ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
BSc Economics & Business
Bachelor Specialisation Financial Economics

Mutual Funds' Performance During Recessions: Time-varying ability of active fund managers

Evidence from the U.S.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Jan Lemmen for providing continuous support and guidance throughout the writing process of this thesis.

Author: Nikola Draganov
Student number: 412030
Thesis supervisor: Dr. Jan Lemmen
Finish date: July, 2017

ABSTRACT

Despite the extensively documented average inferiority of active mutual funds to passive investment strategies, mutual funds constitute one of the fastest evolving type of financial intermediary in the U.S. managing more than $16 trillion at the end of 2016. Some studies present compelling evidence that fund managers are capable of providing superior returns in times of crises when most desired by investors. If managers are able to provide such type of recession insurance, the fast growth of the sector can be duly justified. Using monthly return data this study analyzes the performance of a large sample of U.S. equity mutual funds over the period 2000 – 2017. Employing a modified conditional version of the Carhart four-factor model, the findings fail to provide corroborating evidence for the existence of time-varying ability of fund managers to realize superior recession performance.

Keywords: Mutual funds, time-varying skill, conditional performance, financial crisis, multi-factor models.

Table of Contents

Chapter 1 Introduction ... 4
Chapter 2 Literature Review ... 7
Section 2.1 Brief History of Mutual Funds ... 7
Section 2.2 Classification of Mutual Funds Types ... 7
Section 2.3 Active versus Passive funds – Active managers and value creation ... 9
Section 2.4 Conditional Performance Evaluation (CPE) ... 18
Chapter 3 Methodology ... 21
Section 3.1 Performance measures ... 21
The Sharpe ratio ... 21
The Capital Asset Pricing Model (CAPM) – Jensen's alpha ... 22
The Fama-French Three-factor Model ... 22
The Carhart Measure ... 23
Interpretation of Alpha ... 23
Section 3.2 Conditional state-dependent regression model of performance ... 24
Chapter 4 Data ... 25
1. Data Issues ... 25
2. Data Selection and Definitions ... 26
Chapter 5 Results and Discussion ... 30
Section 5.1 Unconditional Performance Evaluation ... 30
Section 5.2 Conditional Performance Evaluation – Recession Model ... 34
Chapter 6 Limitations and Suggestions for Future Research ... 39
Chapter 7 Conclusion ... 41
Appendix ... 44
Bibliography ... 45

Chapter 1 Introduction

Over the past two decades mutual funds (MFs) have become the main investment vehicle for retail investors offering them liquidity and diversification at a relatively low cost. The first mutual funds date as back as the 18th century, with the first fund emerging in the Netherlands with the primary aim to provide small investors with diversification. Since then, the scope of the benefits MFs offer to investors has expanded to include, among others, and most prominently – liquidity intermediation, denomination intermediation, cost advantages, and managerial expertise which are largely attributable to the elevated fame of MFs.

Nowadays, in the United States alone, investment companies managed more than $19 trillion – almost half of the global funds' assets at year-end 2016 ($40.4 trillion worldwide), with mutual funds accounting for 85 percent or $16 trillion[1] up from 450 million in 1940. In the past two decades, the number of MFs has steadily risen to 8,066 at year-end 2016 as compared to 6,248 in 1996. US households have continuously and increasingly relied on investment companies, representing the largest group of investors in funds holding 89 percent of total MFs assets, which amounts to 22 percent of household financial assets at year-end 2016 up from only 3 percent back in 1980 (Investment Company Institute [ICI], 2017).

While the strong appeal of MFs is undeniable, it is compelling to understand what drives their increasing demand and to explore the underpinnings of this somewhat rapidly evolving trend. It is their popularity among the retail investors that has spurred a significant interest in the scientific communities. Despite the enormous size of the industry, the vast literature on MFs performance has challenged the ability of fund managers to provide superior performance to investors which is perhaps the highest valued 'benefit' among the other already spelled out above. Treynor and Mazuy (1966) famously asked in their seminal paper on fund performance: "Can Mutual Funds Outguess the Market?", and perhaps even most importantly: Can fund managers persistently provide investors with abnormal returns? These two questions are central in the financial literature. A large body of literature attributes performance to momentum strategies and factor investing, while other studies discovered superior managerial

[1] Mutual fund data exclude mutual funds that invest primarily in other mutual funds.

talent in certain types of funds, which often is usually high enough only to cover expenses, without creating value for investors. A compelling case can be made from the findings that active fund managers have been shown to lack superior ability to outguess the market in a systematic manner as documented in the works of Jensen (1968), Malkiel (1995), Gruber (1996), Carhart (1997), and Fama and French (2010), among others. Yet, in the light of the apparent inferiority to passive investment strategies, investors' flows to active funds have been steadily increasing ever since contrary to rational expectations (Glode, 2011). Nevertheless, this seemingly irrational behavior by investors to commit their capital to such underperforming investments is challenged by the findings of Moskowitz (2000), Kosowski (2006), Staal (2006) and Glode (2011) among others, who rationalize such investing behavior by showing that funds tend to outperform the market during economic downturns when returns are of utmost value to investors.

According to behavior finance theory, risk-averse investors' marginal utility of wealth is a concave function, consistent with the law of diminishing returns indicating that returns are dearest to investors when the level of wealth/consumption is low. This is commonly observed during economic busts when consumption decreases and returns to investors matter the most. Therefore, the core idea of this paper is to investigate whether mutual funds can offer investors a safe haven with positive abnormal returns during harsh economic conditions.

The recent financial crisis of 2008 has been labelled the worst crisis since the Great Depression inflicting serious damage on markets and institutions at the core of the financial system. (International Monetary Fund [IMF], 2008). Since its outburst, the USA has experienced an acute contraction in wealth, a surge in risk spreads and worsening of the credit market conditions spurring lack of confidence in investors (Reinhart & Rogoff, 2008). The recent findings that active funds outperform during economic slowdowns and the harsh conditions that followed the recent financial crisis create a compelling case of whether fund managers managed to sustain their ability to perform abnormally. This line of reasoning inspired the research question of this work, namely:

"Did the US equity mutual funds manage to beat the market during the recent financial crisis of 2008?"

The appeal of this questions stems from the idea that if active funds managed to outguess the market and provide a financial cushion for investors during these harsh economic conditions when mostly desired by investors as corroborated in the recent financial literature, then investment behavior in unconditionally underperforming funds can be duly rationalized on economic grounds. Ideally, my findings will shed light on potentially profitable investment endeavors during financial crises which posits the social relevance of the topic.

This research paper finds its main contribution to the existing financial literature on the topic by expanding the sample period to investigate the ability of active fund managers to realize value to investors during the recent financial crisis of 2008. A comprehensive analysis of fund performance before and after expenses is implemented using four different performance measures.

The remainder of the study is organized as follows. Chapter 2 reviews the existing literature on the topic and presents the empirical findings. Chapter 3 outlines the performance measures and the methodology employed in the study, while Chapter 4 describes the properties of the sample data along with common issues which arise in the selection of mutual funds' data. Chapter 5 presents the results of this paper. Chapter 6 discusses the limitations of the research and gives suggestions for future research. Chapter 7 concludes the study.

Chapter 2 Literature Review

This section presents a review of the major findings on the topic of MFs and their performance as documented in the financial literature. The chapter begins with a brief history of MFs, then introduces the main difference between active and passive funds in Section 2.2. Thereafter, the performance of US Equity MFs is reviewed in Section 2.3, while Section 2.4 complements with conditional time-varying performance measures.

Section 2.1 Brief History of Mutual Funds

While the mutual funds raised in popularity in the past decades their origins date back to the dawn of organized stock trading. The foundations of the first investment trust as a closed-end fund have been laid out in 1774 by the Dutch broker and merchant – Abraham van Ketwich, under the name *Eendragt Maakt Magt* – the maxim of the Dutch: “Unity Creates Strength.” (Rouwenhorst, 2004). Close-end funds were peculiar with that they issued a fixed number of shares at initiation and then were traded between investors on the open market. The first modern mutual fund, Massachusetts Investors Trust, with an open-end capitalization allowing for continuing issue and redemption of shares, was created on March 21, 1924 in the United States of America (Rouwenhorst, 2004). For a more detailed outlook on the history of MFs, the reader is referred to Rouwenhorst (2004).

Section 2.2 Classification of Mutual Funds Types

There are two structures which underpin the organization of a mutual fund in terms of their capitalization. The first funds were organized with a closed-end capitalization, where a fixed number of non-redeemable shares are sold at an initial offering and are then traded between investors over-the-counter. The main problem with this type of funds is that once all shares are sold, the fund cannot raise more capital which has largely contributed to the development of the modern open-end structure, allowing for investment at any time by selling redeemable shares.

MFs are primarily classified into five main categories in terms of the assets which they hold: equity funds, bond funds, hybrid funds, money market funds, and index funds. Equity funds invest in stocks and have different investment objectives. Bond funds' primary assets are bonds, while hybrid funds have a combination of both equity and fixed-income instruments. Money-market funds invest as their name suggests in short-term securities with maturities of less than a year. Lastly, index funds are passive funds, which unlike the other classes mimic the stock composition of an index such as the S&P 500, and hence do not require an active management to oversee and allocate their capital (Mishkin & Eakins, 2012). The Securities and Exchange Commission (SEC) in the USA adds three different categories to the ones mentioned above. First, it classifies funds which are designated to be long-term investments for individuals with a particular retirement date in mind as Target Date Funds, also known as Lifecycle funds. In terms of portfolio composition, they invest in stocks, bonds, and other types of securities and investments. These funds are characterized with specific end-dates, and usually shift risk and manage their exposures towards less risky investments with time. Another type of fund which invests in non-conventional asset classes such as real estates or currencies are known as Alternative Funds. These funds offer an extra layer of diversification since their investments are not closely correlated with traditional investments. Also, sometimes known as "hedge funds for the masses", these funds offer exposures similar to those of hedge funds, employing more sophisticated investment strategies, and as a result, these funds usually charge higher costs. Lastly, a variation of index funds – called Smart-Beta, which involves alternative index construction scheme to traditional indexes based on market capitalization. While aiming to track particular passive index, these funds employ alternative weighting practices such as volatility, the rules for which are transparent to the investors. The goal of this fund type is to provide higher than passive performance at a lower cost than actively managed funds.

Morningstar® categorizes equity mutual funds into three broad categories – value, growth, and blend. It further applies a second layer of categorization on the size level, that is whether the main stock holdings are in companies with large, middle or small market capitalizations. This classification will be employed for the purpose of portfolio formation and throughout the text.

Section 2.3 Active versus Passive funds – Active managers and value creation

There stands a large body of literature presenting similar results on active fund performance, however, for the purpose of brevity I will review the main findings of several acclaimed articles on the topic in an attempt to also address evidence from the most recent works.

The topic of value creation is central in the context of active mutual funds, with the rationale that they "add value" relying on manager skill and using private information to achieve abnormal performance. Such performance of course comes at a cost, as French (2008) adds up the expenses during 1980-2006 in the US, it turns out this luxury cost investors about 2/3 of a percent per year. Aragon and Ferson (2007) highlight the important difference between evaluating funds using gross and net returns. If funds can achieve superior performance even after costs, as the authors suggest this constitutes the value creation process by fund managers, which is to be contrasted to the scenario when superior before-costs performance is only achieved but expensed in the form of management fees or washed away through trading costs. In short, gross outperformance signals managers' ability to select portfolios which outperform their passive benchmarks, while net outperformance signals whether their skill rewards investors as well.

