.13]	[−1.05]	[−2.68]	[−2.18]	[−1.31]
8	0.30	−0.97	−1.36	−2.09*	−1.48	−1.31	−1.53	−1.67
	[0.30]	[−0.97]	[−1.11]	[−1.94]	[−1.40]	[−1.47]	[−1.60]	[−1.19]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

Table 2.16: Double-sort Alphas with Alternative Market 2 (Before-fees)

This table documents the (before-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. The market factor in the benchmark is constructed as the value-weighted portfolio of all NYSE and AMEX stocks. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

Before-fees Alphas								
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	3.15***	1.31*	0.72	0.87	0.41	0.01	1.07	0.60
	[4.21]	[1.77]	[1.22]	[1.49]	[0.61]	[0.02]	[1.49]	[0.69]
2	2.03***	1.44*	1.43	0.12	−0.04	1.14	−0.14	0.70
	[2.66]	[1.66]	[1.81]	[0.16]	[−0.06]	[1.48]	[−0.18]	[0.75]
3	2.48***	1.62	1.69***	0.47	0.13	0.26	0.89	0.10
	[2.88]	[1.60]	[2.58]	[0.74]	[0.19]	[0.38]	[1.25]	[0.11]
4	2.33***	0.29	0.41	0.60	−0.96	0.57	0.78	0.37
	[2.64]	[0.32]	[0.46]	[0.78]	[−1.45]	[0.74]	[1.05]	[0.34]
5	1.71*	1.06	0.22	0.67	0.17	0.32	−0.09	−0.08
	[1.79]	[1.23]	[0.22]	[0.84]	[0.26]	[0.40]	[−0.10]	[−0.07]
6	1.96**	0.78	0.47	0.40	0.08	0.57	0.05	−0.39
	[2.12]	[0.85]	[0.46]	[0.41]	[0.10]	[0.62]	[0.06]	[−0.35]
7	1.83**	−0.09	0.42	−0.62	0.46	−0.67	−0.52	−0.06
	[2.04]	[−0.09]	[0.42]	[−0.57]	[0.52]	[−0.80]	[−0.59]	[−0.06]
8	1.68*	0.47	0.15	−0.55	−0.01	0.17	0.03	−0.26
	[1.69]	[0.44]	[0.12]	[−0.46]	[−0.01]	[0.18]	[0.02]	[−0.17]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

Table 2.17: Double-sort Alphas with Alternative Market 2 (After-fees)

This table documents the (after-fees) Carhart 4-factor alphas of the 64 portfolio of active mutual funds, first sorted by Herding then sorted by Trading Intensity. The market factor in the benchmark is constructed as the value-weighted portfolio of all NYSE and AMEX stocks. Each column of the table corresponds to a group of funds categorized by Herding; and each row corresponds to the rank sorted by Trading Intensity. The numbers are in percentage, annualized.

	After-fees Alphas							
1st stage (Herding)	1	2	3	4	5	6	7	8
2nd stage (Trading)								
1	2.03***	0.25	−0.14	−0.14	−0.67	−0.98	−0.19	−0.63
	[2.86]	[0.35]	[−0.24]	[−0.25]	[−0.98]	[−1.60]	[−0.26]	[−0.73]
2	0.77	0.20	0.31	−0.97	−1.19*	0.11	−1.23	−0.54
	[1.03]	[0.23]	[0.40]	[−1.42]	[−1.78]	[0.15]	[−1.61]	[−0.58]
3	1.09	0.30	0.48	−0.64	−0.94	−0.96	−0.26	−1.26
	[1.29]	[0.30]	[0.73]	[−1.00]	[−1.37]	[−1.39]	[−0.37]	[−1.37]
4	0.91	−1.06	−0.74	−0.43	−2.19***	−0.56	−0.37	−0.84
	[1.04]	[−1.20]	[−0.84]	[−0.55]	[−3.19]	[−0.74]	[−0.49]	[−0.77]
5	0.17	−0.12	−1.15	−0.39	−0.90	−0.84	−1.23	−1.32
	[0.17]	[−0.14]	[−1.12]	[−0.49]	[−1.35]	[−1.06]	[−1.42]	[−1.25]
6	0.48	−0.55	−0.88	−0.73	−1.00	−0.65	−1.12	−1.78
	[0.53]	[−0.59]	[−0.89]	[−0.76]	[−1.27]	[−0.73]	[−1.28]	[−1.56]
7	0.44	−1.36	−1.13	−2.00	−0.78	−2.08**	−1.69**	−1.36
	[0.50]	[−1.46]	[−1.15]	[−1.84]	[−0.88]	[−2.47]	[−1.97]	[−1.21]
8	0.34	−0.90	−1.31	−2.00*	−1.40	−1.26	−1.47	−1.54
	[0.34]	[−0.82]	[−1.05]	[−1.69]	[−1.21]	[−1.33]	[−1.44]	[−1.01]

"***" ~ significant at 1% level
"**" ~ significant at 5% level
"*" ~ significant at 10% level

2.8 Conclusion

By installing Herding and Trading Intensity as two filters, I managed to develop a systematic process to identify a group of mutual fund managers who embrace the long-term, fundamental investment philosophy exemplified by Warren Buffett. I show that over the past 20 years, the group of Buffett-like managers were able to outperform the Carhart 4-factor benchmark by 3.06%(1.94%) before(after) fees per year - a magnitude that is both statistically and economically significant. Moreover, rather than evenly spreading out, the outperformances of the Buffett-like managers were concentrated over the 1998-2003 period when there was more cross-sectional heterogeneity in the performances of the entire mutual fund industry. Finally, not only the group of Buffett-like managers were able to deliver high alphas before fees, they also charge lower fees compared with an average manager in the industry. I further show that one can easily replicate their before-fees performances by simply investing in their lagged portfolio holdings.

References

Asness, Clifford S, Andrea Frazzini, and Lasse Heje Pedersen. 2014. "Quality minus junk." *Available at SSRN 2312432* .

Berk, Jonathan B and Richard C Green. 2004. "Mutual Fund Flows and Performance in Rational Markets." *Journal of Political Economy* 112 (6):1269–1295.

Brown, Stephen J and William N Goetzmann. 1995. "Performance persistence." *The Journal of finance* 50 (2):679–698.

Buffett, Hy Warren E. 1984. "The Superi nvestors of Graham-and-D0 ddsville." .

Carhart, Mark M. 1997. "On persistence in mutual fund performance." *The Journal of finance* 52 (1):57–82.

Chirkova, Elena. 2012. "Why is It that I am not Warren Buffett?" *Available at SSRN 2088786* .

Cohen, Randolph B, Joshua D Coval, and L'uboš Pástor. 2005. "Judging fund managers by the company they keep." *The Journal of Finance* 60 (3):1057–1096.

Cremers, KJ Martijn and Antti Petajisto. 2009. "How active is your fund manager? A new measure that predicts performance." *Review of Financial Studies* 22 (9):3329–3365.

Doshi, Hitesh, Redouane Elkamhi, and Mikhail Simutin. 2015. "Managerial activeness and mutual fund performance." *Review of Asset Pricing Studies* 5 (2):156–184.

———. 2010. "Luck versus skill in the cross-section of mutual fund returns." *The journal of finance* 65 (5):1915–1947.

Frazzini, Andrea, David Kabiller, and Lasse H Pedersen. 2013. "Buffett's alpha." .

Grinblatt, Mark, Sheridan Titman, and Russ Wermers. 1995. "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior." *The American economic review* :1088–1105.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser. 1993. "Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988." *The Journal of finance* 48 (1):93–130.

Jegadeesh, Narasimhan and Sheridan Titman. 1993. "Returns to buying winners and selling losers: Implications for stock market efficiency." *The Journal of finance* 48 (1):65–91.

———. 1969. "Risk, the pricing of capital assets, and the evaluation of investment portfolios." *The Journal of Business* 42 (2):167–247.

Jiang, Hao and Michela Verardo. 2013. "Does herding behavior reveal skill? An analysis of mutual fund performance." *An Analysis of Mutual Fund Performance (March 2013)* .

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. 2005. "On the industry concentration of actively managed equity mutual funds." *The Journal of Finance* 60 (4):1983–2011.

