Performance Comparison of Institutional and Retail Mutual Funds Using Multifactor Models

Mehmet Sencicek

Mehmet Sencicek is Associate Professor of Economics and Finance, Department of Business, Economics and Finance, School of Business and Justice Studies, Utica College, 1600 Burrstone Road, Utica, NY 13502. E-mail: mesencic@utica.edu

Abstract

Institutional mutual funds cater specifically to institutional and other large investors. They differ from retail mutual funds in important aspects that could affect their relative performance: very high minimum initial and subsequent investment requirements that put them out of the reach of individual investors, large and discreet flow of funds, and a closer relationship between the fund manager and the institutional investor who closely monitors performance using sophisticated measures. Although a substantial amount of literature exists on retail mutual fund characteristics and performance, institutional mutual funds have received little attention to date. This study examines whether differences in fund characteristics influence the relative performance of institutional funds, before fees and expenses. Using mutual fund historical
data from the Morningstar database, appropriate benchmarks and the established multifactor models, Fama and French three-factor model (FF3F) and Carhart’s four factor model (C4F), performances of institutional and retail mutual funds are compared as overall fund classes, and then in categories derived from Morningstar’s style box. Institutional funds’ superior performance detected with the CAPM disappears when tested with FF3F and C4F. In all models, institutional funds outperform retail funds in the large-cap category with the statistically significant difference being attributable to “alphas” (manager skill) in this category. Evidence is mixed in other categories, and differences are not significant.

**Keywords:** Multifactor models, Mutual Fund Performance, Institutional Mutual Funds.

**JEL Classification:** G11, G12, G23

**Data Availability:** Data used in this study are available from public sources.

**DOI:** https://dx.doi.org/10.15239/j.brcacadjb.2020.10.01.ja01

**Introduction and literature review**

The tremendous growth of the mutual fund industry in the last few decades has made it one of the principal financial intermediaries controlling $17.7 trillion in assets as of the end of 2018 (ICI fact book, 2019). Between 1980 and 2018, the portion of households that owned mutual funds increased from 5.7% to 43.9%. Interest in mutual funds has grown despite numerous academic studies showing that actively managed funds perform as well as a market index at best, and often under-perform (e.g. Jensen (1968), Malkiel (1995), Gruber (1996), and Carhart (1997)). Employer-sponsored retirement plans have become the main source of ownership with 80% of households holding mutual funds within an such a plan in 2018 compared to 52% before 1992. Of these households, 43% also held funds outside these plans and an additional 20% only in outside
sources. This growth of interest in mutual funds by retirement plans and other institutional investors including corporations, not-for-profit organizations, endowments, foundations, and municipalities has been due in great part to the bullish stock market since the mid-1990s, except around the financial crisis of 2008. For example, mutual funds constituted 55% of the assets of defined contribution plans in 2016, compared to 9% in 1990. The fund companies have responded to this growing interest by offering more and more funds geared towards institutional investors. While $206 billion of retirement assets were invested in institutional mutual funds in 1990, this rose to nearly $2,324 billion in 2018, a 1,028% increase in 28 years (ICI Mutual Fund Factbook, 2003 and 2019).