The performance of active managers is broadly assessed along two dimensions – stock picking talent, also known as selectivity, and market timing abilities (Fama, 1972). Stock-picking talent (selectivity) is typically measured by the Jensen's alpha (1968), which indicates the degree to which a fund manager can perform better than the passive benchmark which he has committed to beat. In other words, alpha indicates the ability of the manager to pick stocks in such a manner that provides superior return. Fama (1972) refers to selectivity as the ability of fund managers to deviate deliberately from an index and choose the best securities for a given level of risk in an attempt to outperform the market index. This paper's main focus is along this dimension, and for that purpose only a brief overview of some of the highly acclaimed literature on timing abilities will be presented. Market timing is another skill which has attracted much academic attention. The term "market timing" has two distinctive interpretations in the financial literature. The classical understanding of the term refers to the ability of the manager to adjust his exposure to market risk conditional on his expectations about the state of the market, that is, market exposure goes up before bull markets settle in, while it decreases in the face of a bear market (Fama, 1972). The second use of the term is notoriously connected to the mutual fund scandals in the early 2000s, which exemplifies the attempts of fund investors to engage in trading on stale prices. In the discussion of market-timing here, I stick to the traditional notion of the term. In the early studies, the academics fail to document successful strategies at outguessing the market and conclude that any efforts directed to such endeavours would turn fruitless (Ferson, 2012). One of the first studies to investigate managerial talent decomposed into market-timing abilities is found in the work of Treynor and Mazuy (1966). Using annual data on 57 open-end mutual funds, they devised a statistical test to examine the historical success of funds to anticipate major movements in the equity market, and concluded that managers were unable to significantly "outguess" the market during the period of 1953-1962. In a more comprehensive study, Daniel, Grinblatt, Titman, and Wermers (1997) develop two new performance measures which employ custom benchmarks specifically constructed to match the characteristics of the stocks held by the funds under investigation. Analysing more than 2500 mutual funds over the period 1975 – 1994, their results reaffirm the previous findings on market timing, and instead they find some stock-picking talent among aggressive-growth funds. Wermers (2000) elaborates of the methodological part by decomposing stock returns into selectivity, style timing, transaction costs and expenses. Similarly, he finds supporting evidence for the superior selectivity skills of managers, with lack of support for timing ability. A significant exception is documented in the work of Bollen and Busse (2001) who find significant evidence for positive market timing when analysing daily data. Previous studies mainly focus on exploring timing abilities as compared to the wide equity market, while Swinkels and Tjong-A-Tjoe (2007) argue that mutual funds might enhance their performance by adjusting their exposure to certain investment styles – size, value and momentum. The authors suggest that managers who are able to guess these factors could earn economically significant profits. Employing the daily observations on 153 mutual funds, they find significant empirical evidence that fund managers possess market-timing abilities. Their results on factor timing, indicate the ability of fund managers to guess the direction of the value and momentum factor, though not their magnitude, while no evidence is found for size-factor timing skill.

The Efficient Market Hypothesis (EMH) developed by Fama (1970) predicates that all available information is reflected in stock prices signifying a state of market efficiency. Within this framework, it can be inferred that mutual funds dedicating resources to research in an attempt to capitalize on mispriced assets would be fruitless and only result in inferior returns marked by high expenditures. In other words, in "fully-efficient" markets the possibility to beat the market was negligible and solely marked by chance. In effect, every time researches analyse the performance of mutual funds they essentially test the EMH and its implications for the structure of the market.

The relative attractiveness of active funds is judged along two distinct dimensions as suggested by Jansen (1968), namely i) *the ability of the active manager to boost performance by a correct forecast of future security prices*, and ii) *the ability to minimize the degree of insurable risk through the process of diversification*. In other words, fund managers need to be able to correctly identify under- or overpriced securities, and capitalize on such market discrepancies while on the same time achieving adequate risk hedging by combining stocks in such a manner which drives down idiosyncratic risk to a minimum. The scientific community has expressed its scepticism towards the successful fulfilment of these two criteria by active fund managers.

One of the first contributions to MFs performance is found in the earlier works of Sharpe (1966), where he studies the performance of 34 open-end mutual funds during the period 1954-1963. In his seminal paper, Sharpe reviews the recent developments in the behaviour of stock prices reviewing the theory of random walks which asserts that past stock prices do not qualify as a predictor for future prices, suggesting the cumbersome and expensive task involved in screening the market for mispriced securities. If this line of reasoning is to be followed, the author expresses his scepticism regarding the viability of managers dedicating large amounts of resources in a search of incorrectly priced securities. Advancing on the work of Tobin (1958), the author introduced the "Reward-to-Variability" (R/V) ratio, which subsequently became famously known as the Sharpe ratio, in order to evaluate mutual funds' performance – measuring the reward per unit of risk. As a result, he introduced the notion of the "efficient portfolio" – a portfolio which provides the highest expected return for a given risk level. He further bolstered his results by employing several other purported measures of performance – the Treynor Index (TI), expense ratios and fund size. Fund size *per se* was shown to lack predictive power for future performance. Despite obtaining somewhat good predictability of returns by the R/V ratio, measured by the correlation coefficient and regression analysis, the results were not statistically significant. In contrast, the TI turned out to be the better predictor of performance, showing a higher degree of correlation implying that funds with high(low) TI tend to maintain its high(low) performance. However, after accounting for expenses, Sharpe (1966) conjectured that the only foundation for persistently inferior performance would be associated with the proclivity of funds to commit large amounts of their assets to the relatively fruitless search for mispriced stocks, implying market efficiency and the superiority of diversification.

The sole reliance on relative performance measures such as the Sharpe ratio, and the somewhat inconclusive evidence accompanied by the limits of the relative nature of the performance measures inspired the introduction of an absolute performance measure by Jensen (1968). The author held that the use of relative performance measures was limiting because it only allowed for one dimensional comparisons of the sort "portfolio A is superior/inferior to portfolio B". Instead, what was needed was not only knowing which portfolio is relatively better, but also whether these portfolios are "good or bad compared to some absolute standard". This led to the development of the acclaimed Jensen's alpha which measures the fund performance relative to a passive benchmark portfolio. Jensen (1968) investigated a sample of 115 open-end funds during the period 1955-1964, and tried to find as much data as he could for these funds again during the ten-year period 1946-1954 but complete data was only available for 56 of them. During the later period, he found that that the average mutual fund held less risky portfolios than the market(passive) index. He calculated the alpha for each individual fund using both net and gross returns. Not only did the results show that on average funds underperformed the market by 1.1% but also that neither fund managed to perform better than the market individually regardless of expenses. Akin to the results of Sharpe (1966), Jensen (1968) conjectured from the prevalence of negative alphas that the average fund was unable to engage in a reliable prediction of future security prices sufficiently enough to recoup their total expenditures. Since their inception, the Sharpe ratio and the Jensen's alpha became the cornerstone of performance measures and underpinned the results of many subsequent works.

As a continuation of the period investigated by Jensen (1968), Ippolito (1989) investigated a sample of 146 mutual funds covering the 20-year period from 1965 to 1984 and obtained contrasting results. His sampling method addressed the issue of survivorship bias by including funds that did not survive the entire period. His work was in many aspects similar to Jensen's (1968) and Sharpe (1966) but its highlight was the investigation of the relationship between portfolio turnover and performance. Portfolio turnover refers to the frequency of buying and selling securities by the fund manager or in other words, it portrays how 'active' a manager is. The results showed little evidence that turnover and expenses are associated with inferior performance. Following the Jensen's (1968) CAPM procedure, he obtained that the average mutual fund outperformed the market by 81 basis points after expenses. In addition, he showed that the significantly positive alpha is robust across fund types and the choice of market benchmark. However, the author conjectured that the positive alphas were not sufficiently high so as to compensate investors for the load charges incurred. While load fees were not taken into account by earlier studies such as Sharpe (1966) and Jensen (1968) in estimating performance, the results obtained by Ippolito (1989) give the opposite impression providing a compelling starting point for future research. However, Ippolito's results were shown to be particularly sensitive to the choice of the benchmark. Using his data and a benchmark index on non-S&P equities, Elton, Gruber, Das, and Hlavka (1993) showed that the positive after expense returns vanished. Employing a different approach, Grinblatt & Titman (1992) find similar evidence supporting the ability of fund managers to provide superior and persistent performance. However, Malkiel (1995) called into question the validity of their results by pointing out that the authors' sampling procedure suffered from survivorship bias. According to Malkiel (1995), Grinblatt & Titman (1992) significantly overestimated the returns received by investors by claiming that the effect of survivorship bias was relatively small. On the other hand, the *hypothetical* estimation of gross returns was also seen as problematic by Malkiel (1995) who claimed that the authors did not calculate actually achieved returns because they were computed stock by stock instead for the actual portfolio assuming out portfolio rebalancing throughout the period. Following the CAPM framework, Malkiel (1995) analyzed the performance of all mutual funds in existence during the 20-year period of 1982-1992 and showed the severe distortions survivorship bias may introduce in the performance evaluation. Analyzing the returns of all funds, the author bolstered the widely-held notion that fund managers were unable to provide excess returns even before expenses noting the relative superiority of passive indexes marked by market efficiency. Akin to Fama and French (1992), Malkiel (1995) could not find significant evidence to support the risk-return relationship commanded by the CAPM noting the deficiency of the model and raising the appeal for a more rigorous means of performance evaluation. In their study, Fama and French (1992) found out that the risk-return relationship posited by the CAPM did not hold over their period of investigation 1962-1989, which contributed to the development of the Fama-French Three-Factor Model (1992, 1993) by augmenting the specification of the CAPM with two new risk factors. The authors showed that the equity risk is multidimensional and stock returns can be better explained by two other factors, a size, and a value factors. Parallel to the developments by Fama and French, Jegadeesh and Titman (1993) obtained a momentum strategy of buying past year winning stocks and selling past year losing stocks rendering investors significant positive returns. This momentum anomaly was then described by Chan, Jegadeesh, and Lakonishok (1996) as a market inefficiency occurring due to a market-wide underreaction to information. Nevertheless, it produced robust effects across time (Jegadeesh & Titman, 1993) and countries (Asness, Liew & Stevens, 1997). In his seminal paper, Carhart (1997) added the momentum factor to the Fama-French (1993) three-factor model to arrive at the Carhart four-factor model. Making use of his model, Carhart employed a survivor bias free dataset, analyzing the performance of 1896 diversified equity funds in the period 1962-1993 grouping them in portfolios by investment objective. The author found that the 4-factor model significantly improved the average pricing errors by the 3-factor model and its predecessor the CAPM, signifying its good fit in explaining cross sectional variation in average stock returns. On performance, Carhart reports that only the top-decile funds managed to earn back their investment costs, while most mutual funds underperform by about the size of their costs, with bottom-decile funds underperforming on even larger scale by about twice their expenses. Furthermore, the author finds a significant negative correlation between performance and expense ratios, portfolio return, and load fees. A preceding study by Gruber (1996) employing a different 4-factor model, finds supporting evidence that mutual funds underperformed their benchmarks by approximately 65 basis points. With the average expense ratio being 113 basis points per year, Gruber suggests the ability of funds to provide superior returns, but that managers charge more than the magnitude of the created value, resulting in the well documented inferior performance of active funds. Wermers (2000) conducts a comprehensive analysis on the mutual fund industry, including data on 1788 funds over the period 1975 – 1994. His results support the average net underperformance of funds by 1% on average. However, he takes a different approach and decomposes the returns into stock-picking, transaction costs and expenses. He reports that the average manager achieves a 1.3 percent gross outperformance, which results in a 2.3 percent gap with the net underperformance. Of this gap, he attributes 0.7 to underperforming non-stock holdings such as cash and bonds, which usually funds keep aside in the unfortunate event of a liquidity crisis triggered by unexpected high-scale stock redemption by investors – signifying the high hedging costs associated with liquidity problems. The other 1.6 percent is attributable to transaction costs and management fees. Finally, the author found a positive relationship between stock selectivity and turnover, used to measure the activeness of the manager. He reports that despite of the substantially higher total expenses of high-turnover funds, their choice of stocks yields much higher returns than those of low-turnover funds. A more recent study by Fama and French (2010) documents the performance of mutual funds using a bootstrap method in the period 1984-2006 using the CRSP survivor bias free database. Their results show that the average fund manager cannot provide superior returns to investors as analyzed by CAPM, 3-factor, and 4-factor benchmarks. They estimate the magnitude of their underperformance at about the costs in expense ratios. Following the bootstrap simulation, they strive to distinguish luck from skill in fund performance. Only a few funds surface with ability to cover their costs, while the estimates of a four-factor true *alpha* are significantly negative. Therefore, if a number of managers are capable of covering their costs, they are obscured in the aggregate by the mass of managers with inferior skill.