———. 2008. "Unobserved actions of mutual funds." *Review of Financial Studies* 21 (6):2379–2416.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White. 2006. "Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis." *The Journal of finance* 61 (6):2551–2595.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny. 1992. "The impact of institutional trading on stock prices." *Journal of financial economics* 32 (1):23–43.

Linnainmaa, Juhani T. 2013. "Reverse survivorship bias." *The Journal of Finance* 68 (3):789–813.

Ljungqvist, Alexander and William J Wilhelm. 2003. "IPO pricing in the dot-com bubble." *The Journal of Finance* 58 (2):723–752.

Loughran, Tim and Jay Ritter. 2004. "Why has IPO underpricing changed over time?" *Financial management* :5–37.

Malkiel, Burton G and Eugene F Fama. 1970. "Efficient capital markets: A review of theory and empirical work." *The journal of Finance* 25 (2):383–417.

Martin, Gerald S and John Puthenpurackal. 2008. "Imitation is the sincerest form of flattery: Warren Buffett and Berkshire Hathaway." *Available at SSRN 806246* .

Statman, Meir and Jonathan Scheid. 2002. "Buffett in foresight and hindsight." *Financial Analysts Journal* 58 (4):11–18.

Wermers, Russ. 1997. "Momentum investment strategies of mutual funds, performance persistence, and survivorship bias." *University of Colorado. Working Paper* .

———. 1999. "Mutual fund herding and the impact on stock prices." *The Journal of Finance* 54 (2):581–622.

Chapter 3

A Global Risk Factor Is Needed to Price Global Assets

3.1 Introduction

Modern asset pricing theories often start from the Euler equation. The Fundamental Theorem of Asset Pricing (FTAP) states that absence of arbitrage is equivalent to the existence of a positive stochastic discount factor (SDF) which prices all existing traded assets. However, the SDF is not directly observable without additional assumptions. Due to its importance in pricing assets, lots of efforts have been made to recover the unobservable SDF either semi-parametrically or using structural models with equilibrium conditions. For example, Ross (2015) shows that the SDF can be recovered from the prices of financial derivatives in an economy with time homogeneity; whereas Mehra and Prescott (1985) proposes the equity risk premium puzzle by relating the SDF to the marginal utility of consumption of the representative agent. I show, in this paper, that in the context of international finance, countries' SDFs can be recovered from exchange rates under the following three assumptions: 1) the Euler equation holds internationally; 2) the movements of exchange rates are driven by a few common factors; 3) there does not exist a special global risk factor which has identical influence on all countries' SDFs. I then design an empirical test using exchange rates and equity returns of 28 countries from 1988 to August of 2014 to show that the SDFs recovered from exchange rates are unable to price these countries' equities. The failure the exchange-rate-recovered SDFs to price countries' assets reflects the violation of my assumptions. In particular, the test result highlights the importance of a global risk factor with identical influence on all countries' SDFs to price assets in an arbitrage-free international financial market.

My test methodology relies on the close link between bilateral exchange rates and SDFs implied by the Euler equation[1] (see Bansal (1997), Bekaert (1996)). Since exchange rates are observable and SDFs can be used to form asset pricing moment conditions with the Euler equation, such an implication on exchange rates makes the recovery of SDF possible. However, the relation between exchange rates and countries' SDFs is less than ideal for a direct empirical test because exchange rates are bilateral so that the SDFs of two countries are intertwined in one rate between them. Yet, recent discussions in the foreign exchange rate literature, such as Lustig et al. (2011), Lustig et al. (2014), and Verdelhan (2011), reveal the common factor structure among bilateral exchange rates. Thanks to these discoveries, I am able to disentangle country specific risk from bilateral exchange rates through factor analysis techniques so as to, at least, partially recover countries' SDFs to form a moment condition.

My recovery procedure consists of three steps. The first step is to convert bilateral exchange rates to country specific **unilateral appreciation rates** (UARs). The UAR is a new construction in this paper. It captures a country's currency appreciation relative to a basket of many different currencies. The second step is to recover factors from the UARs by performing robust principle component analysis. The last step is to form moment conditions according to the Euler equation. My analysis shows that in the absence of a global risk factor with identical influence on all countries' SDFs, the countries' SDFs can be represented with UARs and a linear combination of the currency factors.

The technique to extract common factors from exchange rates is similar to Litterman and Scheinkman (1991), where principle component analysis is applied to interest rates corresponding to different maturities. However, the direct application of PCA on exchange rates is less successful than on interest rates because PCA is susceptible to outliers so that large country specific idiosyncratic shocks can substantially blur the factor analysis through PCA. On the other hand, latest development in statistics and computer science suggests the possibility of new PCA techniques that are robust

[1]According to the Euler equation: $\frac{M_{i,t+1}}{M_{j,t+1}} = \frac{Q^j_{i,t+1}}{Q^j_{i,t}}$, where $M_{i,t+1}$ and $M_{j,t+1}$ are the SDFs of country i and j respectively; and Q^j_i is the bilateral exchange rate between the two countries.

to outliers and data corruption. Here I follow Candès et al. (2011) to perform robust PCA on UARs. The robust PCA technique proves to be reliable with the compositions of exchange rate factors being stable during the financial crisis. Since the exchange rate factors have to be recovered from robust PCA in the first stage rather than being directly observable, the distribution of the J test statistic is non-standard. I resolve this issue by designing a block bootstrap procedure which preserves the covariances among the exchange rates and equity returns, and takes into account of the approximations in the theoretical derivations.

In terms of related literature, the factor analysis component of my project is mostly related to the work of Lustig et al. (2011), Lustig et al. (2014), and Verdelhan (2011), with the difference being that they manage to extract factors from exchange rates by forming portfolios whereas I take advantage of the latest robust PCA technique. The comparisons between my approach and theirs are: 1) robust PCA does not require prior knowledge of the sorting characteristics in order to form portfolios so that the recovered factors contain stronger economic signals if the sorting characteristics are not perfectly aligned with the underlying economic forces in exchange rates ; 2) robust PCA is able to reveal the total number of factors that are present among exchange rates as well as their relative importance. The testing component of my project is similar to Mehra and Prescott (1985), where I replace their SDF, constructed from marginal utility of consumption, with the ones that are recovered from exchange rates. Finally, the attempt to empirically test the international Euler equation with exchange rates echoes the earlier finding by Brandt et al. (2006), where they show that the international Euler equation implies that the correlations between countries SDFs have to be high given the magnitudes of the observed equity risk premia.

The rest of the paper is arranged as follows. Section 3.2 presents the theoretical derivations of my test. Section 3.3 documents the factor structure among exchange rates and equity returns during the sample period as well as the empirical results of the test. Section 3.4 summarizes the empirical findings and concludes the paper.

3.2 Design of the Test

3.2.1 The Factor Structure among Log SDFs

Consider an economy with N countries. Denote $M_{i,t}$ as the SDF of country i at time t; and $Q^j_{i,t}$ as the bilateral exchange rate between country i and j at time t, defined as the number of units of country j's currency that one unit of country i's currency can buy. The Euler equation implies:

$$\frac{M_{i,t+1}}{M_{j,t+1}} = \frac{Q^j_{i,t+1}}{Q^j_{i,t}} \tag{3.1}$$

It is important to note that Equation (3.1) holds in both complete markets and incomplete markets. In complete markets where all contingent claims are available, the SDFs are unique and must satisfy Equation (3.1). However, even in incomplete markets where SDFs are not unique, Equation (3.1) still holds for one set of valid SDFs. Therefore, Equation (3.1) only comes from the assumption that risks in different countries are accessible to an international investor and there is no arbitrage.

Taking log on both sides:

$$m_{i,t+1} - m_{j,t+1} = q^j_{i,t+1} - q^j_{i,t} \equiv \Delta q^j_{i,t+1} \tag{3.2}$$

where m is the log SDF, q is the log exchange rate.