Institutional (mutual) funds are a share class of mutual funds intended for institutional and other large investors, and are regulated under the Investment Company Act of 1940 (ICI fact book, 2019). Their target market makes them different from “retail” mutual funds in several respects. First, institutional investors have the resources to conduct in-depth research and use highly specialized analytics tools to evaluate the markets giving them a considerable advantage over the retail investor who buys and sells securities through brokerage firms. Accordingly, the presumably sophisticated institutional investors can closely monitor the performance of a fund manager with whom they have a closer relationship. In their paper titled “Investor monitoring and differences in mutual fund performance”, James & Karceski (2006) found that institutional investors do, in fact, closely monitor the fund manager’s performance. Another important difference arises in the sensitivity of fund flows. In “Costly Search and Mutual Fund Flows”, Sirri & Tufano (1998) demonstrate that retail investors, who face high cost and information costs, “chase returns, flocking to funds with the highest recent returns”. Harless & Peterson (1998) find similar results, and argue that this behavior, which does not take risk and expenses into account, coupled with extremely slow adjustment of fund holdings, allows poorly performing mutual funds to persist. In contrast, flows into institutional funds are sensitive to measures such as risk adjusted returns; institutional investors, not facing high search
costs, can focus on more sophisticated investment selection criteria, and punish poorly performing managers by withdrawing assets (e.g., Del Guercio and Tkac 2002; Goyal and Wahal 2008). A more recent study (Evans & Fahlenbrach, 2012) found that institutional investor flows were more sensitive to poor risk-adjusted performance than retail flows, and retail investors responded more strongly to counterproductive signals like past total (not risk adjusted) return. As a second difference, institutional fund inflows are discrete and large compared to continuous and small inflows of retail funds as they come from institutional investors such as retirement plans where contributions are made periodically rather than continuously. Finally, institutional funds require much higher minimum initial and subsequent investments often in millions and hundreds of thousands of dollars, respectively, effectively putting them out of the reach of individual retail investors. Some funds labeled as “institutional” sell shares to retail investors through brokers; these funds typically have a much lower minimum initial and subsequent investment requirements, and therefore are not considered “institutional” for the purposes of this study.

Given that the relatively more sophisticated institutional investors use risk-adjusted measures, can closely monitor fund managers, and punish poor performance by moving funds, institutional funds have a stronger incentive to produce better returns, and poor performance should not persist. Cremers et al (2009) do indeed find that, as manager incentives increase and interests become more aligned with investor interests, fund performance increases. This, in turn, should theoretically increase the average performance of institutional funds relative to retail funds. Grossman and Stiglitz’s (1980) model of market efficiency predicts that informed investors earn higher returns than uninformed investors. Given the well-established differences between institutional and retail investors as outlined above, and the literature finding evidence that these differences can affect relative performances, the aim of this study is to answer the research question: “Do institutional mutual funds outperform retail mutual funds?”
Performance evaluation of retail mutual funds has been studied extensively for some time (e.g. Jensen, 1968; Grinblatt and Titman, 1992 & 1993; Brown and Goetzmann, 1995; Malkiel, 1995; Gruber, 1996; Carhart, 1997), and continues to be of interest to investors and scholars (e.g. Guercio and Reuter, 2014). Despite their dramatic growth both in the number of funds and assets under management, institutional mutual funds have received little special attention. The sparse literature on the topic includes a study by Evans and Fahlenbrach (2012) that finds retail funds with an institutional “twin” outperform other retail funds presumably benefiting from the advantage of better management of an institutional fund due to differences between the two classes. Baker and Haslem (2009) study the relationship between characteristics and performance of institutional equity funds, but do not compare institutional and retail fund performance. Similarly, Bird et al (2011) examine the connection between the management’s active decisions and the subsequent performance outcomes in institutional funds, but do not address performance differences with retail funds. Applying the traditional performance measures based on the original capital asset pricing model (CAPM), the Sharpe ratio, the Treynor Index and Jensen’s “alpha”, Sencicek (2018) found that institutional funds outperformed retail funds in large-cap and, to a lesser extent, in small-cap categories, but slightly underperformed in mid-cap categories; the evidence was mixed when funds were compared by book-to-market ratios (value, blend and growth).

This study aims to contribute to the performance evaluation of this growing segment of mutual funds geared towards a fast growing market of retirement plans and other institutional investors by applying multifactor models. The measures used in the previous study by the author that compared performances of institutional and retail funds (Sencicek, 2018) were all based on the original CAPM. Despite still being the most widely used model by practitioners and in MBA classes, the original CAPM has been widely criticized for its highly restrictive and unrealistic assumptions, and its questionable adequacy as a single factor model to explain portfolio performance. Two multifactor models are currently
center stage in academic research on asset pricing/portfolio evaluation: *Fama and French’s (1993) three-factor model and Carhart’s four factor model.* (These models and the criticism of the original CAPM model are discussed in detail in section 2.2.). The current study uses these two dominant multifactor models to evaluate the performance of institutional mutual funds relative to their appropriate benchmarks and compare their performance to the corresponding retail funds (as reference group) in categories based on the Morningstar® style box.