The majority of empirical findings on active fund performance is well summarized in the work of Cuthbertson, Nitzsche, and O'Sullivan (2010). Their rigorous and extensive review indicates that ex-post there are around 0-5% of the top performing US and UK equity funds with truly positive *alpha*-performance (net of fees), and around 20% of funds marked by truly negative *alpha*-performance, while the rest 75% of active managers earn on average just to cover their expenses (zero *alpha*-performance). They report that key predictors of relative performance are load fees, expenses and turnover. In general, they conclude that only sophisticated investors should pursue active investment strategies while sensible for the most investors would be to stick to low cost index funds and avoid holding active funds with poor past performance.

The need for new approaches has arisen to better address the growing importance of active funds in spite of the conflicting results of their average underperformance. As it appears from the literature review so far, outperformance clearly cannot be ruled out, the sheer fact that in the average, returns are smoothed out especially after accounting for expenses, does not mean that superior performance does not exist altogether. It is better to tilt the analysis more into the underlying characteristics peculiar to active managers capable of providing excess returns. A good starting point would be to understand what distinguishes capable active managers from the average. A recent study by Cremers and Petajisto (2009) took such a direction and introduced a new measure of active portfolio management called "Active Share"

(AS) which measures the degree to which a fund's portfolio is differentiated from its respective benchmark. The authors gauged the activeness of US mutual funds from 1980 to 2003 and related it to size, expenses, and turnover. They found out that the new measure fits its role as performance predictor well, reporting that the most active managers significantly outperformed their benchmarks both before and after costs persistently, while the opposite held true for the least active funds. In his later work, Petajisto (2013) elaborates on his previous work and introduces a two-dimensional framework for gauging the activeness of managers. He introduces the complementary measure as "tracking error" (TE) which measures the excess risk over the benchmark portfolio, and argues that these two measures underlie different aspects of activeness, namely stock selection as proxied by AS, and systematic factor risk as proxied by TE. He then categorizes equity funds into four classes depending on their score on the two measures and examines performance. Looking at gross performance, Petajisto finds that the average fund managed to provide superior performance by 96 basis points, and 31 basis points when analyzed with the Carhart (1997) 4-factor model. The author concluded that only funds with high TE did not manage to outperform their benchmarks even before costs, while net of costs, only funds with the highest AS provided a statistically and economically significant outperformance, rendering them the only class among funds capable of creating value to investors. Next to "AS" as performance predictor, expenses and fund age showed significant negative relationship with performance, indicating that higher expense ratios signal poor fund quality, while older funds exhibit slight underperformance.

Kacperczyk and Zheng (2006) took a different approach by examining the "unobserved actions" by mutual funds. They argued that investors do not see the full picture behind funds management despite the rigorous aims of regulatory institutions regarding transparent disclosure. The authors estimated the scale of the asymmetrical information using the "return gap" – a measure which captures the difference between reported fund return and the return on a portfolio that invests in the previously disclosed fund holdings. The authors analysed more than 2500 unique US equity funds in the period 1984-2003. Despite observing an average return gap very close to zero, their main results showed a significant positive relationship between past return gaps and performance, indicating that funds with high past return gaps tend to show persistence in providing excess returns both before and after differences in risk and styles. Following similar logic as Cremers and Petajisto (2009), Amihud and Goyenko (2013) take a different track to gauge the activeness or "stock selectivity" of mutual funds. They suggest that fund performance can be predicted by its R^2 resulting from a regression of its returns on a multifactor benchmark model. R^2 is a statistical metric, which measures the variation in fund returns explained by the variation in the different factors. In other words, it can serve as a metric to gauge the joint exposure funds have to these factors, which indicates the differentiability of the fund portfolio or in other words, as the authors propose - the selectivity of the fund manager measured by $1 - R^2$, indicating the residual or idiosyncratic variation in fund returns. Thus, lower R^2 signals greater activity or selectivity in the fund's investment. The authors find that R^2 is a significant predictor of performance, with a negative relationship with *alpha*. Their results hold even when controlling for fund characteristics and are robust to different benchmark factor models. Amihud and Goyenko (2013) show that investing in the portfolio with the lowest R^2 and highest *alpha* yields a significant *alpha* of 3.8% or higher in the following period.

Somewhat contrary to the vast majority of empirical findings on fund performance, a recent study by Berk and van Binsbergen (2015) employs a different approach to studying performance and shows supporting evidence for the existence of managerial skill. The authors ascertain that many prior studies have utilized net alpha to measure managerial talent, whereas other have acknowledged the gross *alpha* as a better indicator. However, Berk and van Binsbergen claim that the *alpha* on its own is not an appropriate measure because competition between rational investors drivers net *alphas* to zero in equilibrium as suggested by the model of Berk and Green (2004). They argue that *alpha* is a return measure, not a value measure and its interpretation should not be carried out on its own but relative to the size of the fund because 1% abnormal return on $1-million fund is far different from that on a 10$-million fund from an economic perspective. Therefore, authors measure value by examining how much money managers can extract from the financial markets. A sample period covering 1962 to 2011, shows that the average fund is able to create value by extracting on average $3.2 million per year from financial markets. However, caution must be exercised when interpreting these findings as here the process of value creation refers to managerial expertise. Despite the fact that value created is not passed on to investors as extensively documented, they believe that managerial talent is better measured on such a base. Lastly, the authors emphasize the surprising results that investors appear to identify managerial skill and compensate it. In other words, as evident from their results, current expenses serve as predictor for future performance.

Section 2.4 Conditional Performance Evaluation (CPE)

The review of the empirical findings until now employs the traditional approach to estimating unconditional alphas which assumes constant variation in average returns and risk measured over a given period of evaluation. In other words, alphas are estimated without taking into consideration variations in the state of the financial markets or the general economy. An immediate weakness of this methodology surfaces when considering evaluation consistent with business cycles, whose regularity and influence on the financial markets and consequently the behaviour of the market agents are well documented. For instance, if a study investigates performance during economic downturns, but extrapolates its results in the future where market is likely to take on upwards, the external validity of the results becomes questionable (Ferson, 2012). The problems with unconditional alpha evaluations have long been recognized in the financial literature (Jensen, 1972), (Grant, 1977).

In the framework of the Conditional Performance Evaluation (CPE), risk exposures and the related market premiums are allowed to vary with the state of the economy (Ferson, 2012). It is documented by Ferson and Schadt (1996) in their empirical study, that unconditional alphas differ significantly from conditional alphas for equity style funds implying a time-varying ability of fund managers. As a result, being able to distinguish properly between the states of the economy increases the precision of models to accurately detect abnormal performance insofar as performance varies across business cycles.

The literature reviewed in the previous section extensively documents the lack of ability for active manager to provide superior returns to their investors. However, as already posited by the CPE these studies might not paint the entire picture, since they employ the unconditional *alpha*-performance methodology. Studies investigating the time-varying ability of fund managers appear to build up a different rationale for performance evaluation when properly accounting for time-variations in managers' ability. Moskowitz (2000) conjectures that managers might create value for investors by performing better during recessions. He shows that over the 1975-1994 period, the average alpha was 1% higher during recessions. Ferson and Qian (2004) carry out an empirical investigation over the conditional performance of US mutual funds in the period 1973-2000, and find evidence for time-varying ability in the CPE framework when ability is allowed to vary over economic ups and downs. Lynch, Wachter, and Boudry (2002) develop a new methodology using an *ex-ante* available information to model conditional performance. Using dividend yield as an indicator of business cycles, they find that all types except growth, perform better during recessions than in expansions. Staal (2006) studies the conditional performance of US equity mutual funds during the period 1962-2002, and finds evidence of a negative relationship between *alpha*-performance and the Chicago Fed National Activity (CFNA) Index, an indicator for the state of the economy. As a result, low levels of CFNA index were associated with better performance, a sign of time-varying ability. Kosowski (2006) implements one of the most comprehensive studies to date employing Markov regime-switching model to investigate conditional fund performance in economic expansions and contractions during the period 1962 – 2005 as documented by the National Bureau of Economic Research (NBER). Alpha-performance measures were computed using the CAPM, 3-factor Fama-French (1993), and 4-factor Carhart (1997). He argues that the previously documented empirical findings on poor unconditional performance is concentrated in economic upturns, while performance conditional on recessionary periods is positive. The difference in the risk-adjusted return measured by *alpha* between economic busts and booms is statistically and economically significant with a magnitude of 3-5% per year. The finding of a higher performance during recessions is shown to be robust to model specification and conditioning on turnover, front load, expenses and percentage of stock holdings. Kosowski's (2006) findings indicate that traditional approaches of estimating fund performance underestimate the true scale of value added during recessions and mark the importance of time-varying ability to performance evaluation. Another study by Kacperczyk, Van Niewerburgh, and Veldkamp (2014) employs a holdings-base model to estimate conditional performance. They exhibit the relationship between different components of performance and an indicator of economic expansions and contractions. Their results show a stronger concentration of stock-picking (*alpha*) ability in booms, while a stronger timing-ability during busts. Glode (2011) develops a simple theoretical elaboration on the Berk and Green (2004) model, where fund managers concentrate their skill during recessions. The intuition from the model stems from the fact that an active manager would be better-off if he focuses his ability on such high-value states, when investors are willing to pay more for it. The author further calibrates the model to the US economy and reproduces quantitatively the underperformance using data on 3,147 actively manages U.S. funds, finding consistence with the results of his theoretical model predictions. He constructs a binary variable which indicates the NBER recession dates, and interacts it with the factors of the Carhart (1997) model. The scholar finds that on average equity mutual funds with the worst (best) four-factor unconditional alphas reverse to positive (negative) alphas in recessions. A subsequent study by Badrinath and Gubellini (2012) , elaborates on his work by employing a conditional CAPM that jointly captures time variation in abnormal performance and risk premiums. They accentuate the use of *ex-ante* available to investors data to model the different economic states. The academics voice their concern that Glode's (2011) work does not provide information whether the funds in his sample shared any common characteristics. Instead, they group funds based on investment objective, and find that only small- and mid-cap growth equity funds exhibit negative performance during expansions and significant positive performance during contractions, while value funds do not.