The difficulty to recover countries' SDFs from Equation 3.2 is that Equation 3.2 is a bilateral relation so that the SDFs of two countries are intertwined in one exchange rate between them. To extract the country specific risk of country i, I hold the base currency i constant in Equation (3.2) meanwhile averaging across all the counter-party countries in the economy:

$$m_{i,t+1} - \frac{1}{N}\sum_{j=1}^{N} m_{j,t+1} = \frac{1}{N}\sum_{j=1}^{N} \Delta q^j_{i,t+1}$$

Define $\bar{m} \equiv \frac{1}{N}\sum_{j=1}^{N} m_{j,t+1}$ as the average log SDF; and $\Delta q_i \equiv \frac{1}{N}\sum_{j=1}^{N} \Delta q^j_{i,t+1}$ as the **unilateral appreciation rate** (UAR henceforth) of country i. The last equation can be thus rewritten as:

$$m_{i,t+1} - \bar{m}_{t+1} = \Delta q_{i,t+1} \tag{3.3}$$

By averaging across counter-party countries, I turn a bilateral relation into a country specific relation so that I get one step closer towards recovery. A few important observations can be made from Equation (3.3). First, $\Delta q_{i,t+1}$ is directly observable from exchange rates. Therefore, Equation (3.3) states that in order to recover countries' log SDFs, the object of interest, one only needs to recover $\bar{m}_{t+1}$ - the average log SDF across many countries. Second, suppose there is a factor structure among exchange rates, then, due to the tight connection between the SDFs and the exchange rates, there should be a factor structure among log SDFs as well. Third, when the number of countries N is large enough, $\bar{m}_{t+1}$ would only contain systematic risks in log SDFs according to the Law of Large Number. And presumably, these systematic risks might also drive the movements in exchange rates so that they are potentially observable. To carefully investigate these observations, I impose a factor structure upon log SDFs.

I assume the following factor structure of log SDFs in the cross section:

$$m_{i,t} = f_{0,t}^q + \sum_{k=1}^{K} \beta_{i,k}^q f_{k,t}^q + \epsilon_{i,t}^q, \quad \forall i \in \{1, \cdots, N\} \tag{3.4}$$

The factor structure satisfies the following conditions:

1. $\forall i, \quad E\left(\epsilon_{i,t}^q\right) = 0.$
2. Orthogonality:
   (a) $\forall k_1,\ k_2 \in \{0, 1, \cdots, K\}, \quad Cov\left(f_{k_1,t}^q, f_{k_2,t}^q\right) = 0$
   (b) $\forall i,\ j \in \{1, \cdots, N\}, \quad Cov\left(\epsilon_{i,t}^q, \epsilon_{j,t}^q\right) = 0$
   (c) $\forall i \in \{1, \cdots, N\},\ k \in \{0, 1, \cdots, K\}, \quad Cov\left(f_{k,t}^q, \epsilon_{i,t}^q\right) = 0$
3. Normalization:
   (a) $\forall k \in \{0, 1, \cdots, K\}, \quad \frac{1}{N}\sum_{i=1}^{N} \beta_{i,k}^q = 1$
   (b) $\forall k \in \{1, 2, \cdots, K\},\ \exists i, j \in \{1, 2, \cdots, N\}, \quad \textit{so that } \beta_{i,k}^q \neq \beta_{j,k}^q$

The definition of the idiosyncratic risks and the orthogonality conditions are standard features of a factor structure. The normalization conditions distinguish between $f_{0,t}^q$ and other factors. For the lack of a better term, I refer to $\left\{f_{0,t}^q\right\}$ as the **Global Risk Factor** in this paper. This term is in effect a misnomer. The specialty of $f_{0,t}^q$ is not just that all countries have exposures to it, but, more strongly, that it has **identical** influence on all countries. In contrast to $f^q_{0,t}$, the other factors in the global financial market all have heterogeneous impacts on different countries.

I highlight the difference between $f^q_{0,t}$ and other factors because it is not observable from exchange rates. Indeed, exchange rates are related to the differences between log SDFs. Therefore, unlike other factors, $f^q_{0,t}$ does not drive the movements of exchange rates. However, the existence of $f^q_{0,t}$ is a knife-edged case. It requires all countries in the world to have the identical exposure to a systematic risk factor. I will posit the nonexistence of $f^q_{0,t}$ as a formal assumption later on in this exercise.

Economically, the factor structure is only meaningful when the number of the countries is greater than the number of the factors, i.e. $N > K$, so that the bulk of the risks of different countries' SDFs can be summarized by a few common factors. However, this is not required as an assumption. Yet, as I will show in the next section, this will be the case in data, consistent with the findings of Lustig et al. (2011), Lustig et al. (2014), and Verdelhan (2011) about the factor structure among bilateral exchanged rates.

With the factor structure representation in Equation (3.4), it is easy to see that the average log SDF $\bar{m}$ only loads on systematic components $\{f^q_{k,t}\}^K_{k=0}$ when the number of countries N is large enough for the idiosyncratic shocks to cancel out.

$$\begin{aligned}
\bar{m}_t &= \frac{1}{N}\sum_{j=1}^{N} m_{j,t} \\
&= f^q_{0,t} + \frac{1}{N}\sum_{j=1}^{N}\left(\sum_{k=1}^{K}\beta^q_{j,k}f^q_{k,t} + \epsilon^q_{j,t}\right) \\
&= f^q_{0,t} + \sum_{k=1}^{K}\left(\frac{1}{N}\sum_{j=1}^{N}\beta^q_{j,k}f^q_{k,t}\right) + \frac{1}{N}\sum_{j=1}^{N}\epsilon^q_{j,t} \\
&= f^q_{0,t} + \sum_{k=1}^{K} f^q_{k,t} + \frac{1}{N}\sum_{j=1}^{N}\epsilon^q_{j,t} \\
&\approx f^q_{0,t} + \sum_{k=1}^{K} f^q_{k,t}
\end{aligned} \tag{3.5}$$

where the last equality is implied by the Law of Large Number.

Combining Equation (3.3), (3.4) and (3.5):

$$\begin{aligned}
\Delta q_{i,t} &= m_{i,t} - \bar{m}_t \\
&\approx \left(f_{0,t}^q + \sum_{k=1}^{K} \beta_{i,k}^q f_{k,t}^q + \epsilon_{i,t}^q \right) - \left(f_{0,t}^q + \sum_{k=1}^{K} f_{k,t}^q \right) \\
&= \sum_{k=1}^{K} \left(\beta_{i,k}^q - 1 \right) f_{k,t}^q + \epsilon_{i,t}^q
\end{aligned} \tag{3.6}$$

Equation (3.6) reveals that there is a factor structure among UARs as well, with $\left(\beta_{i,k}^q - 1\right)$ being the loading of country i's UAR on factor k. The factor structure of UARs is inherited from the factor structure of SDFs directly due to the close link between SDFs and exchange rates according to the Euler equation. Intuitively, the UAR averages across the changes in bilateral exchange rates (Δq_i^j) of all counter-party countries so that the idiosyncratic risks of these counter-parties are averaged out and only systematic risks and the country specific risk of the base currency are remained. Therefore, unlike the bilateral exchange rate, the unilateral appreciation rate of country i only embodies the specific risks of country i alone.

Equation (3.6) is at the center of the construction of my test. It speaks to the fact that the K systematic risk factors $\left\{f_{k,t}^q\right\}_{k=1}^K$ in log SDFs are potentially observable through factor analysis on the UARs $\{\Delta q_{i,t}\}_{i=1}^N$. Moreover, comparing Equation (3.6) from Equation (3.4), it is evident that the loading of $\{\Delta q_{i,t}\}_{i=1}^N$ on $\left\{f_{k,t}^q\right\}_{k=1}^K$ are closely related to the loadings of log SDFs $\{m_{i,t}\}_{i=1}^N$ on systematic factors $\left\{f_{k,t}^q\right\}_{k=1}^K$.

Combining Equation (3.4) and Equation (3.6):

$$\begin{aligned}
m_{i,t} &= f_{0,t}^q + \sum_{k=1}^{K} \beta_{i,k}^q f_{k,t}^q + \epsilon_{i,t}^q \\
&= f_{0,t}^q + \sum_{k=1}^{K} \left(\beta_{i,k}^q - 1 \right) f_{k,t}^q + \epsilon_{i,t}^q + \sum_{k=1}^{K} f_{k,t}^q \\
&\approx \Delta q_{i,t} + f_{0,t}^q + \sum_{k=1}^{K} f_{k,t}^q
\end{aligned} \tag{3.7}$$

Since the UARs $\{\Delta q_{i,t}\}_{i=1}^N$ are observable, the K systematic risks $\left\{f_{k,t}^q\right\}_{k=1}^K$ are potentially estimable, it is possible to recover the log SDFs $\{m_{i,t}\}_{i=1}^N$ when the global risk factor $\left\{f_{0,t}^q\right\}$ is absent.