Based on the research question “Do institutional mutual funds outperform retail mutual funds?” and evaluation models to be used in the study, the research hypothesis is follows:

\[ H_0: \alpha_i \leq \alpha_r. \] (Institutional funds do not outperform retail funds).

\[ H_1: \alpha_i > \alpha_r. \] (Institutional funds outperform retail funds).

where \(\alpha_i\) and \(\alpha_r\) are alphas (representing manager skill) for institutional and retail funds, respectively, in each of the models used.

The hypothesis will be tested by deriving alphas for each class of fund using each of the three models (CAPM, FF3F and C4F) and comparing alphas of institutional funds to those of retail funds as classes without separating them into categories. It is possible that institutional funds outperform (underperform) retail funds in certain categories whether or not there is any significant performance difference overall. Accordingly, each class will be further be separated into size categories (large cap, midcap and small cap) and book-to-market ratio categories (value, blend, growth) to see if there are performance differences within categories. This approach is consistent with the multifactor models used and will account for different investment strategies. The rest of the paper is organized as follows: section 2 reviews and discusses the traditional measures and the multifactor models of fund (portfolio) performance; section 3 describes the data and methodology; section 4 presents results and conclusions.
A review of performance measures

The Traditional Measures
The most well-known definition of risk with knowable probabilities is due to Frank Knight. In his book based on his dissertation “Risk, Uncertainty and Profit” (1921), Knight made his famous distinction between “risk” and “uncertainty” (randomness with unknowable probabilities). The two most commonly used variables to assess risk in performance evaluation are beta and standard deviation. The beta for (the manager of) a fund ($\beta_p$) is the covariance of the fund's risk premium with the benchmark’s risk premium divided by the variance of the benchmark’s risk premium. In economic terms, beta reflects a fund’s volatility relative to the market or the only market-related portion of a fund's risk (systematic or undiversifiable risk) rather than the total risk (represented by standard deviation of the fund’s returns) which includes both systematic and unsystematic risk. Beta can be estimated by the regression equation below:

$$ (R_{pt} - r_t) = \alpha_p + \beta_p (R_{bt} - r_t) + \epsilon_t, \quad (1) $$

where $R_{pt}$, $R_{bt}$ and $r_t$ are the fund’s return, the benchmark’s return and the risk free rate, respectively.

Another commonly used measure of risk, standard deviation measures the volatility of returns around their average. Standard deviation is not dependent on an index, and therefore can be used to compare funds across asset classes. Since it is a measure of total risk, both systematic and unsystematic, it captures risk more comprehensively than beta. Standard deviation can be calculated readily through a software package or a program such as Microsoft Excel, or by the formula given in the appendix. A third measure, “alpha”, is a measure of the differential earned in an
average period, given the fund's systematic risk (beta). It is estimated by regressing fund returns against benchmark returns.

The first rigorous method applicable to mutual fund evaluation was due to Markowitz’s Mean-Variance Portfolio Theory (1952, 1959). Prior to Markowitz’s paper “Portfolio Selection” (1952), standard investment advice was to construct a portfolio from securities that offered the best opportunities for gain with the least risk, which could lead to a portfolio of assets all from a single sector. Instead of compiling portfolios from securities that each individually have attractive risk-reward characteristics, Markowitz advocated selecting a portfolio based on the overall risk-reward characteristics of the portfolio itself! William Sharpe introduced the Capital Asset Pricing Model (CAPM) in (1964); Lintner (1965), and Mossin (1966) performed parallel work. CAPM extended Markowitz’s (1952, 1959) portfolio theory by introducing systematic and specific risk. Sharpe later received the 1990 Nobel Prize in Economics for his work on CAPM, which he shared with Harry Markowitz and Merton Miller. In this model, if investors have homogeneous expectations and optimally hold mean-variance efficient portfolios, then, in the absence of market friction, the market portfolio itself will be a mean-variance efficient portfolio. In equilibrium, all investors will hold combinations of the market portfolio of risky assets and the risk-free asset. The expected return on an asset is a linear function of the asset’s systematic risk, beta: the expected return of an asset equals the risk-free rate plus the portfolio’s beta multiplied by the expected excess return of the market portfolio.