Another study by De Souza and Lynch (2012) calls these results into question by criticizing the methodological aspects of previous studies. They argue that some of the above studies employ conditional information which is only available to investors *ex-post*, such as the NBER recession dates in Kosowski's (2006) and Glode's (2011) works. Instead, they develop a model which conditions only on publicly available information and fail to find supporting evidence of time-varying performance over the business cycle. They show that not all fund styles perform better during recessionary periods conditioning on *ex-ante* publicly available instruments such as dividend yield and term-spread. These contrasting results serve as a compelling case and a foundation for future research into time-varying ability.

The thought-provoking empirical evidence and theoretical models reviewed so far inspire the following hypotheses through which the main research question will be investigated.

H01*: U.S. Equity Mutual Funds can unconditionally outperform their passive benchmarks after expenses during the period 2000-2017.*

H02*: There is a significant time-varying ability of fund managers to provide superior results over times of crisis.*

The hypotheses will be tested with 4 acclaimed performance measures in the financial literature in order to check the robustness of the results to different model specifications. The next chapter discusses the models in detail.

Chapter 3 Methodology

Section 3.1 Performance measures

This section presents the performance measures underpinning the investigation of **H01**. Section 3.2 shows the modified version of the Carhart (1997) four-factor model used to analyse **H02**.

The Sharpe ratio

The Sharpe Ratio (SR) is perhaps the simplest risk-adjusted measure of historical performance, developed by Sharpe (1966) to analyse mutual fund performance, and nowadays frequently used by practitioners for the evaluation of hedge funds and other investment products (Ferson, 2012). It measures the return an investor earns in excess of the risk-free investment per unit of risk. In other words, it shows the compensation an investor receives for bearing risk. For a given portfolio p, it is defined as:

$$SRp = (Rp - Rf\ )/\sigma(Rp\ )\ , \qquad (1)$$

where Rp is the return on the portfolio, and Rf is the return to the safe investment, and $\sigma(Rp\ )$ is the sample standard deviation of the portfolio's excess returns. The higher the Sharpe ratio, the better the fund's historical risk-adjusted performance. Since this paper employs monthly returns, the above formula produces the monthly SRs. In the tables and throughout the text, annualized SRs will be reported. The calculation and annualization of the SRs follows the Morningstar procedure and multiplies the monthly SRs by the square root of twelve, which is, in essence, equivalent to converting both the numerator and the denominator in annual terms by multiplying the average monthly excess returns by twelve, and the sample standard deviation by the square root of twelve.

The Sharpe ratio needs to be applied with caution when the distribution of the returns is *non*-normal. Since the standard deviation is sensitive to outliers and non-normal distributions, it can produce highly distorted results which can seriously impinge on the relative interpretation of the measure.

The Capital Asset Pricing Model (CAPM) – Jensen's alpha

In his seminal paper, Jensen (1968) employed the CAPM, developed independently by Sharpe (1964) and Lintner (1965), to study fund performance and his work contributed to making *alpha* the most widely used performance measure among practitioners and academics (Ferson, 2012). Jensen's alpha is derived from the CAPM as follows:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i\left(R_{MKT,t} - R_{f,t}\right) + \epsilon_{i,t}\ , \tag{2}$$

where $R_i - R_f$ stands for the fund return in excess of the risk-free rate at time t, $R_{MKT} - Rf$ is the return on the market index in excess of the risk-free rate, β_i measures the systematic risk that the fund takes towards the market, and ϵ_i is the residual return of the fund at time t. In the CAPM framework, a fund can only increase its expected return if it increases its exposure to the market (β_i). *Alpha* (α_i) is the intercept from Equation (2) obtained from a time-series regression of $(R_i - R_f)$ on $\left(R_{MKT,t} - R_{f,t}\right)$.

The Fama-French Three-factor Model

Fama and French (1992, 1998) along with Lakonishok, Shleifer and Vishny (1994), among others, discovered that value stocks perform significantly better than growth stocks, resulting in a value premium. Value stocks are defined as stocks possessing high book-to-market ratios, whereas growth stocks have low book-to-market ratios. Additionally, Fama and French (1992) accentuated on another phenomenon where small stocks tended to exhibit higher average returns than big stocks, giving rise to a size premium. Small stocks are defined as stocks with low market capitalization, while big stocks have high market capitalization. These market anomalies led to the development of the value and size investment strategies. The value strategy entails buying value stocks and selling growth stocks, while the size tactic involves purchasing small stocks and selling big stocks. These phenomena along with the findings of Fama and French (1992) regarding the poor predictability of the CAPM beta led to the development of the Fama-French three-factor model (FF3F), defined as follows:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}\left(R_{MKT,t} - R_{f,t}\right) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \epsilon_{i,t}\ , \tag{3}$$

where SMB_t (Small Minus Big) is the return difference between low and high market capitalization stocks, and HML_t (High Minus Low) is the return difference between high and low book-to-market stocks, and β_i measures the exposure to the corresponding risk factor.

The Carhart Measure

The Carhart four-factor model (C4F) is an extension of the Fama-French 3-factor model with one additional factor which captures the one-year momentum anomaly documented by Jegadeesh and Titman (1993). The momentum strategy involves buying past-year winning stocks (stocks with high prior-year return) and selling past-year losing stocks (stocks with low prior-year return). The model is defined as:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_{1,i}\left(R_{MKT,t} - R_{f,t}\right) + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}PR1YR_t + \epsilon_{i,t}\ , \qquad (4)$$

where SMB_t, HML_t, and $PR1YR_t$ are the returns on value-weighted, zero-investment, factor-mimicking portfolios for size, value, and one-year momentum in stock returns or alternatively, the return difference between high prior-year return and low prior-year return stocks. The momentum factor is also commonly known as WML (Winners Minus Losers), and just MOM (momentum). In this study, I employ the former use of the term throughout the analysis and the tables presenting the results. The intercept of the model, α_i, is the Carhart measure of abnormal performance. Carhart (1997) leaves risk interpretations to the reader, and uses the model to explain returns.

Interpretation of *Alpha*

The alpha from Equations (2), (3), and (4) is the average return left unexplained by the respective equation factors (benchmarks). In the CAPM framework, a positive α_i may signal the ability of the manager to construct such a portfolio which on average outperforms the market return by the magnitude of its coefficient. A positive Carhart α_i indicates that the fund manager’s portfolio outperformed a passive benchmark constructed on the market, size, value, and momentum factors. In common practice, α_i represents the ability of the manager to find mispriced securities (Pastor & Stambaugh, 2002). In general, a positive α_i constitutes the added value to the fund investors, while a negative α_i refers to value destruction without confining ourselves to interpretations of selectivity.

Section 3.2 Conditional state-dependent regression model of performance

To investigate **H02**, a simple modification of the four-factor Carhart (1997) model will be implemented to account for time-varying ability in fund managers. This procedure follows the methodology of Glode (2011) and adds a binary state-dependent variable indicating the different economic states. The model is defined as:

$$R_{i,t} - R_{f,t} = \alpha_i + \gamma_1 FC + \beta_{1,i}\left(R_{MKT,t} - R_{f,t}\right) + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} PR1YR_t + \gamma_2 FC * RMRF_t + \gamma_3 FC * SMB_t + \gamma_4 FC * HML_t + \gamma_5 FC * PR1YR_t + \epsilon_{i,t}\ , \quad (5)$$

where FC is state-dependent *dummy* variable which follows NBER recession dates to capture the period of the recent financial crisis. FC assumes values of 1 for the period between December 2007 to June 2009 – the NBER reported period encompassing the crisis, and for the rest of the periods it assumes values of 0. In addition, interaction terms are added between FC and the four factors to account for time-varying exposures to the different investment styles and the market index. The exposure to the market is expected to increase in bull periods and decrease in bearish periods (Treynor & Mazuy, 1966). The regression intercept (α_i) here refers to the conditional *alpha* during normal times, and the coefficient (γ_1) of the FC binary indicator gives the marginal contribution to abnormal performance attributable to the time-varying ability of the fund manager during times of crisis. The sum of the intercept (α_i) and the coefficient of FC gives the conditional *distressed* alpha. This term will be frequently employed throughout the text to denote the alpha performance during the financial crisis.

Chapter 4 Data

1. Data Issues

This chapter discusses data selection problems which in turn might impinge on the validity of the research. Early studies on fund performance suffered from survivorship biases or have underestimated its true effect. Survivor bias is a form of selection bias which arises when subsets of the population of funds are systematically excluded from the study. Ferson (2012) identifies two channels through which survivorship bias might be transmitted. Firstly, database limitations might introduce survivorship bias in many forms. For instance, a database may include only surviving funds, thus leaving “dead” funds, which might have ceased operations or merged with other funds due to poor performance, out of the sample. As Malkiel (1995) suggests, a fund manager that takes very high risk will have a high probability of failure. However, if that fund survives, it implies a superior performance, while unsuccessful funds are likely to drop out, leaving a majority of high-return surviving funds which give rise to an illusionary setting with high performance. Alternatively, survivor bias can occur when only fund managers with good records decide to report their returns to a database. In such situations, any study which selects data from such sources, is inherently destined to suffer from upward bias, or in other words, grossly overstate the performance of fund managers. Malkiel (1995) accentuates that using only the surviving funds from his sample in his evaluation would introduce an upward bias of 1.5%. Elon, Gruber, and Blake (1996) estimate an average bias of 0.7-0.9% per year in a sample of mutual funds. A recent comprehensive study by Rohleder, Scholz, and Wilkens (2010) investigates the distortionary effect survivor bias exerts on evaluation studies on mutual fund performance. In their study, they analyse the performance of funds in two identical samples only different in that one of them omits non-surviving funds, while the other includes all funds during the evaluation period 1993 – 2006. They find a 1.57% difference in performance between the two samples attributable to survivor bias, which substantially overstates the performance in the surviving sample. The second transmission channel of survivor bias originates from the design of the study. Studies of performance persistence are typical example where survivor bias is virtually unavoidable (Ferson, 2012). In such studies, usually the evaluation period is divided into two subperiods, and in each period performance is measured. Thus, funds must survive after the initial evaluation period in order to test for persistence in performance. While modern databases have largely addressed the issue of survivor bias, researchers still need to be cautious of biases induced by the design of a study.

Another problem arises as a result of infrequently traded assets which introduce stale prices. If stale prices play a role in the formation of the Net Asset Value (NAV) of a mutual fund, Zitzewitz (2009) argues that the NAV will experience a lagged adjustment which might be predictable. Such inefficiencies might be profitably exploited by sophisticated traders at the expense of the current shareholders.

Another related form of bias in managed databases is backfilling, also commonly known as "incubation bias". Backfilling refers to a situation where a database includes returns that precede a fund's entry into that database. Evans (2010) discusses incubation bias in the context of mutual funds, where incubation is a strategy for setting up new funds, where multiple funds are initiated privately and evaluated during a certain period, at the end of which only successful funds are opened to the public and enter the database. He documented that "incubated" funds outperform their "non-incubated" counterparts by 3.5% and as a result attract higher investor flows. However, after that this outperformance vanishes, and precisely this performance reversal gives rise to the incubation bias by overstating performance. The author accentuates that sorting out funds on age and ticker creation date filter eliminates the bias. Leaving out the first year of observations for ned funds and setting a minimum value for assets under management (AUM) before funds can enter the sample are other techniques aimed at attenuating the effect of the bias (Ferson, 2012).