Logic of the Test

So far, the logic of the test can be already summarized with the following equations:

$$m_{i,t} - \bar{m}_t = \Delta q_{i,t} \iff m_{i,t} = \Delta q_{i,t} + \bar{m}_t \tag{3.3'}$$

$$m_{i,t} = \Delta q_{i,t} + f^q_{0,t} + \sum_{k=1}^{K} f^q_{k,t} \tag{3.7}$$

$$m_{i,t} = f^q_{0,t} + \sum_{k=1}^{K} \beta^q_{i,k} f^q_{k,t} + \epsilon^q_{i,t} \tag{3.4}$$

$$\Delta q_{i,t} \approx \sum_{k=1}^{K} \left(\beta^q_{i,k} - 1\right) f^q_{k,t} + \epsilon^q_{i,t} \tag{3.6}$$

Equation (3.3') states that log SDF is simply the UAR with a compensation $\{\bar{m}_t\}$. Equation (3.7) further illustrates the the compensation $\{\bar{m}_t\}$ is in fact the systematic components in countries' log SDFs. Finally, Equation (3.4) and Equation (3.6) show that log SDFs and UARs share very similar factor structure in the cross section so that except for the common global risk factor $\left\{f^q_{0,t}\right\}$, the systematic factors in log SDFs $\left\{f^q_{k,t}\right\}_{k=1}^{K}$ can be recovered from factor analysis on the UARs. Therefore, the logic of the test is to recover the SDFs by conservatively estimating the systematic compensation $\{\bar{m}_t\}$ from the UARs to best fit the Euler equation.

3.2.2 Design of the Test

According to Equation (3.7), the unobservable common global risk factor $\left\{f^q_{0,t}\right\}$ is the only object that prevents the log SDFs $\{m_{i,t}\}_{i=1}^{N}$ from being estimated with $\{\Delta q_{i,t}\}_{i=1}^{N}$ and $\left\{f^q_{k,t}\right\}_{k=1}^{K}$. I hereby simplify the problem by assuming away $\left\{f^q_{0,t}\right\}$ to derived the test.

Without the common global risk factor $\left\{f^q_{0,t}\right\}$, Equation (3.7) can be rewritten as

$$m_{i,t} \approx \Delta q_{i,t} + \sum_{k=1}^{K} f^q_{k,t} \tag{3.7'}$$

Equation (3.7') is very powerful. It implies the possibility to recover SDFs through UARs using the Euler equation. Therefore, any test built upon Equation (3.7') will be a test on the Euler equation, the factor structure in exchange rates, together with the simplifying assumption that $\{f^q_{0,t}\}$ does not exist.

With the help of Equation (3.7'), a moment condition is readily at hand by the unconditional Euler equation:

$$E\left[M_{i,t}R_{i,t}\right] \approx E\left[\exp\left(\Delta q_{i,t} + \sum_{k=1}^{K} f^q_{k,t}\right) R_{i,t}\right] \approx 1 \tag{3.8}$$

where $R_{i,t}$ can be any asset return in country i.

To finish the derivation, rewrite Equation (3.6) in the form of a PCA:

$$\begin{aligned} \Delta q_{i,t} &\approx \sum_{k=1}^{K} \left(\beta^q_{i,k} - 1\right) f^q_{k,t} + \epsilon^q_{i,t} \\ &= \sum_{k=1}^{K} \gamma^q_{i,k} \cdot PC^q_{k,t} + \epsilon^q_{i,t} \end{aligned} \tag{3.9}$$

where PC^q_k is the kth principal component, and $\gamma^q_{i,k}$ is Δq_i's loading on PC^q_k.

In PCA, PC_k is perfectly correlated with $f^q_{k,t}$, i.e. $\left|Corr\left(PC_{k,t}, f^q_{k,t}\right)\right| = 1$. Yet the magnitude of $f^q_{k,t}$ is indeterminate from PCA.

Specifically, define PC amplifier:

$$C^q_k \equiv \frac{f^q_{k,t}}{PC^q_{k,t}}$$

so that

$$f^q_{k,t} = C^q_k PC^q_{k,t} \tag{3.10}$$

The combination of Equation (3.8) and Equation (3.10) yields a estimable moment condition:

$$\begin{aligned} & E\left[\exp\left(\Delta q_{i,t} + \sum_{k=1}^{K} f^q_{k,t}\right) R_{i,t}\right] \approx 1 \\ \iff & E\left[\exp\left(\Delta q_{i,t} + \sum_{k=1}^{K} C^q_k PC^q_{k,t}\right) R_{i,t}\right] \approx 1 \\ \iff & E\left[\exp\left(\Delta q_{i,t} + \sum_{k=1}^{K} C^q_k PC^q_{k,t}\right) R_{i,t} - 1\right] \approx 0 \end{aligned} \tag{3.11}$$

Equation (3.11) is my moment condition where $\{C_k^q\}_{k=1}^K$ are treated as free the parameters to be estimated from the asset pricing formula (the Euler equation). Apparently, my moment condition is non-standard as the principal components of the UARs $\left\{PC_{k,t}^q\right\}_{k=1}^K$ are not directly observable and have to be estimated in the first stage. Nevertheless, bootstrap techniques would allow me to characterize the finite sample distribution of the test statistic so as to constitute a valid test.

3.2.3 Intuition of the Test

The derivations of my test essentially rely on two observations.

The first observation is that, suppose the Euler equation holds internationally, then there exists an international investor who is indifferent between assets in different countries after adjusting for currency and systematic risks. To see that, recall Equation (3.3):

$$m_{i,t+1} - \bar{m}_{t+1} = \Delta q_{i,t+1} \tag{3.3}$$

From it, one can rewrite the Euler equation as:

$$E\left(\underbrace{\exp\left(\bar{m}_{t+1}\right)}_{\substack{\text{Risk}\\ \text{Adjustment}}} \; \underbrace{\exp\left(\Delta q_{i,t+1}\right)}_{\substack{\text{Currency}\\ \text{Adjustment}}} R_{i,t+1} \right) = 1 \tag{3.12}$$

I refer to Equation (3.12) as the global Euler equation. It suggests that the international investor employs $\exp\left(\bar{m}_{t+1}\right)$ as the global pricing kernel to evaluate asset in different countries.

Another way to interpret the same equation is to understand how this international investor compares assets in two different countries:

$$E\left(\underbrace{\exp\left(\bar{m}_{t+1}\right)}_{\substack{\text{Risk}\\ \text{Adjustment}}}\ \underbrace{\exp\left(\Delta q_{i,t+1}\right)}_{\substack{\text{Currency}\\ \text{Adjustment}}}\ R_{i,t+1}\right)$$

$$=E\left(\underbrace{\exp\left(\bar{m}_{t+1}\right)}_{\substack{\text{Risk}\\ \text{Adjustment}}}\ \underbrace{\exp\left(\Delta q_{j,t+1}\right)}_{\substack{\text{Currency}\\ \text{Adjustment}}}\ R_{j,t+1}\right) \tag{3.13}$$

Equation (3.13) can be regarded as a modified version of the uncovered interest-rate parity (UIP). $R_{i,t+1}$ and $R_{j,t+1}$ are asset returns from countries i and j, respectively. $\Delta q_{i,t+1}$ and $\Delta q_{j,t+1}$ reflect the movements in exchange rates. And $\bar{m}_{t+1}$ acts as the risk adjustments. Equation (3.13) states that, in equilibrium, an international investor would be indifferent in investing in the assets of either country after the adjustments of exchange rate and systematic risks. Alternatively, the standard UIP condition might fail, because assets in countries i and j might have different covariances with the global pricing kernel $\exp\left(\bar{m}_{t+1}\right)$ after the currency fluctuations have been taken into account.

The second observation also comes from Equation (3.3). This equation illustrates the close connection between countries log SDFs and exchange rates. Therefore, if there is a factor structure among the log SDFs, then the same factors would also drive the movements of the exchange rates. With that observation, I manage to replace

$\bar{m}_{t+1}$ with a linear combination of the UARs in the absence of the global risk factor:

$$E\left(\underbrace{\exp\left(\sum_{k=1}^{K} C_k^q PC_{k,t}^q\right)}_{\substack{\text{Risk}\\ \text{Adjustment}}} \underbrace{\exp\left(\Delta q_{i,t+1}\right)}_{\substack{\text{Currency}\\ \text{Adjustment}}} R_{i,t+1}\right) = 1$$

3.2.4 Assumptions and Approximations

I review the assumptions and approximations that have been made so far before the empirical section.