\[ E(R_i) = R_f + [E(R_m)-R_f] \beta_i \]  

(2)

where, \( E(R_i) \) is the expected return on the \( i \)th asset, \( R_f \) is the risk-free rate, \( E(R_m) \) is the expected return on the market portfolio and \( \beta_i = \text{Cov}(R_i,R_m)/\text{Var}(R_m) \) is the systematic risk of asset \( i \), and measures the volatility of the security relative to the market.

If we let \( R_p \) be a portfolio’s simple return, and let now denote the portfolio’s beta, we obtain
The essential conclusion of CAPM is that a stock’s (or portfolio’s) excess expected return depends on its beta and not its volatility, in other words, on systematic risk and not on total risk. A one factor model, CAPM describes a linear relationship between expected returns and systematic risk as measured by ‘beta’, the idea being that investors demand additional expected return (called ‘the risk premium’) to accept additional risk. It provides a benchmark against which mutual funds and actively managed portfolios can be evaluated. The risk adjusted performance measures used in an earlier study by the author (Sencicek, 2018), the Treynor Index (1965), the Sharpe Ratio (1966) and Jensen’s ‘alpha’ (1968), were based on Capital Asset Pricing Model and Markowitz’s Portfolio Theory. All three measures, known as the “traditional performance measures”, try to reduce the risk-reward dimensions of portfolio performance to a single measure that indicates a risk-adjusted return. Since they are based on Capital Asset Pricing Model, which assumes a single period investment horizon, Sharpe Index, Treynor Index and Jensen’s model are single period models. In reality, investors have multiple investment horizons; this implies that the functional form of the relationship between expected returns and systematic risk will be nonlinear. Despite the highly restrictive and theoretical assumptions of CAPM, and the wide criticism leveled by academicians and proponents of other methods (e.g. Fama and French, 1996), the CAPM based traditional models are still the most widely used measures of performance by academicians and practitioners, and in MBA classes.

**Multifactor Models and other alternatives to CAPM**

The main result of the CAPM model is a statement of the relation between the expected risk premiums on an asset and its systematic risk. Almost since its inception, the model has been controversial. While some scholars and practitioners hailed it as revolutionary concepts in finance, others criticized its highly restrictive and unrealistic assumptions. These
assumptions include a single period transaction horizon, investors’ ability to borrow and lend at the risk free rate, and a perfect capital market in which there are taxes or transaction costs, but there are large number of buyers and sellers, who are rational and risk averse, trying to maximize their own utility, and to whom perfect information is freely available which would imply that all investors have the same expectations. Early tests of the CAPM showed that higher beta stocks generally had higher stock returns, as the model predicted. Empirical work by Black et al. (1972), Miller and Scholes (1972) and Fama & McBeth (1973) documented a clear relationship between an asset’s beta and its returns. However, the returns on stocks with higher (lower) betas were systematically less (more) than predicted by the CAPM. In response, Black et al (1972) proposed a two-factor model (with factor loadings on the market and a zero-beta portfolio). Thus, it was claimed that the CAPM could be fixed by substituting the risk-free rate in the model with the rate of return on a portfolio of stocks with zero beta.