2. Data Selection and Definitions

This study examines the performance of U.S equity mutual funds in the period 2000 - 2017 using monthly fund returns and the Fama-French market portfolio and benchmark factors covering size, value and momentum premiums. The Fama-French factors are readily available on Kenneth French's web data library[2]. Since, the analysis covers solely U.S. Equity MFs, I stick to the factors specially constructed for North America. In this text, a short overview of the definitions will be presented. For a detailed description of the market index and the four

[2] See: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

factors, the reader is referred to the website of Kenneth French. The market index measures the return on the region's value-weighted market portfolio minus the U.S. one-month T-bill rate. The size factor (SMB) is the equal-weight average of the returns on the three small stock portfolios for the region minus the average of returns on the three big stock portfolios,

$$SMB = 1/3 * (Small\ Value + Small\ Neutral + Small\ Growth) - 1/3 * (Big\ Value + Big\ Neutral + Big\ Growth) \quad (6)$$

The value factor (HML) is the equal-weight average of the returns for the two high B/M portfolios for a region minus the average of the returns for the two low B/M portfolios,

$$HML = 1/2 * (Small\ Value + Big\ Value) - 1/2 * (Small\ Growth + Big\ Growth) \quad (7)$$

The momentum factor (WML) is the equal-weight average of the returns for the two winner portfolios for a region minus the average of the returns for the two loser portfolios,

$$WML = 1/2 * (Small\ High + Big\ High) - 1/2 * (Small\ Low + Big\ Low) \quad (8)$$

The summary statistics on the factor portfolios reported in Table 1 indicate low cross correlations between the different factors which mitigates the distortionary effect multicollinearity can exert on the coefficient estimates and also represent the good fit of the model. The high mean returns of MktRF and HML indicate that these factors should be able to account for much of the cross-sectional variation in the mean return on the fund portfolios, while the relatively high standard deviations of the four factors should be able to explain much of the time-series variation.

Table 1. Factors Summary Statistics, January 2000 to May 2017

Factor Portfolio	Monthly Excess Return	Std. Dev	Sharpe Ratio	*t*-value for Mean = 0	Correlation Matrix			
					MktRF	SMB	HML	WML
MktRF	0.41	4.43	0.32	1.35	1.00			
SMB	0.21	3.11	0.23	0.96	0.30	1.00		
HML	0.39	3.52	0.38	1.60	-0.17	-0.34	1.00	
WML	0.17	5.44	0.11	0.45	-0.26	0.24	-0.19	1.00

This table portrays summary statistics for the Fama-French four factors, where MktRF is excess return on the market proxy. SMB, HML, and WML are the factor-mimicking portfolios for size, value, and momentum. The Sharpe Ratio is annualized by multiplying the monthly Sharpe Ratio by $\sqrt{12}$.

The data for the fund returns is collected from the Morningstar Direct database. A sample of 1925 open-end U.S. domestic equity mutual funds is constructed in the period January, 2000 – May 2017, covering data on both monthly gross and net returns, investment objective as categorized by Morningstar, Morningstar rating, fund size, average manager tenure, turnover ratio, management fees, and annual expense ratios. The sample selection process follows the setup procedure of Fama and French (2010) in order to address the potential biases outlined in the previous section. The Morningstar Direct database includes data on both surviving investments (live) and obsolete (dead) investments, thus eliminating the issue of survivorship bias. Table 2 outlines the filtering criteria applied during the sample selection process.

Table 2. Sample Selection Filtering Morningstar Direct -- Open End Funds

Filtering			# surviving funds	# all funds
US Category Group	=	U.S. Equity	7622	19740
Domicile	=	United States	7608	19726
Oldest Share Class	=	Yes	2279	5055
Inception Date	<	1/1/2012	1850	4511
Fund Size in $USD	>	$5,000,000.00	1833	1925

This table illustrates the filtering criteria used for the data selection process, and the number of funds after the application of each filter.

Funds are selected based on their Oldest Share Class, which denotes the share class in the fund that has the longest history, thus contributing to better performance-comparisons. To alleviate the effect of incubation bias funds are filtered on Inception Date. Thus, I avoid having lots of recently incepted funds with short return histories, and only use funds that appear in the database at least 5 years before the end of the sample period. Thus, we smooth out potential prior superior performance of “incubated funds” as post-inclusion superior performance disappears (Evans, 2010). The Assets Under Management (AUM) test is employed where funds need to pass a $5-million threshold to be included in the sample, in order to filter out “incubated funds” which are likely to have low AUM values during the pre-release period.

The funds are then sorted into investment objective portfolios. Table 3 presents an overview of the number of funds in each category, along with averages of several fund characteristics. The highest average fund size is observed in the Large Blend Category, with an average fund size of $6 billion, while the lowest fund size is observed in funds that invest in small cap growth stocks. In general, it appears that investors favour large cap to small cap stocks, and further they prefer portfolios consisting of both growth and value stocks to single portfolios consisting only of one type. Managers of Small Value funds appear to have the highest average tenure, while the lowest is found in managers of Mid Value funds. In terms of Turnover Ratio (TR), managers of the Small Cap category appear to be the most active, with an average TR of 80%. As it can be expected, management fees exhibit a positive relationship with TR, that is, the more active a fund manager is the higher commissions they earn. Not surprisingly, the same relationship appears to hold between TR and expense ratios, which capture the total expenses incurred by a fund.

Table 3. Descriptive statistics - Fund (averages)

Fund Category	Fund Size USD	Number of Funds	Manager Tenure (Average)	Turnover Ratio %	Management Fee	Expense Ratio
US Fund Large Blend	$6,122,813,662.98	375	7.1	53.8	0.56	0.83
US Fund Large Growth	$3,494,923,987.32	389	8.1	62.3	0.68	1.02
US Fund Large Value	$3,404,711,865.61	284	8.2	54.6	0.62	0.91
US Fund Mid Blend	$2,652,756,079.62	121	8.2	65.5	0.67	1.00
US Fund Mid Growth	$1,537,839,973.31	170	8.0	65.6	0.79	1.11
US Fund Mid Value	$2,000,498,217.19	96	6.9	64.9	0.74	1.05
US Fund Small Blend	$1,339,161,578.02	197	7.4	60.3	0.76	1.08
US Fund Small Growth	$1,087,598,751.11	191	7.7	84.3	0.84	1.19
US Fund Small Value	$1,298,700,139.07	102	8.6	93.2	0.84	1.16

This table presents the summary averages of several fund characteristics partitioned by fund category as designated by Morningstar.

Chapter 5 Results and Discussion

The research question of this thesis will be answered through the separate investigation of the underlying hypotheses constructed at the end of Chapter 2. Section 5.1 evaluates the unconditional performance of fund managers. Section 5.2 addresses the time-varying ability of fund managers during the recent 2008 financial crisis.

Section 5.1 Unconditional Performance Evaluation

The first part of the analysis addresses **H01** by evaluating the unconditional stand-alone performance of the entire sample of mutual funds in the context of CAPM, Fama-French 3-factor model, and Carhart 4-factor model. For all funds in existence during the sample period 2000 – 2017, *alpha* estimates are calculated through the above-mentioned models. The regression results are summarized in Table 4.

Table 4. Analysis of the Performance of 1925 Equity Funds

	Net Returns	**Gross Returns**
	Panel A: CAPM 1-Factor Alphas	
Average α %	0.432	1.452
Zero α	1647	1487
Positive & Significant α	180	415
Negative & Significant α	98	23
	Panel A: Fama and French 3-Factor Alphas	
Average α %	0.192	1.236
Zero α	1710	1559
Positive & Significant α	112	352
Negative & Significant α	103	14
	Panel A: Carhart 4-Factor Alphas	
Average α %	0.156	1.196
Zero α	1683	1534
Positive & Significant α	123	375
Negative & Significant α	119	16

This table compares the number of significant αs under the different model specifications. The significance level used is 5%. Average monthly αs are annualized (12*α). Insignificant alphas are classified as Zero αs.

The table above compares the number of significant alphas under the different model specifications. Under the CAPM, we observe the highest number of significant alphas, while the number of alphas decreases when we look at the multi-factor models. As it can be expected, gross alphas substantially outnumber net alphas. An interesting observation occurs when we compare the number of alphas under the 3-factor model and the 4-factor model. Surprisingly, the amount of both positive and negative alphas increases under the 4-factor model, while the average net and gross alphas decrease monotonically with the number of factors. To test the joint significance of all the alphas under the different model specifications, two equally weighted portfolios are created, one for net returns, and one for gross returns.

Table 5. Equal-weight Net & Gross Portfolios - Alpha Estimates (%)

	α	β MktRF	β SMB	β HML	β WML	R^2
Panel A: Net Returns						
CAPM	0.085	1.01***				0.97
	(1.54)	(81.74)				
Fama-French	0.02	0.973***	0.226***	0.085***		0.99
	(0.63)	(129.03)	(20.06)	(8.87)		
Carhart	0.019	0.975***	0.224***	0.086***	0.0037	0.99
	(0.59)	(119.79)	(18.96)	(8.83)	(0.57)	
Panel B: Gross Returns						
CAPM	0.171***	1.011***				0.97
	(3.12)	(82.15)				
Fama-French	0.107***	0.975***	0.225***	0.084***		0.99
	(3.33)	(129.39)	(19.99)	(8.79)		
Carhart	0.106***	0.976***	0.223***	0.085***	0.0035	0.99
	(3.28)	(120.09)	(18.91)	(8.75)	(0.53)	

This table reports the regression alphas from the equal-weight net and gross portfolios under the different models. Alpha estimates are given in percentages. In the final column, the R^2 represents the amount of variation in portfolio returns explained by the respective model factors. Numbers in parenthesis are *t*-values. Finally, *, **, *** denote a coefficient that is statistically significant at 10%, 5%, 1% level.

The regression results for the two portfolios under the three different models are reported in Table 5. The annualized gross alphas in Panel B indicate that the average fund managed to significantly outperform before expenses by 2.05%, 1.28%, and 1.27% under the CAPM, Fama-French 3-factor model, and Carhart 4-factor model, respectively. This observation indicates that exploiting investment strategies on value, size, and momentum contributes on average 78 basis points to the mean fund return. In the context of the 4-factor model, a fund manager delivering 1.27% superior returns to its investors would create value as long as the expense ratio is lower than that amount. However, the superior results vanished once expenses were taken into account, as it can be inferred from the insignificant alpha estimates under all model specifications in Panel A. As expected, alpha estimates monotonically decrease with the number of factors employed. The gross results are consistent with the individual alpha estimates in Table 4, as implied by the prevalence of positive alphas over negative alphas. The same intuition holds for the net results, where the number of non-zero alphas is substantially lower, with almost even distribution between positive and negative estimates. The momentum factor does not appear to add any significant explanatory power for predicting return as its coefficient is indistinguishable from zero. It appears that the average fund manager did not incorporate the momentum strategy in the portfolio formation. This finding is also consistent with the observation from Table 4, where the number of both positive and negative alphas increased when adding the momentum factor. The significantly positive average gross αs seem to offer support of the Ippolito's (1989) findings that fund managers earn sufficient gross returns to cover their research expenses. In the context of Grossman and Stiglitz (1980), such findings are consistent with the conception of market efficiency which rewards information gathering. The Berk and Green (2004) rational model long-term prediction that most funds exhibit sufficient skill to cover their costs fares well as the net alphas do not indicate that fund managers destroyed value as would be implied by significant net underperformance. The results reported in Table 5 serve to reject **H01**. In the aggregate, fund managers could not create significant value for investors, with the results being robust to the choice of model used to estimate *alpha* performance.