In terms of assumptions, I have made three assumptions to make my test possible:

1. The Euler equation holds internationally: $\frac{M_{i,t+1}}{M_{j,t+1}} = \frac{Q_{i,t+1}^j}{Q_{i,t}^j}$.
2. There is a factor structure among countries' log SDFs.
3. The global risk factor $\{f_{0,t}^q\}$ does not exist.

The Euler equation is useful in this context because it establishes the connection between countries' SDFs and the observable exchange rates. The factor structure helps because it reduces dimensionality and leads to over-identification in the test. Last of all, the absence of the global risk factor makes sure that all the factors among the log SDFs also drive the movements of the exchange rates so that they are potentially observable.

As for approximation, I have made the approximation that idiosyncratic risks are canceled out with each other by the Law of Large Number as appears in Equation (3.5). Such an assumption is not worth of concern as the finite sample imperfection of this approximation will be taken into account in the bootstrap procedure introduced in the next section to determine the distribution of the test statistic.

3.3 Empirical Findings

In this section, I show that the moment condition derived in the previous section can be rejected in the full test with exchange rates and equity returns of 28 countries from 1988 to August of 2014.

3.3.1 Data

Spot Exchange Rates

I use spot daily exchange rates to the US dollar from 1988 to August 2014 of the following 27 countries/regions: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Hong Kong, Ireland, Italy, Japan, Korea, Mexico, Netherlands, New Zealand, Norway, Philippines, Portugal, Singapore, Spain, Sweden, Switzerland, Taiwan, Thailand and United Kingdom.[2] Therefore, including the United States, my test assets span 28 countries/regions. The spot daily exchange rates are collected from Datastream. I conduct my test at weekly frequency to maximize the number of periods but avoid noises such as bid/ask spread at daily frequency. I first construct UARs from daily exchange rates then compound them into weekly frequency.

Table 3.1 summarizes the first few moments of weekly UARs of the sample countries during 1988-2014. A large average UAR means that the currency of the country is appreciating on average. Table 3.1 shows that there is considerable heterogeneity of average UAR across different countries. Moreover, the volatility of an UAR is in general much larger than its mean. The noisiness of the exchange rates builds up the challenge to reject the null as the moment conditional would be noisy as well. Figure 3.1 plots the average UARs in decreasing order.

[2]Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Portugal and Spain joined the European Union in 1999 and their currencies merged into the Euro ever since.

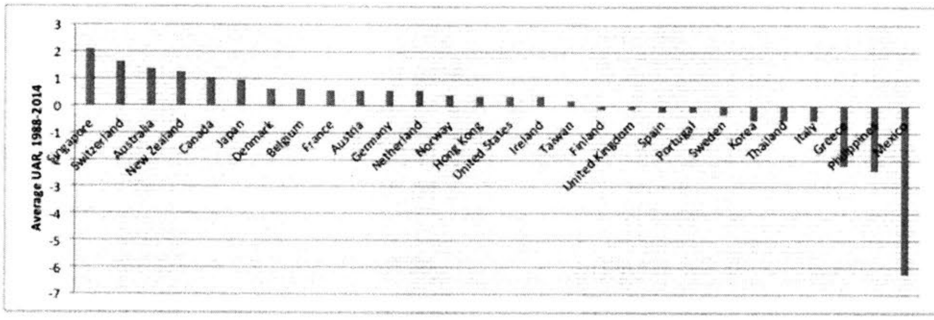
Table 3.1: Moments of UARs, 1988-2014

This table summarizes the first few (annualized) moments of weekly UARs of the sample countries during 1988-2014.

Countries	mean(%)	std(%)	ar1
Australia	1.35	9.81	-0.11
Austria(Euro)	0.57	4.66	-0.85
Belgium(Euro)	0.62	4.46	-0.05
Canada	1.04	7.54	-0.04
Denmark	0.62	4.41	-0.05
Finland(Euro)	-0.10	5.92	-0.08
France(Euro)	0.57	4.43	-0.05
Germany(Euro)	0.57	5.56	-0.04
Greece(Euro)	-2.24	4.79	-0.10
Hong Kong	-0.36	6.53	0.04
Ireland(Euro)	-0.31	4.83	-0.10
Italy	-0.57	5.04	-0.06
Japan	0.94	10.56	-0.06
Korea	-0.05	11.63	-0.05
Mexico	-6.29	12.45	0.04
Netherlands(Euro)	0.57	4.50	-0.04
New Zealand	1.25	9.75	-0.11
Norway	0.42	6.28	-0.10
Philippines	-2.39	8.73	-0.03
Portugal(Euro)	-0.21	4.81	-0.14
Singapore	2.13	5.13	-0.06
Spain(Euro)	-0.21	5.60	-0.14
Sweden	-0.36	6.79	-0.08
Switzerland	1.61	6.90	-0.08
Taiwan	0.21	6.14	0.02
Thailand	-0.52	8.60	0.06
UK	-0.10	6.68	-0.05
US	0.42	6.62	0.04

Figure 3.1: Average UARs, 1988-2014

This figure plots the average (annualized) weekly UARs of the sample countries during 1988-2014. The units are in percentage.



Equity Returns

I compute equity returns from the MSCI aggregate market indices downloaded from Datastream. I compound daily returns into weekly frequency. Table 3.2 reports the mean, standard deviation, and correlation with UAR of countries' equity returns in local currencies. It is interesting to see that UAR and equity returns are in general negatively correlated in the time series for a given country. Therefore, the popular view that a country's currency value responds positively to the health of the country's economy does not have much empirical support at weekly frequency.

Table 3.2: Moments of Equity Returns 1988-2014

This table reports the (annualized) mean, standard deviation, and correlation with UAR of countries' equity returns in local currencies during 1988-2014.

Countries	mean(%)	std(%)	$Corr\,(rs_{i,t}, \Delta q_{i,t})$
Australia	5.77	15.55	-0.35
Austria	5.30	23.46	-0.12
Belgium	6.76	19.27	-0.20
Canada	6.97	15.40	-0.08
Denmark	12.95	21.31	-0.55
Finland	11.02	30.54	-0.17
France	8.16	20.32	-0.27
Germany	8.48	21.60	-0.21
Greece	8.11	32.54	-0.12
Hong Kong	10.97	23.78	-0.21
Ireland	4.94	23.91	-0.24
Italy	4.94	22.85	-0.05
Japan	1.35	23.99	-0.57
Korea	11.08	26.65	-0.04
Mexico	32.76	26.10	-0.26
Netherlands	8.11	19.35	-0.31
New Zealand	1.51	19.65	-0.39
Norway	10.24	23.88	-0.28
Philippines	15.24	25.02	-0.10
Portugal	1.92	20.14	-0.13
Singapore	5.15	19.78	-0.02
Spain	8.01	22.19	-0.17
Sweden	13.99	25.02	-0.31
Switzerland	9.20	22.98	-0.70
Taiwan	9.26	29.94	0.04
Thailand	12.22	31.85	-0.12
UK	7.44	19.26	-0.45
US	9.31	16.48	-0.16

3.3.2 The Factor Structure among UARs

Robust PCA

The technique of principal component analysis has long been popular in the finance literature. As one of the most eminent example, Litterman and Scheinkman (1991) illustrates the success of such a technique when it is applied to interest rates. However, the direct application of PCA on exchange rates turns out to be less ideal than on interest rates because PCA is susceptible to outliers so that large country specific idiosyncratic shocks can substantially blur the factor analysis. The search for PCA-like techniques that are robust to outliers and data corruption has always been an active research topic in the literature of statistics. In a recent breakthrough, Candès et al. (2011) shows that under mild assumptions, it is possible to recover the principal components of a data matrix even though a positive fraction of its entries are arbitrarily corrupted or missing by solving a **Principal Component Pursuit** program that separates the data matrix into a low-rank component and a sparse component.

I relegate the technical details of such an influential new discovery to their paper, yet here I briefly describe how such a technique is implemented on UARs.