Although CAPM is theoretically sound, its empirical validity and the adequacy of a single factor model to explain portfolio performance has been called into question by several studies starting in the 1970s. Scholars have tried to get more generalized asset pricing models by relaxing some of these assumptions. Black et al. (1972) provided the minimum-variance zero-beta portfolio to address the problem if there is no risk-free asset that has constant returns in every state of nature. Merton (1973) relaxed the single-period assumption of CAPM to develop the intertemporal CAPM (ICAPM) model with stochastic investment opportunities where expected return on any asset is deduced from a multi-beta version of CAPM in a continuous-time model. The empirical work by researchers including Black et al. (1972), Miller & Scholes (1972), Fama & MacBeth (1973) and Blume & Friend (1973) showed that the original CAPM underestimate (overestimate) the real returns for high (low) beta stocks. Some studies (e.g. Roll, 1977) found evidence that contradict the CAPM and reject the hypothesis that the intercept of the security market line in CAPM equals the risk-free rate (meaning a regression of excess returns on the excess
Performance Comparison of Mutual Funds

return of the market is zero). Other researchers identified factors other than beta that explain the portion of returns not captured by beta. For example, Brennan (1970) proposed an extended form of the single period CAPM model that accounted for the differential taxation of dividends over capital gains. Extending Brennan’s model by assuming both margin and income constraints on borrowing, Litzenberger and Ramaswamy (1979) found higher rate of returns were associated with high dividend yields. Basu (1977) found that low price/earnings portfolios show higher rate of returns than what could be explained using CAPM. Banz (1981) and Reinganum (1981) showed that smaller firms earn high abnormal returns.

The ability of other variables in explaining the portion of returns unexplained by CAPM indicates that CAPM is misspecified and needs modification to explain security returns. Based on their finding that “investors have an aversion for variance and a preference for positive skewness”, Kraus and Litzenberger (1976) extended the single factor CAPM to incorporate the effect of preference for skewness. Empirically, they showed that the two factor (three moment) CAPM is a better predictor of returns than its original one-factor CAPM. Subsequently, they constructed a three moment (mean, variance, and skewness) consumption-oriented CAPM to correctly price a subset of assets. They argued that CAPM’s assumptions of risk averse investors with identical expectations and utility functions in an efficient capital market risk imply that these investors exhibit preference for expected return, aversion to variance of return, and preference for positive skewness (Kraus & Litzenberger, 1983). Breeden (1979) used the same continuous-time economic framework as Merton (1973) to show Merton’s multi-beta pricing equation can be collapsed into a single-beta equation, and as such, the expected return on any asset is proportional to its beta with respect to aggregate consumption alone. Grinols (1984) extended Merton’s intertemporal capital asset pricing model with multiple consumers by including a description of the supply of traded securities.
Empirical work by Fama and French (1992) showed that “when the tests allow for variation in "beta" that is unrelated to size, the relation between market "beta" and average return is flat, even when "beta" is the only explanatory variable.” A model that incorporated book-to-market ratio and market capitalization explained the cross-section variation of average returns more appropriately than single factor CAPM, leading the authors to conclude that market capitalization and that book-to-market ratio should replace beta altogether. Subsequently, Fama and French (1993) proposed a model that modified CAPM by adding two more factors to CAPM, namely SMB (small minus big), the difference between the returns on portfolios of small cap stocks and large cap stocks) and HML (high minus low; the difference between the returns on portfolios of high book-to-market ratio and low book-to-market ratio stocks). In a paper published the same year, Jegadeesh and Titman (1993) documented that “strategies which buy stocks that have performed well in the past and sell stocks that have performed poorly in the past generate significant positive returns over 3 to 12 month holding periods”, and that the profitability of these strategies were not due to their systematic risk or to delayed stock price reactions to common factors. They also found that part of the abnormal returns generated in the first year after portfolio formation dissipated in the following two years. Carhart (1997) used Fama & French’s (1993) three-factor model and added a fourth factor that captured Jegadeesh and Titman’s (1992) so-called “one year momentum anomaly” thereby constructing his 4-factor model (henceforth “C4F”).