So far, we have observed that fund managers are capable of providing superior performance, albeit only sufficient to cover their information-gathering expenses. In order to gain a better insight into fund performance, funds are separated into 9 equally weighted portfolios based on their investment style as designated by Morningstar. Consequently, each portfolio returns are regressed on the market index and the respective factors under the two different model specifications. The resulting regression output is presented in Table 6.

Table 6. Category Portfolio Analysis – Gross Returns

	CAPM			Carhart 4-factor Model					
	α	β MktRF	R^2	α	β MktRF	β SMB	β HML	β WML	R^2
Large Blend	0.07***	0.936***	0.988	0.058***	0.954***	-0.064***	0.053***	-0.011***	0.995
Large Growth	0.007	1.028***	0.953	0.081*	1.006***	0.026	-0.197***	0.033***	0.98
Large Value	0.205***	0.887***	0.893	0.097**	0.929***	-0.083***	0.296***	-0.047***	0.979
Mid Blend	0.303***	1.012***	0.928	0.175***	0.98***	0.285***	0.213***	0.00002	0.968
Mid Growth	0.168	1.095***	0.874	0.133*	1.016***	0.428***	-0.083***	0.076***	0.966
Mid Value	0.409***	0.99***	0.888	0.243***	0.964***	0.236***	0.361***	-0.074***	0.968
Small Blend	0.363***	1.058***	0.836	0.143***	0.959***	0.664***	0.32***	-0.00041	0.979
Small Growth	0.224	1.157***	0.805	0.117*	1.015***	0.755***	-0.0002	0.063***	0.977
Small Value	0.479***	1.026***	0.799	0.211***	0.931***	0.632***	0.492***	-0.081***	0.971

This table summarizes the regression results under CAPM and Carhart 4-factor Model. Gross *alphas*, given in percentages, are estimated for each fund category equal-weight portfolio. *, **, *** denote a coefficient that is statistically significant at 10%, 5%, 1% level.

Summary statistics and Sharpe ratios are calculated for the 9 portfolios, and presented in Table A1 in the Appendix. The highest risk-adjusted historical performance measured by the Sharpe Ratio is observed for the Small Cap group with an average annualized SR of 0.46, neatly followed by the Mid Cap category with SR of 0.45, and finally the Large Cap group with SR of 0.33. In this risk-adjusted terms, the average Mid- and Small Cap fund managed to beat the market index and the factor-mimicking portfolios, while the Large Cap category could not outperform only the value factor. On the investment style level, the top performing category appears to be the Value group, followed by the Blend.

Table 6 shows that all fund categories except growth funds managed to outperform the market gross of expenses at the 5% significance level. The aggregate insignificant exposure to the momentum factor observed in Table 5, appears to be driven by Mid- and Small-Blend funds which seem to exclude this particular investment strategy from their portfolios in the aggregate. The top performing category of funds seems to be the Value funds. In each size category, these funds outpaced their Blend counterparts by 47 to 82 basis points on an annualized average basis. Funds investing in mid-cap stocks delivered the highest gross abnormal return to their investors, with Mid-Value funds reaching almost 3% on yearly basis.

Table 7. Category Portfolio Analysis – Net Returns

	CAPM			Carhart 4-factor Model					
	α	β MktRF	R^2	α	β MktRF	β SMB	β HML	β WML	R^2
Large Blend	-0.001	0.935***	0.988	-0.014	0.953***	-0.064***	0.054***	-0.012***	0.995
Large Growth	-0.083	1.026***	0.952	-0.009	1.005***	0.029*	-0.196***	0.034***	0.98
Large Value	0.125	0.886***	0.892	0.017	0.928***	-0.084***	0.297***	-0.047***	0.979
Mid Blend	0.213**	1.011***	0.928	0.085	0.979***	0.285***	0.213***	0	0.969
Mid Growth	0.072	1.093***	0.874	0.036	1.015***	0.429***	-0.082***	0.075***	0.966
Mid Value	0.321***	0.988***	0.887	0.155**	0.962***	0.236***	0.362***	-0.075***	0.967
Small Blend	0.274*	1.056***	0.835	0.053	0.956***	0.662***	0.322***	-0.001	0.979
Small Growth	0.122	1.155***	0.805	0.014	1.013***	0.754***	0.001	0.063***	0.977
Small Value	0.379**	1.025***	0.8	0.11*	0.93***	0.631***	0.491***	-0.081***	0.971

This table summarizes the regression results under CAPM and Carhart 4-factor Model. Net *alphas*, given in percentages, are estimated for each fund category equal-weight portfolio. *, **, *** denote a coefficient that is statistically significant at 10%, 5%, 1% level.

After expenses, only Mid-Blend, Mid- and Small-Value funds showed significant superior ability when performance was measured against the market index alone, with Small-Value funds delivering the highest risk-adjusted abnormal performance with an annualized *alpha* of 4.5%, followed by Mid-Value (3.8%) (See Table 7 above). When accounting for the different investment styles under the 4-factor model, significant net outperformance remained only in the Mid- and Small-Value categories[3]. Mid-Value funds remained the ultimate winner with an annualized superior net performance of 1.86%, followed by Small-Value funds (1.32%). These findings give rise to an interesting reversal in performance between the two categories. Relative to Small-Value, its Mid-cap counterpart had higher market exposure and lower exposure to factor investing which might have partially contributed to that change in positions as can be observed in the performance differential between the CAPM and Carhart alphas of the two fund categories. Small-Value had a differential of 3.18 percentage points attributable to its highest exposure to the investment styles, while the relatively lower exposure to these known strategies of Mid-Value funds resulted in 1.94 percentage points performance differential. This observation highlights the superior idiosyncratic stock selection ability of this particular fund category, which is not captured by the market index and the three investment styles. I call the alpha difference between CAPM and the Carhart model – the factor alpha differential (FAD). This factor differential can be used in the context of measuring fund managers' activeness. In other words, an inverse relationship between the FAD and activeness is expected. That is, the lower the factor differential, the lower performance attributable to the known investment strategies, the more active a fund manager is. To the best of my knowledge, this factor differential has not yet been employed as a measure of activeness in the financial literature. Furthermore, it is interesting to see if FAD can serve as a predictor for fund performance. Does lower FAD correspond to higher returns? Are funds with lower FADs more likely to perform better than funds with higher FAD levels? For the sake of investigating that relationship, the FAD should better be calculated as the ratio of the 1-factor alpha to the 4-factor alpha, as two funds having equal FAD ratio can have substantially different FADs. For example, a FAD ratio of 50%, can correspond to 1.8% FAD, but also to a 0.5% FAD.

Section 5.2 Conditional Performance Evaluation – Recession Model

The previous section covered the unconditional analysis of the sample mutual funds in the context of two different asset pricing models in the period 2000 to 2017, a period covering the financial crisis of 2008 whose damaging effect has been unmet by any other crisis since the

[3] The alpha of Small-Value was only significant at the 10% level.

Great Depression. The initial analysis assumed constant variation in returns and risk throughout the evaluation period. However, the literature has documented significant variation with time, most prominently observable in the course of the business cycles giving rise to time-varying ability of fund managers in performance terms. In this section, we employ the Recession model from Equation 5 as a direct test of the second hypothesis in order to investigate whether fund managers managed to provide insurance against the financial crisis by providing superior performance. If active funds can beat the market during times of crisis this would rationalize investment behaviour in the light of their average underperformance documented in the financial literature.

Table 8 reports the regression results from the recession model. In the context of this conditional model, the FC is the binary indicator for the duration of the financial crisis. Interaction terms are set between FC and the four factors of the Carhart model in order to account for time-varying exposure to the different factors, which is likely to occur in high volatility periods. Here the regression intercept (α) refers to the alpha conditioned on normal times, or in other words, the entire sample period beyond the period entailing the financial crisis, while the FC variable adjusts the intercept to the *alpha* performance during the crisis period. In other words, the FC variable indicates the sensitivity of fund managers' ability to provide superior results in economic contractions. Throughout the text, the terms conditional crisis alpha and distressed alpha will be used interchangeably to denote the *alpha* during the financial crisis. A positive and significant FC coefficient indicates that funds delivered better performance in times of crises as compared to regular times, while the opposite is true for negative coefficients. In Panel A, the Growth category is the only one with insignificant alpha performance before costs[4], while Value funds remain the top performer with conditional "normal" *alpha* ranging from 1.43% to 2.9% in annualized terms. Similar to the unconditional results from the four-factor model in the prior section, the only statistically significant conditional net alpha remains for the Mid-Value category with an annualized value of 1.87%. What is more important from the results is the striking sensitivity of fund performance to the financial crisis, albeit statistically insignificant at the 5% level. All fund categories except the Small-Blend and Small-Value, exhibit negative sensitivity to the financial crisis indicating

[4] At a significance level of 5%.

poorer average performance during the crisis. Despite the statistical insignificance of these results, which is not surprising given the small crisis sample used to calculate the coefficients and the higher volatility during that period, the economic implications seem acute. The Mid-Growth category underperformed by 2.06% per year during the distressed period relative to the normal period. Interestingly, even the top-performer category Mid-Value delivered inferior recessionary performance amounting to just 0.31% per year as compared to the non-recessionary period of 1.87% per year. Despite the negative recession outlook, the Small-Value category emerges as the only group able to provide economically significant superior performance during the crisis period, with a conditional distressed alpha of 2.7%, however not statistically significant. Once again, the short sample period covering the financial crisis coupled with increased volatility limits the explanatory power of the results, and does not give any reliable guidance as to what investment strategy might prove viable for investors seeking to weather the harsh conditions of a financial crisis by channelling their capital to a certain category of funds able to provide a safe haven.

Contrary to expectation, no category significantly changed its exposure to market risk during the crisis, with main prevalence of positive coefficients indicating an increased appetite for systematic risk. The mean market beta of all fund categories is close to unity, indicating that funds held portfolios with volatility equal to that of the market, which is not uncommon phenomenon for the well diversified equity funds comprising the sample. With respect to the investment strategies, no significant change in the momentum factor was present for any of the fund categories during the crisis, while an interesting pattern is observed in how funds in aggregate managed their exposures to the size and value factors during the recession. All fund categories except Mid-Growth, increased their exposure to the size factor, albeit only significant for Large-Blend, Mid- and Small-Value, and Small-Blend. The overall negative reaction to the value exposure was only significant for the Middle-Cap group. The significant change in value and size exposure for these funds, implies that during the crisis period funds increased their portfolio holdings with small stocks at the expense of high book-to-market (value) stocks. Elgammal, Bas, Gough, Shah and van Dellen (2016) analysed the relationship between financial crises, while differentiating between liquidity crisis and financial distress, and value and size premiums in the period 1982 to 2011. They found that the increased default risk during times of crisis increases the value and size premiums. Furthermore, liquidity crisis was found to have significant predictive power when examining the volatility of large stocks' value premium and size premiums. In particular, they conjectured that a liquidity crisis may increase the volatility of the size and value factors. The increased volatility of these stock premiums might partially explain why we observed an overall negative exposure to the value factor. This argument is corroborated by Fama and French (1996) who show that value stocks exhibit higher vulnerability than growth stocks. This, however, does not explain the positive change in the size factor. One explanation might be that managers believed that size premiums would rise higher than value premiums during the crisis, thus increasing size exposure at the expense of value exposure. Another explanation from the perceived favour of small stocks during financial distress might be that they are prone to less analyst coverage since analysts mainly might direct their attention more towards bigger stocks during such times. The relatively lower coverage for small stocks implies that their returns might be less negatively impacted.