Denote Δq as an $N \times T$ matrix with the time series of country i's UAR $\{\Delta q_{i,t}\}$ being its ith column. The principal component pursuit algorithm decomposes Δq into two matrices:

$$[\Delta q_A, \Delta q_E] = PCP(\Delta q, \lambda)$$

where Δq_A is a low-rank component that embodies the principal components of Δq, and Δq_E is a sparse matrix that contains the outliers and corruptions in Δq. λ is a parameter in the PCP algorithm. I choose λ following the recommendation in Candès et al. (2011).

Being able to separate outliers from the UAR matrix, I then conduct the standard PCA on Δq_A to extract factors in the currency market.

The Factor Structure among UARs

Figure 3.2 illustrates the principal components extracted from Δq_A. Consistent with the findings in Lustig et al. (2011), Lustig et al. (2014), and Verdelhan (2011), there a strong factor structure in the currency market. Before the Euro Zone era, the first PC explains as much as 94% of the total variations in Δq_A, while the second and third PC explain 4% and 1%, respectively. Thus, the first three components explain 99% of the total variations in Δq_A. After the establish of the Euro Zone, the first three PCs explain 75%, 17% and 5% of the total variation in Δq_A respectively, with the combined explanatory power as high as 97%. The compositions of the PCs have distinct geographic features and are stable throughout the whole sample period.

Specifically, the first PC can be interpreted as having the European countries on one side, and having US and Asia-Pacific countries on the other side. The second PC can be viewed as a Australia-New Zealand Versus Europe-US factor. The third factor holds Japan and Switzerland on one end, and some of the European countries on the other end.

Figure 3.2: Principal Components of UARs

This figure plots the compositions of the first three principal components extracted from Δq_A before and after the Euro Zone. The compositions of the PCs before the Euro Zone are plotted in Panel a, while the compositions of the PCs after the Euro Zone are plotted in Panel b. Before the Euro Zone, the first principal component explains 94% of the total variations in Δq_A, while the second and third principal components explain 4% and 1% respectively. After the Euro Zone, the three PCs explain 75%, 17% and 5% of the total variations in Δq_A, respectively. The observations in Δq_A are of weekly frequency.

(a) 1988-1998 (b) 1999-2014

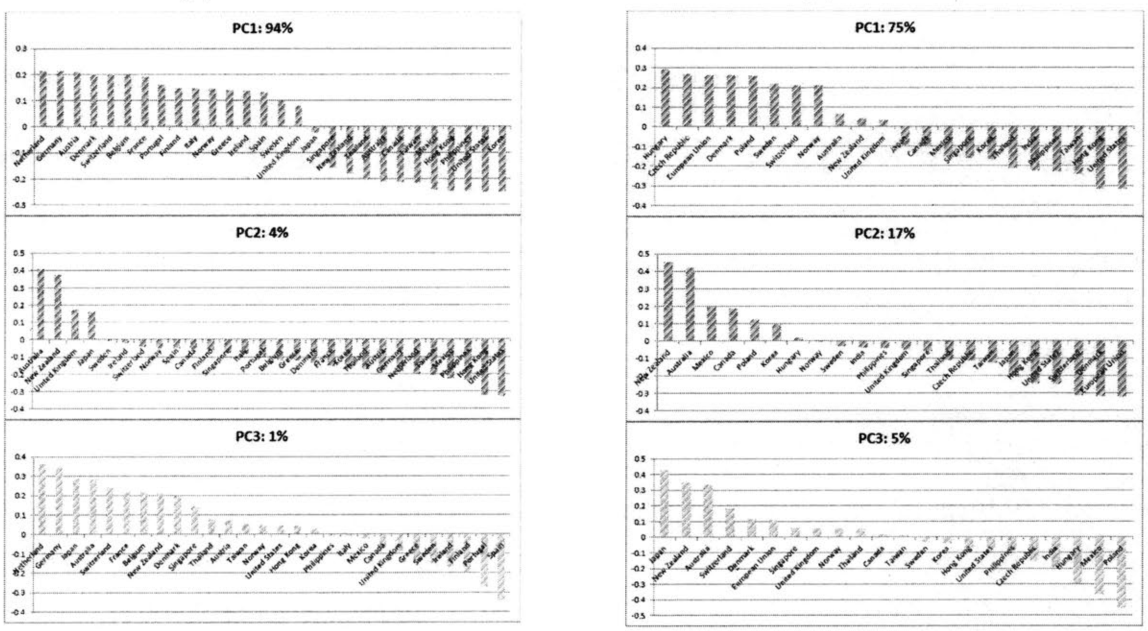


Table 3.3 and Table 3.4 relate the PCs to the findings in Lustig et al. (2011) and Lustig et al. (2014). In particular, Lustig et al. (2011) and Lustig et al. (2014) discovered the Carry and Dollar factors in the cross section of bilateral exchange rates. The Carry factor is defined as a long-short portfolio with positive positions in high interest rate currencies and negative positions in low interest rate currencies. The

Dollar factor is defined as longing the US dollar and shorting the currencies in the rest of the world. Table 3.3 documents the correlations between the principle components and the Carry and Dollar factors. It is obvious that the Carry and Dollar factors are related to the principle components with intuitive correlations. Table 3.4 compares the principle components and the Carry, Dollar factors in the their explanatory power. Not surprisingly, the principle components dominate the Carry and Dollar factors in terms of explanatory power by construction.

Table 3.3: Correlations between PCs and the Carry, Dollar Factors, 1988-2014

This table documents the correlations between the principle components and the Carry and Dollar factors. The Carry factor is defined as longing the high interest rate currencies and shorting the low interest rate currencies. The Dollar factor is defined as longing the US dollar and shorting the currencies in the rest of the world.

Equity Returns\UARs	PC_1	PC_2	PC_3
Carry	-0.28	0.25	0.43
Dollar	-0.74	-0.43	-0.01

Table 3.4: Explanatory Power of PCs and the Carry, Dollar Factors, 1988-2014

This table compares the principle components and the Carry, Dollar factors in the their explanatory power. The explanatory power is defined as the ratio of the explained variance of the total variance in Δq. The Carry factor is defined as longing the high interest rate currencies and shorting the low interest rate currencies. The Dollar factor is defined as longing the US dollar and shorting the currencies in the rest of the world.

	Variance Explained
PC_1	35.9%
PC_2	12.3%
PC_3	12.2%
$PC_1 + PC_2$	48.0%
$PC_1 + PC_2 + PC_3$	55.8%
$Carry$	7.0%
$Dollar$	21.7%
$Carry + Dollar$	29.8%

3.3.3 The Factor Structure among Global Equity Returns

The application of robust PCA on equity returns reveals that there is a significant factor structure among global equity market as well. Figure 3.3 plots the composition of the first three principal components extracted from global equity returns in local currencies. The first PC resembles the market factor in CAPM, and explains as high as 92% of the total variance in global equity returns. The second and third factors are interesting. They both have Asian countries on one side and European countries on the other side. The second factor is influenced by the non-Euro-Zone European countries, whereas the third factor is related to the risks of the Euro-Zone European countries. The both explain about 3% of the total variance in global equity returns. The compositions of the equity PCs are similar to the International CAPM framework proposed in Dumas and Solnik (1995) (also see the summary in Brusa et al. (2014)).

Figure 3.3: Principal Components of Equity Returns, 1988-2014

This figure plots the composition of the first three principal components extracted from global equity returns in local currencies. The first principal component explains 92% of the total variations, while the second and third principal components both explain about 3% of the total variations. The observations in equity returns are of weekly frequency.