Although numerous models have been introduced to improve upon the CAPM since the 1970s, Fama and French (1993) three-factor model (henceforth FF3F) was considered to be the one model that explains fund behavior better than CAPM and other proposed models. The model contained a value-weighted market proxy and included two additional risk proxies, namely the returns on size and book-to-market-sorted equity portfolios, mitigating average CAPM pricing errors. However, FF3F was not able to explain the cross-sectional variation in momentum-sorted portfolio returns. By adding a fourth factor that captures the Jegadeesh
and Titman (1993) momentum anomaly, Carhart (1997) made his model consistent with a market equilibrium model with four risk factors. Before and since the introduction of FF3F and C4F, numerous models have been proposed that purport to improve upon the original static, one period, one factor CAPM, and/or its variants, or to offer significantly different alternatives for asset pricing. Some of the better known of these models include (1) Intertemporal CAPM consumption (or production) based models that modify Merton (1972) and/or Breeden’s (1979) models (e.g. Jagannathan and Wang, 1996; Lettau and Ludvigson, 2001; Balvers and Huang, 2007 & 2009), (2) equilibrium models with heterogeneous time horizons or beliefs and investors that relax the original CAPM’s time horizon or homogeneous investors and expectations assumption (e.g. Harvey and Siddique, 2000; Basak, 2005; Levy et al., 2006; Krasny, 2009), (3) Liquidity based models that attempt to capture “liquidity premium” in asset returns (e.g. Pastor and Stambaugh, 2003; Acharya and Pedersen, 2005), (4) Behavioral finance models that relax the original CAPM’s assumptions of utility maximization by or risk aversion of investors (e.g. Barberis et al, 2001; Barberis and Huang, 2008; Levy, 2010), and (5) the supply-side effect model by Lee et al (2009) that extends Black’s (1976) dynamic, simultaneous CAPM model by incorporating the supply effect in the asset pricing process. A comprehensive review of alternative theoretical models of capital asset pricing as well as the empirical tests of CAPM can be found in Shih et al (2014). More recently, motivated by the evidence offered by Novy-Marx (2013) and Titman, Wei, and Xie (2004) that FF3F’s three factors miss much of the variation in average returns related to profitability and investment, Fama and French (2015) added profitability and investment factors to the three-factor model, thus constructing their five factor model.

Despite the many alternatives to the original CAPM, French and Fama’s (1993, 1996) three factor model (FF3F) and Carhart’s (1997) four factor model (C4F), these three models remained the dominant models in academia with the original CAPM still being the choice of practitioners and MBA classes. FF3F and C4F have taken center stage
in recent years despite the numerous models introduced since their inceptions as described above, including French and Fama’s (2015) five factor model which the authors claim performs better than the their three factor model. However, with little empirical work to test it, the five-factor model has yet to prove an improvement over FF3F OR C4F. Many scholars and practitioners still use the three-factor model partly for this reason and because methods seem to take some time before people adopt them. Accordingly, this study uses CAPM and the two dominant multi-factor models, Fama and French’s three factor model and Carhart’s four factor model, to evaluate institutional MFs and compare their performances to those of retails funds in the same category.

DATA AND METHODOLOGY

There are two well-known sources of mutual fund data; one is Morningstar Principia ® by Morningstar Inc., and the other by the Center for Research in Security Prices (CRSP®), a financial research center at the University Of Chicago Graduate School Of Business. Because CRSP creates and maintains historical U.S. databases for mutual funds and contains live funds as well as funds that have been discontinued, its US Mutual Fund Database is free of the survivorship bias documented by Brown, Goetzmann, Ibbotson, and Ross (1992) and Brown and Goetzmann (1995). It has one important disadvantage, however, in that it is subject to omission bias. Elton et al (2001) compared CRSP Survivor Bias Free U.S. Mutual Fund Database to Morningstar’s database. Although all mutual funds were listed in CRSP, return data were missing for many funds, and the researchers found that the CRSP database’s omission bias causes the same effects as survivorship bias. They also noted that returns in Morningstar and CRSP were a problem for older data and small funds. Morningstar mutual fund data (Morningstar, 2017) and profiles are publicly available online at the Yahoo! Finance website, where historical data can be downloaded free of charge. Although data on dead funds are not available (leading to survivorship-bias), the historical prices
are complete for the period they are available except for occasional missing daily prices. The profiles of the funds provide information on the current lead manager including his/her start date. Other relevant data include fund overview, fund summary, fund operations, fees, expenses, and investment information including minimum initial and subsequent investment amounts required.