Finally, the results reported in Table 8 and discussed above serve to reject **H02**. None of the fund categories managed to provide higher alpha performance during recessions, but the exact opposite, the dominance of negative coefficients implies that fund managers in fact realized inferior performance during times of crises, albeit statistically insignificant. Nevertheless, the results pose important economic implications for the general underperformance of funds during the harsh economic conditions brought about the recent financial crisis.

Taking together the conclusions from the two hypotheses, the following conjecture regarding the answer of the research question emerges. There is no significant and conclusive evidence that the U.S. Equity Mutual Funds in the sample selected to represent the population managed to beat the market during the financial crisis of 2008 when controlling for size, value, and momentum investment strategies. None of the fund categories in aggregate managed to provide superior alpha performance during the crisis period, with the majority of funds realizing economically significant underperformance reflecting the serious impact on the economy and the financial sector as a whole. Finally, it can be conjectured that investors would be better off by just committing their funds to low-cost passive fund portfolios constructed on the market index and the three investment strategies in times of crisis as an economically superior strategy providing financial cushion in times of crises.

Table 8. Recession Model - Time-varying ability of fund managers

Panel A: Gross Returns

Category	α	FC	β MktRF	β SMB	β HML	β WML	β FCMktRF	β FC SMB	β FCHML	β FCWML	R^2
Large Blend	0.069***	-0.147*	0.951***	-0.069***	0.052***	-0.012**	-0.01	0.124***	-0.007	0.008	0.995
Large Growth	0.071	-0.196	1.012***	0.016	-0.191***	0.039***	-0.042	0.121	-0.121*	-0.049*	0.981
Large Value	0.119***	-0.202	0.922***	-0.078***	0.299***	-0.056***	0.022	0.079	-0.084	0.031	0.98
Mid Blend	0.174***	-0.174	0.973***	0.281***	0.226***	-0.002	0.032	0.173*	-0.285***	-0.016	0.972
Mid Growth	0.142*	-0.213	1.008***	0.431***	-0.076***	0.078***	0.041	-0.069	-0.235**	-0.057	0.967
Mid Value	0.242***	-0.124	0.956***	0.23***	0.373***	-0.078***	0.035	0.26***	-0.257***	0.013	0.972
Small Blend	0.136**	0.015	0.955***	0.648***	0.319***	0.007	0.003	0.213**	-0.043	-0.011	0.98
Small Growth	0.11*	-0.112	1.018***	0.752***	0.004	0.067***	-0.017	0.009	-0.094	-0.039	0.977
Small Value	0.19***	0.142	0.934***	0.616***	0.492***	-0.076***	-0.019	0.25**	0.006	0.014	0.973

Panel B: Net Returns

Category	α	FC	β MktRF	β SMB	β HML	β WML	β FCMktRF	β FCSMB	β FCHML	β FCWML	R^2
Large Blend	-0.003	-0.149*	0.951***	-0.069***	0.052***	-0.012**	-0.01	0.124***	-0.007	0.008	0.995
Large Growth	-0.019	-0.196	1.011***	0.019	-0.191***	0.041***	-0.041	0.119	-0.12*	-0.05*	0.981
Large Value	0.039	-0.203	0.921***	-0.08***	0.3***	-0.056***	0.022	0.081	-0.085	0.031	0.98
Mid Blend	0.084	-0.174	0.972***	0.28***	0.226***	-0.002	0.032	0.173*	-0.285***	-0.016	0.972
Mid Growth	0.045	-0.217	1.006***	0.432***	-0.074***	0.077***	0.042	-0.067	-0.236**	-0.058	0.967
Mid Value	0.156**	-0.13	0.953***	0.23***	0.374***	-0.08***	0.037	0.26***	-0.258***	0.016	0.971
Small Blend	0.047	0.005	0.952***	0.647***	0.32***	0.006	0.004	0.214**	-0.044	-0.01	0.98
Small Growth	0.009	-0.117	1.016***	0.752***	0.006	0.066***	-0.016	0.009	-0.095	-0.038	0.977
Small Value	0.089	0.14	0.933***	0.616***	0.491***	-0.075***	-0.019	0.251**	0.007	0.014	0.973

This table reports the results from the Conditional Performance Evaluation Model. Alphas are given in percentages. Fund alphas are both estimated gross and net of expenses. FC is the binary indicator for the financial crisis, while the FCMktRF, FCSMB, FCHML, and FCWML are the interaction terms which represent the marginal change in the estimates during the period of the financial crisis. The regression intercept (α) plus the coefficient of the FC variable return the conditional *distressed* alpha. *, **, *** denote a coefficient that is statistically significant at 10%, 5%, 1% level.

Chapter 6 Limitations and Suggestions for Future Research

The organization of the study does not come without limitations. Firstly, there is one inherent limitation which arises from the treatment of costs in analysing performance, and especially measuring value added by managers from net *alphas*. As we saw earlier investing in active funds comes at a cost amounting about 2/3 of a percent more per year as compared to the low-cost passive investing. Ferson (2012) maintains that most performance measures are rough in how they treat the various expenses associated with active management. He argues that in most academic studies the benchmark strategy does not account for such costs, as for instance, "S&P 500 and the CRSP indexes pay no cost when their composition changes", which in essence gives a comparison between two different "kettles of fish". Therefore, investors need to be cautious when following investment strategies which are based on such comparisons as they might portray a misguided picture not fully reflecting all underlying costs. To get a better sense of performance after costs, investors need to take into account the potential trading costs involved. Hence, it makes sense in future studies to adjust the passive benchmarks to reflect costs of trading. Only then, would we achieve a fairer comparison and mitigate the potential of misleading investment advice. As a suggestion for future research, the market index can be replaced by the returns of a Closet Indexers, a type of passive mutual fund which closely tracks the market index as defined by Cremers and Petajisto (2009). In addition, another limitation concerns the benchmark choice. This study employs one market index as a benchmark for all categories. While it would be optimal to compare each fund with its prospectus benchmark to evaluate performance, as in theory that is the index the fund manager has officially committed to beating, it is cumbersome to construct a large sample of funds and compare each to its designated benchmark. A study by Sensoy (2009) condemns the use of self-designated benchmarks for performance evaluation as the author finds that nearly 1/3 of actively managed U.S. Equity mutual funds choose opportunistically a "mismatched" benchmark index in an attempt to improve flows. In order to counter such strategic behaviour by managers, another suggestion for future research might be the use of category matching benchmarks which are readily available for each fund category used in this study by Morningstar.

Secondly, the results indicate that the average fund managed to earn significantly high returns to cover the research costs without resulting in a statistically significant net underperformance, which appears to contradict the frequently maintained notion of underperformance after costs. The results might suffer from model misspecification. In a recent study, Jordan and Riley (2015) identify a volatility “vol” anomaly factor constructed on the difference in returns on portfolios of low and high volatility stocks. The authors argue that the omission of this factor can lead to substantial mismeasurement in managerial talent, as supported by their results where after controlling for the volatility factor abnormal performance is washed away. Therefore, the results of my research should be interpreted with caution as they might be subject to model misspecification. As a result, future works might make use of including this volatility anomaly to address specification issues and mitigate misestimation of fund performance.

Thirdly, methodology used in this study employs the factor-mimicking portfolios from Fama and French. Cremers, Petajisto, and Zitzewitz (2012) revisit the performance evaluation and obtain that the standard Fama-French and Carhart pricing models assign economically and statistically significant non-zero alphas, even for passive indices such as the S&P 500 and Russel 2000. They attribute the emergence of these alphas to disproportionate weights attached to the value and size factors, and as a result they propose a small methodological modification to the construction of the factors along with tradable benchmark indices in order to improve performance evaluation of actively managed funds. As a consequence, one suggestion for future study, would be to employ the modified versions of these factors to account for alpha mismeasurements.

Fourthly, the conditional evaluation model is based on NBER recession dates. This approach has been criticized as consisting of information which is only available to investors *ex-post*. Instead, *ex-ante* information such as dividend yields and term spreads can be incorporated in future studies to address this limitation (Lynch et al., 2002). Furthermore, the study only investigates the effect of the financial crisis of 2008. In essence, recession periods generally differ and are triggered by different factors and impact the economy on a varying scale. Thus, concentrating only on one recessionary period imposes limits to the generalizability of the results as to whether funds in general can provide superior performance in recessions. One suggestion for future research is to include more recessionary periods similar in characteristics and impact to the financial crisis employed in this study, in order to improve the external validity of the research. In terms of internal validity, the simple recession model might not be sophisticated enough to provide conclusive results. Therefore, as an elaboration and robustness check, a conditional multi-factor model, allowing both for risk and performance to be a function of information available at the beginning of the period (Lynch et al., 2002), can be employed as a way to provide more rigorous conditional alpha estimation.

Chapter 7 Conclusion

Starting with Jensen (1968), the average after costs underperformance of mutual funds as compared to passive market proxies is extensively documented in the financial literature. Later studies, however, indicated that fund managers are capable of earning high enough returns to cover their research expenditures (Grinblatt & Titman, 1992), (Ippolito, 1989). However, Malkiel (1995), Gruber (1996) and Fama and French (2010), among others, argue that most of the earlier studies suffered from survivor bias which introduced an upward bias in the performance evaluation. After controlling for this bias, they conclude that most active funds underperformed with the amount of their expenses, raising the investment preference for low cost passive index funds. Yet, despite this apparent inferiority to passive investments, the mutual fund industry has thrived to manage $16 trillion at the end of 2016. The question as to what justifies their increased popularity remains in the light of their inability to perform better than passive investment strategies. However, the traditional study designs estimate unconditional performance, assuming away time-varying variation in risk and returns. The development of the Conditional Performance Evaluation addresses this issue, and provides results appearing to rationalize investors' behaviour of investing in actively manged funds. For instance, Kosowski (2006) estimates the difference in risk-adjusted return between recessions and expansions of mutual funds to be between 3% to 5%. If mutual funds are able to provide insurance against bear markets when returns are most valued by investors, investing in active funds on average would not prove such a relatively fruitless endeavour.