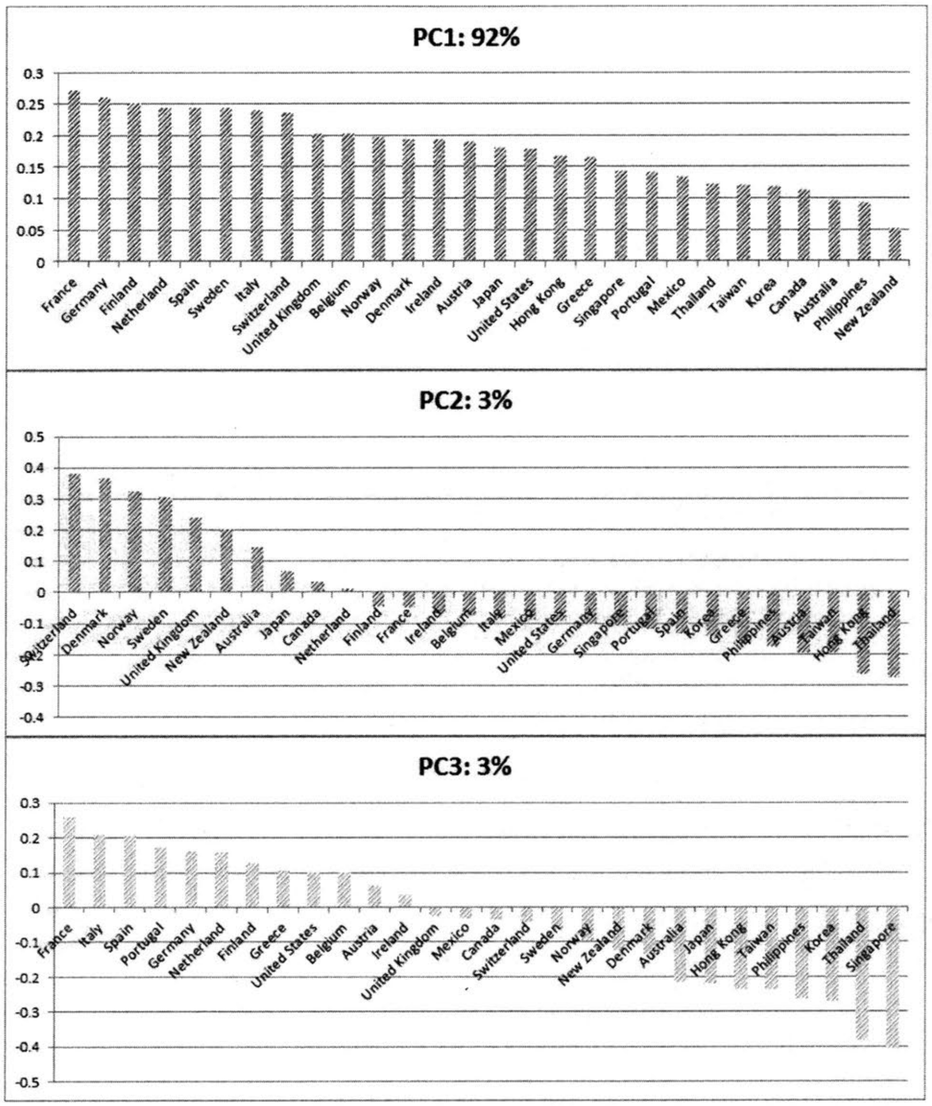


Table 3.5 reports the correlation between the PCs in the global equity returns (in local currencies) and the UARs. This table attempts to capture the degree of connection between the global equity market and the currency market. It is interesting that the second equity principle component is negatively correlated with the first and

second currency principle components. I leave the interpretation of such a finding to future research.

Robustness checks for sub-sample periods are reported in Appendix 3.A

Table 3.5: Correlations between Equity Return PCs and UAR PCs, 1988-2014

This table reports the correlations between the PCs of global equity returns and the PCs of the UARs, with each equity return PC taking one row and each UAR PC taking one column.

Equity Returns\UARs	PC_1	PC_2	PC_3
PC_1	-0.33	0.28	0.21
PC_2	-0.54	-0.42	0.12
PC_3	0.13	0.16	-0.08

3.3.4 Estimation, Testing and Bootstrap

As discussed in detail in Section 3.2.2, Equation (3.11) provides me with an estimable moment condition for the 28 countries in my sample with K parameters $\{C_k^q\}_{k=1}^K$. Contrast to a standard GMM (Hansen (1982)) estimation, the principle components $\{PC_{k,t}^q\}_{k=1}^K$ in the moment condition have to be estimated with robust PCA in the first stage. In addition, before I can estimate the loadings on the PCs $\{C_k^q\}_{k=1}^K$, I need to decide the total number of PCs K to retain. The seminal work in Horn (1965) proposes **Parallel Analysis** as a factor retention criterion in exploratory factor analysis. As a Monte-Carlo based simulation method, parallel analysis proves to be one of the most accurate method in determining the total number of factors in factor analysis, and has been gaining popularity in management studies during recent years (see Hayton et al. (2004)). In order to convincingly reject the null hypothesis, I set the confidence level conservatively at 1% in the parallel analysis and recovered three significant factors in $\{\Delta q_{i,t}\}$, i.e $K = 3$.

Since the number of moment conditions $N = 28$ is significantly larger than the number of parameters $K = 3$, over-identification allows me to test the null hypothesis that Euler equation holds internationally. I conduct the J-test proposed in Sargan (1958), Hansen (1982) with a minor adjustment. Instead of adopting the 2-stage Feasible GMM estimates in the J-test, I simply construct the test statistic with the first-stage GMM estimates under the identity spectral matrix. The standard J-test requires the 2-stage Feasible GMM estimates for the test statistic to converge to χ^2 distribution asymptotically. However, my estimation is non-standard as the components in my moment conditions $\left\{PC^q_{k,t}\right\}_{k=1}^K$ are from estimations rather than directly observable. Therefore, the 2-stage GMM estimation would compound the estimation errors in $\left\{PC^q_{k,t}\right\}_{k=1}^K$ twice, lowering the power of the test as a result. In addition, given the estimation errors in $\left\{PC^q_{k,t}\right\}_{k=1}^K$, the test statistic does not converge to χ^2 distribution anyway. Hence, replacing the 2-stage Feasible GMM estimates with the first-stage GMM estimates is an improvement without cost. To obtain a valid p value of the test, I follow a bootstrap procedure to characterize the asymptotic distribution of the test statistic.

I summarize my bootstrap strategy with the following steps:

1. Assume away global risk factor $\left\{f^q_{0,t}\right\}$ and adopt the unconditional test to estimate the PC amplifiers $\left\{\hat{C}^q_k\right\}_{k=1}^K$ with equity returns according to the moment condition in Equation 3.11.

2. Construct empirical SDFs according to Equation 3.7 and Equation equation (3.10), ignoring global risk factor $\left\{f^q_{0,t}\right\}$, i.e. $\hat{m}_{i,t} = \Delta q_{i,t} + \sum_{k=1}^K \hat{f}^q_{k,t} = \Delta q_{i,t} + \sum_{k=1}^K \hat{C}^q_k \cdot \widehat{PC}^q_{k,t} \iff \hat{M}_{i,t} = \exp\left(\Delta q_{i,t} + \sum_{k=1}^K \hat{C}^q_k \cdot \widehat{PC}^q_{k,t}\right)$.

3. Shift the sample mean of the asset returns so that the unconditional Euler equation holds exactly in sample: $\hat{R}_{i,t} = R_{i,t} + \frac{1-\hat{E}\left(\hat{M}_{i,t}R_{i,t}\right)}{\hat{E}\left(\hat{M}_{i,t}\right)} \Rightarrow \hat{E}\left(\hat{M}_{i,t}\hat{R}_{i,t}\right) = 1$.

4. Conduct the standard bootstrap simulations in the constructed economy with B passes, and record the distribution of the bootstrap-simulated test statistic $\left\{\hat{J}^b\right\}_{b=1}^B$.

My bootstrap strategy possesses the following appealing features:

1. The Euler equation holds exactly: $\hat{m}_{i,t} - \hat{m}_{j,t} = \Delta q_{i,t} - \Delta q_{j,t} = \Delta q^j_{i,t}$.

2. The covariance matrix between SDFs and asset returns is preserved: $\widehat{Cov}\left(\hat{M}_{i,t}, \hat{R}_{j,t}\right) = \widehat{Cov}\left(\hat{M}_{i,t}, R_{j,t}\right)$, $\forall i, j \in \{1, 2, \cdots, N\}$.

3. The covariance matrix between asset returns is preserved: $\widehat{Cov}\left(\hat{R}_{i,t}, \hat{R}_{j,t}\right) = \widehat{Cov}\left(R_{i,t}, R_{j,t}\right)$, $\forall i, j \in \{1, 2, \cdots, N\}$.

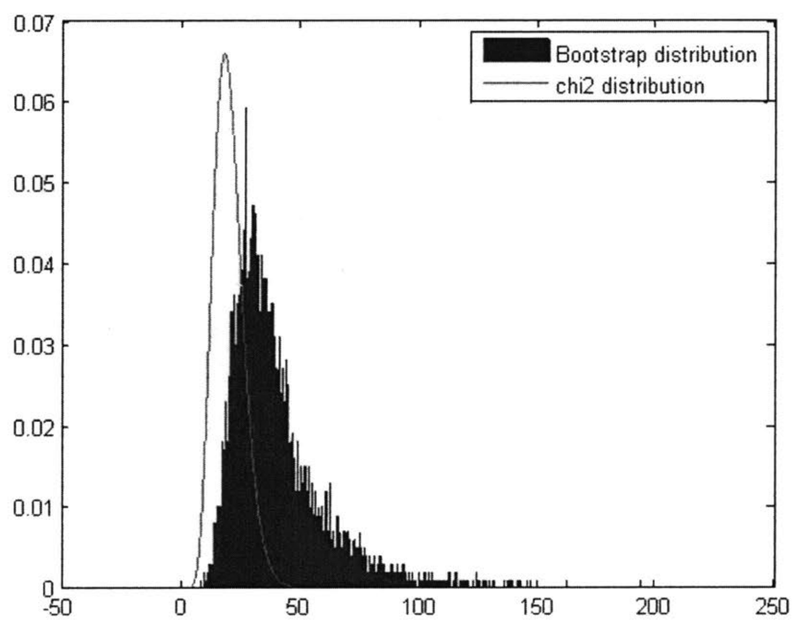
4. Unconditional Euler equation holds by construction: $\hat{E}\left(\hat{M}_{i,t}\hat{R}_{i,t}\right) = 1$, $\forall i \in \{1, 2, \cdots, N\}$.

3.3.5 The Test Result

Figure 3.4 compares the bootstrap distribution of the J-test statistic versus the standard χ^2 distribution for the whole sample. The estimation errors in $\left\{PC_{k,t}^{q}\right\}_{k=1}^{K}$ shift the distribution to the right of the χ^2 distribution and contribute to a larger right tail.

Figure 3.4: Bootstrap Distribution VS χ^2 Distribution, Conditional Test, 1988-2014

This figure compares the bootstrap distribution of the J-test statistic versus the standard χ^2 distribution for the whole sample.



Finally, Table 3.6 summarizes the result of the test. From 1988 to August 2014, the sample contains 1390 observations for each country at weekly frequency. 28 countries are included in the test and 3 PCs are found significant according to parallel analysis. The value of the test statistic is as high as 81.46. With a p value at 2.66%, the null hypothesis that Euler equation holds internationally is rejected at conventional confidence levels.

Table 3.6: Result of the Test, 1988-2014

This table reports the result of the test.

Sample	# of countries	# of factors	p value	# of obs
1988-2014	28	3	2.66%	1390

3.4 Conclusion and Interpretation

In this exercise, I propose a methodology to recover countries' SDFs under three assumptions: 1) the Euler equation holds internationally; 2) there is a factor structure among the exchange rates; 3) the special global risk factor with identical influence on all countries' SDFs does not exist. The intuition of the recovery builds on two observations. First, if Euler equation holds internationally, then there exists an international investor who is indifferent between different countries' assets after taking into account of currency fluctuations and systematic risks. Second, the same factors drive the movements of the SDFs and the exchange rates. Therefore, the pricing kernel employed by the international investor can be represented as a linear combination of the currency factors when the special global risk factor does not exist. Using the exchange rates and equity returns in 28 countries from 1988 to August 2014, I formally test the ability of the exchange-rate-recovered SDFs to price countries' equities. My test takes advantage of the recent development of the robust PCA technique to deal the noises and data imperfection in the exchange rates. I also design a bootstrap strategy to accommodate the non-standard distribution of the test statistic due to the fact that the currency factors are estimated in the first stage rather than being directly observable. I show that the moment conditions are rejected in the data.

The test result indicates the violation of my assumptions. Mechanically, the rejection is caused by the inability of the currency factors to act as risk adjustments to account for the large heterogeneity in the equity risk premia across different countries. Economically, since the factor structure among the exchange rates is directly observed, the rejection of the test suggests that either the Euler equation fails in the international financial market or the global risk factor with equal influence on all countries does exist. If the explanation is that the Euler equation does not hold, then the international financial market is segmented and profitable opportunities are not fully exploited by the arbitrageurs. On the other hand, if the explanation is that the special global risk factor exists, then theories that relate pricing kernels to fundamentals and preferences need to address how such a universal, widespread source of risk could arise. I leave these questions to future research.

References

Bansal, R.: 1997, An exploration of the forward premium puzzle in currency markets, *Review of Financial Studies* **10**(2), 369–403.

Bekaert, G.: 1996, The time variation of risk and return in foreign exchange markets: A general equilibrium perspective, *Review of Financial Studies* **9**(2), 427–470.

Brandt, M. W., Cochrane, J. H. and Santa-Clara, P.: 2006, International risk sharing is better than you think, or exchange rates are too smooth, *Journal of Monetary Economics* **53**(4), 671–698.

Brusa, F., Ramadorai, T. and Verdelhan, A.: 2014, The international capm redux, *Available at SSRN 2462843* .

Candès, E. J., Li, X., Ma, Y. and Wright, J.: 2011, Robust principal component analysis?, *Journal of the ACM (JACM)* **58**(3), 11.

Dumas, B. and Solnik, B.: 1995, The world price of foreign exchange risk, *The journal of finance* **50**(2), 445–479.

Hansen, L. P.: 1982, Large sample properties of generalized method of moments estimators, *Econometrica: Journal of the Econometric Society* pp. 1029–1054.

Hayton, J. C., Allen, D. G. and Scarpello, V.: 2004, Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis, *Organizational research methods* **7**(2), 191–205.

Horn, J. L.: 1965, A rationale and test for the number of factors in factor analysis, *Psychometrika* **30**(2), 179–185.

Litterman, R. B. and Scheinkman, J.: 1991, Common factors affecting bond returns, *The Journal of Fixed Income* **1**(1), 54–61.

Lustig, H., Roussanov, N. and Verdelhan, A.: 2011, Common risk factors in currency markets, *Review of Financial Studies* p. hhr068.

Lustig, H., Roussanov, N. and Verdelhan, A.: 2014, Countercyclical currency risk premia, *Journal of Financial Economics* **111**(3), 527–553.

Mehra, R. and Prescott, E. C.: 1985, The equity premium: A puzzle, *Journal of monetary Economics* **15**(2), 145–161.

Ross, S.: 2015, The recovery theorem, *The Journal of Finance* **70**(2), 615–648.

Sargan, J. D.: 1958, The estimation of economic relationships using instrumental variables, *Econometrica: Journal of the Econometric Society* pp. 393–415.

Verdelhan, A.: 2011, The share of systematic variation in bilateral exchange rates, *manuscript, MIT* .

Appendix

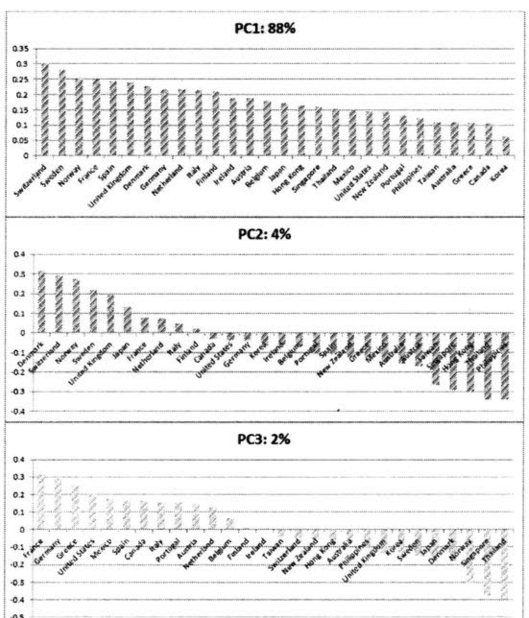
3.A Robustness Checks

3.A.1 Robust PCA of Equity Returns during Sub-sample Periods

Figure 3.5: Principal Components of Equity Returns during Sub-sample Periods

This figure plots the compositions of robust principal components during sub-sample period before and after the establishment of Euro Zone.

(a) Principal Components of Equity Returns, 1999-2003



(b) Principal Components of Equity Returns, 1999-2014