In this thesis, I revisit the performance of mutual funds employing both unconditional and conditional evaluation approaches to investigate whether active fund managers are able to beat the market unconditionally on average, and whether funds were able to provide financial cushion in the recent 2008 financial crisis, inspired by the earlier findings of Kosowski (2006). A sample free of survivor bias is constructed from the Morningstar database covering monthly data on returns and various fund characteristics over the period 2000 to 2017 for 1925 U.S. equity mutual funds. Firstly, unconditional performance is estimated for each fund, as well as for two equal-weigh portfolios reflecting the net and gross returns of all funds, and also for 9 equally-weighted portfolios grouped by fund investment objective. The findings show that the average fund managed to realize an unconditional statistically significant four-factor gross alpha of 1.3% per year. In addition, the dominance of positive and significant alphas further corroborates with the earlier results of Grinblatt and Titman (1992) and Ippolito (1989) that active fund managers are able to cover their research expenses without destroying value. Controlling for expenses, the average fund could not provide significant superior performance and create value to investors. Overall, the findings are consistent with a state of the market which remunerates research gathering efforts. In an attempt to achieve a better insight of which funds perform better and to account for the possibility that certain group of funds can outperform the market after costs but their results are smoothed in the average, the funds were grouped by their investment objective as designated by Morningstar in 9 equal-weight portfolios. The results ranked the Mid-Cap funds as the top performing category, with the Mid-Value group reaching a four-factor annualized gross alpha of 2.9%. An interesting observation emerged for the Growth category, for any size level, growth funds could not provide statistically significant gross performance at the 5% level. Additionally, the Large-Cap category had the lowest average alphas during the sample period. Another thought provoking observation among the investment styles is that fund managers had the least relative exposure to the momentum strategy, which might be supported by the fact that it also provided the lowest risk-adjusted return measured by the Sharpe Ratio in Table 1, implying that fund managers were able to correctly predict this development and adjust their exposures strategically. On average, the fund categories had a higher exposure to the size strategy relative to its value counterpart. This, however, is not supported by the factors' Sharpe Ratios over the sample period, as the value portfolio had a higher risk-adjusted return. After controlling for costs, only the Mid-Value category sustained its positive and significant outperformance totalling an average annualized four-factor alpha of 1.87%. In addition, this group maintained a higher exposure to the value factor relative to the size factor, correctly guessing its higher risk-adjusted return. It appears that the most skilled managers were concentrated in that category.

The results from the conditional recession model partially side with the findings of De Souza and Lynch (2012) that not all fund styles are able to provide superior performance during times of crises. In my study, not a single fund category managed to provide superior performance during the financial crisis. Albeit statistically insignificant, the economic implications for the ability of fund managers to provide recessionary insurance seem severe. For all fund groups except for the Small Blend, Mid- and Small-Value funds, the average recession alpha was negative, indicating that these funds underperformed the market during the crisis as evidenced by the highest annualized average underperformance reaching 2.6%. Interestingly, the only fund categories which performed better during the recession were Small Blend and Value. The Small Value category performed on average 1.7% per year better during the bear market than in normal states of the market. However, this result was not significantly different from zero. Nevertheless, this magnitude of superior performance in market distress is of economic significance.

The general negative outlook during the financial crisis marks the inferior ability of fund managers to perform better during recessions with weak evidence for taming a distressed market. In aggregate, the study fails to provide supporting evidence for significance time-varying ability of fund managers. As a result, an investor seeking safer shelter for his funds during times of crises would be better off investing his capital into funds managing passive portfolios constructed on the market index and the three investment strategies. However, the reader needs to be cognizant of the various limitations discussed in the preceding chapter when interpreting the results. Possible model misspecification and an inappropriate choice of the market benchmark might introduce distortions and plague the validity of the results. Future research addressing the limitations of this study is needed to provide a better insight and guidance for profitable investment strategies.

Appendix

Table A1. Summary Statistics & Sharpe ratios for the 9 equal-weight portfolios.

Fund Category	Mean Excess Return	Std. Dev.	Sharpe Ratio	Min	Max
Large Blend	0.39	4.17	0.32	-17.32	11.06
Large Growth	0.34	4.66	0.25	-17.38	11.59
Large Value	0.49	4.16	0.41	-17.47	11.09
Total Large	0.41	4.33	0.33	-17.39	11.25
Mid Blend	0.63	4.65	0.47	-20.54	13.73
Mid Growth	0.52	5.19	0.35	-20.16	16.26
Mid Value	0.73	4.65	0.54	-20.79	15.8
Total Mid	0.63	4.83	0.45	-20.5	15.26
Small Blend	0.71	5.12	0.48	-20.8	16.37
Small Growth	0.6	5.71	0.36	-20.92	18.99
Small Value	0.8	5.08	0.55	-20.4	18.69
Total Small	0.7	5.3	0.46	-20.71	18.02

This table presents the summary statistics, and the annualized Sharpe Ratios for the 9 equal-weight portfolios grouped by investment objective as designated by Morningstar. The above calculations are performed for the net monthly returns over the entire sample period.

Bibliography

Amihud, Y., & Goyenko, R. (2013). Mutual fund's R 2 as predictor of performance. *The Review of Financial Studies*, *26*(3), 667-694.

Aragon, G. O., & Ferson, W. E. (2007). Portfolio performance evaluation. *Foundations and Trends® in Finance*, *2*(2), 83-190.

Asness, C. S., Liew, J. M., & Stevens, R. L. (1997). Parallels between the cross-sectional predictability of stock and country returns. *The Journal of Portfolio Management*, *23*(3), 79-87.

Badrinath, S. G., & Gubellini, S. (2012). Does conditional mutual fund outperformance exist? *Managerial Finance*, *38*(12), 1160-1183.

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, *112*(6), 1269-1295.

Berk, J. B., & Van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, *118*(1), 1-20.

Bollen, N. P., & Busse, J. A. (2001). On the timing ability of mutual fund managers. *The Journal of Finance*, *56*(3), 1075-1094.

Brown, S. J., & Goetzmann, W. N. (1997). Mutual fund styles. *Journal of Financial Economics*, *43*(3), 373-399.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, *52*(1), 57-82.

Cremers, K. M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *The Review of Financial Studies*, *22*(9), 3329-3365.

Cremers, M., Petajisto, A., & Zitzewitz, E. (2012). *Should benchmark indices have alpha? Revisiting performance evaluation* (Working paper No. 18050). National Bureau of Economic Research.

Cuthbertson, K., Nitzsche, D., & O'Sullivan, N. (2010). Mutual fund performance: measurement and evidence. *Financial Markets, Institutions & Instruments*, *19*(2), 95-187.

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance*, *52*(3), 1035-1058.

De Souza, A., & Lynch, A. W. (2012). *Does mutual fund performance vary over the business cycle?* (Working paper No. 18137). National Bureau of Economic Research.

Elgammal, M. M., Bas, T., Gough, O., Shah, N., & van Dellen, S. (2016). Do financial distress and liquidity crises affect value and size premiums? *Applied Economics*, *48*(39), 3734-3751.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivor bias and mutual fund performance. *The Review of Financial Studies*, *9*(4), 1097-1120.

Elton, E. J., Gruber, M. J., Das, S., & Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *The review of financial studies*, *6*(1), 1-22.

Evans, R. B. (2010). Mutual fund incubation. *The Journal of Finance*, *65*(4), 1581-1611.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, *25*(2), 383-417.

Fama, E. F. (1972). Components of investment performance. *The Journal of Finance*, *27*(3), 551-567.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, *47*(2), 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3-56.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The journal of finance*, *51*(1), 55-84.

Fama, E. F., & French, K. R. (1998). Value versus growth: The international evidence. *The Journal of Finance*, *53*(6), 1975-1999.

Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *The Journal of Finance*, *65*(5), 1915-1947.

Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *The Journal of Finance*, *51*(2), 425-461.

Ferson, W. E., & Qian, M. (2004). Conditional performance evaluation, revisited. *Research Foundation of the CFA Institute,* 1-85.

Ferson, W. (2012). Investment performance: A review and synthesis. *Handbook of Economics and Finance*, *2*, 969-1010.

French, K. R. (2008). Presidential address: The cost of active investing. *The Journal of Finance*, *63*(4), 1537-1573.

Glode, V. (2011). Why mutual funds “underperform”. *Journal of Financial Economics*, *99*(3), 546-559.

Grant, D. (1977). Portfolio performance and the “cost” of timing decisions. *The Journal of Finance*, *32*(3), 837-846.

Grinblatt, M., & Titman, S. (1992). The persistence of mutual fund performance. *The Journal of Finance*, *47*(5), 1977-1984.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, *70*(3), 393-408.

Gruber, M. J. (1996). Another puzzle: The growth in actively managed mutual funds. *The Journal of Finance*, *51*(3), 783-810.

Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988. *The Journal of Finance*, *48*(1), 93-130.

Investment Company Institute. (2017). 2017 Investment Company Factbook: A Review of Trends and Activities in the Investment Company Industry. Retrieved from www.icifactbook.org

Ippolito, R. A. (1989). Efficiency with costly information: A study of mutual fund performance, 1965–1984. *The Quarterly Journal of Economics*, *104*(1), 1-23.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, *48*(1), 65-91.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, *23*(2), 389-416.

Jensen, M. C. (1972). Optimal utilization of market forecasts and the evaluation of investment performance.

Jordan, B. D., & Riley, T. B. (2015). Volatility and mutual fund manager skill. *Journal of Financial Economics*, *118*(2), 289-298.

Kacperczyk, M., Sialm, C., & Zheng, L. (2006). Unobserved actions of mutual funds. *The Review of Financial Studies*, *21*(6), 2379-2416.

Kacperczyk, M., Nieuwerburgh, S. V., & Veldkamp, L. (2014). Time-varying fund manager skill. *The Journal of Finance*, *69*(4), 1455-1484.

Klapper, L., & Vittas, D. (2004). The development of mutual funds around the world. *Emerging Markets Review*, *5*(1), 1-38.

Kosowski, R. (2006). Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions. *The Quarterly Journal of Finance*, *1*(03), 607-664.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The Journal of Finance*, *49*(5), 1541-1578.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13-37.

Lynch, A. W., Wachter, J. A., & Boudry, W. (2002). Does mutual fund performance vary over the business cycle? *AFA 2004 San Diego Meetings*.

Malkiel, B. G. (1995). Returns from investing in equity mutual funds 1971 to 1991. *The Journal of Finance*, *50*(2), 549-572.

Mishkin F. S., & Eakins S. G. (2012). *Financial markets and Institutions*. Boston, MA: Pearson Education

Pástor, Ľ., & Stambaugh, R. F. (2002). Investing in equity mutual funds. *Journal of Financial Economics*, *63*(3), 351-380.

Petajisto, A. (2013). Active share and mutual fund performance. *Financial Analysts Journal*, *69*(4), 73-93.

Reinhart, C. M., & Rogoff, K. S. (2008). *Is the 2007 US sub-prime financial crisis so different? An international historical comparison* (Working paper No. 13761). National Bureau of Economic Research.

Rohleder, M., Scholz, H., & Wilkens, M. (2010). Survivorship bias and mutual fund performance: Relevance, significance, and methodical differences. *Review of Finance*, *15*(2), 441-474.

Rouwenhorst, K. G. (2004). The origins of mutual funds. *Yale ICF Working Paper*, 04-44.

Tobin, J. (1958). Liquidity preference as behaviour towards risk. *The Review of Economic Studies*, *25*(2), 65-86.

Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics*, *92*(1), 25-39.

Staal, A., 2006. *Essays in empirical finance* (Doctoral dissertation). Northwestern University.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, *19*(3), 425-442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, *39*(1), 119-138.

Swinkels, L., & Tjong-A-Tjoe, L. (2007). Can mutual funds time investment styles? *Journal of Asset Management*, *8*(2), 123-132.

Treynor, J., & Mazuy, K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*(4), 131-136.

Wermers, R. (1997). Momentum investment strategies of mutual funds, performance persistence, and survivorship bias. *Unpublished Working Paper, University of Colorado, Retrieved from http://bus. colorado. edu/faculty/wermers/*.

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

Zitzewitz, E. W. (2009). Prosecutorial discretion in mutual fund settlement negotiations, 2003-7. *The B.E. Journal of Economic Analysis & Policy*, *9*(1).