Despite their impressive growth, institutional mutual funds are still a very small subset of mutual funds in terms of number of funds although growing interest by retirement plans has led to inception of numerous funds in this subset in recent years. Some funds, although labeled as Institutional Mutual Funds do (also) cater to the individual investors by selling shares through brokers; these funds typically have a much lower minimum initial and subsequent investment requirements, and therefore are not included in this study. All of the funds included invest primarily in domestic equities, have at least 10 years of historical prices available as of June 30th, 2017, have minimum initial investment requirement of $100,000 (most over $1 million) and are designated as “Institutional” even if this qualifier is not in the fund’s name. The fact that only funds with at least 10 years of life are chosen may lead to selection bias, and since dead funds are not included in the database, there is survivorship-bias in the data. Morningstar ratings and funds’ fees and expenses are not considered, although it is worth noting that institutional funds typically have significantly lower fees and expense ratios than retail funds. Since expense ratios lower a fund’s net return for the investor, the net return for an institutional fund will be higher than a retail fund that has the same (excess) return before expenses.

Although other classifications have been offered (see e.g. Brown and Goetzmann, 1997), the most commonly used fund classification tool is Morningstar style box which. The “style” of a fund or its manager refers to the approach fund managers use for security selection and is considered to have an influence on his/her performance. Morningstar’s styles classification allocates funds into different categories based on
fund’s investment methodology based on book-to-market value ratios (BM) and the market capitalization (size) of the companies in which it invests. It places funds in large-cap, medium-cap or small-cap categories based on the market capitalization of the companies whose equities the fund holds. Funds are also characterized as growth oriented, value oriented and blend based on book-to-market value ratios. The combination of orientation and market capitalization yields nine categories: Large-Cap Growth, Mid-Cap Growth and Small-Cap Growth; Large-Cap value, Mid-Cap Value and Small-Cap Value; Large-Cap Blend, Mid-Cap Blend and Small-Cap Blend. However, for the purposes of this study, funds are classified into categories based on investment styles consistent with the size and book-to-market ratios (BM) factors in the models used for evaluation. Rather than evaluating funds in nine separate categories, it makes sense to place funds into three categories based on size (large cap, midcap, small cap), and three other categories based on book-to-market value ratios (value, growth, blend). This is because the factors in the multi-factor models that will be used for evaluation are directly derived from differences in size and book-to-market value ratios. Accordingly, funds will be evaluated in six categories that are consistent with both Morningstar style box and the factors in the models used for evaluation.

If funds are divided into several categories, it makes little sense to compare the performances of all funds to a single benchmark, such as the S&P 500 index. Instead, funds should be compared to benchmarks compatible with their classification. Accordingly, this study uses different benchmarks, one for each category of funds based on market capitalization and book-to-market ratios. Russell indices are chosen because of their similarity in construction to Morningstar’s style categorization and for the stated purposes of the style indexes “to provide investors with accurate benchmarks for measuring [manager performance]”. For overall performance comparison between institutional and retail funds, a market wide Russell (Russell 3000) index is used as the benchmark. A list of the Russell indices used as a benchmark for each category is provided in Web Appendix B.
The data consisted of otherwise comparable, actively managed institutional and retail funds. Institutional funds were chosen first for each of the categories based on the criteria described above. As the reference group, retail funds were then selected to be otherwise comparable in investment objectives to the included institutional funds. As the first step of performance evaluation, monthly fund prices and benchmark indices were obtained from Yahoo! Finance website for the 10 year period between July 2007 and June 2017, and monthly returns for each fund ($R_{it}$) were computed in MS Excel as a monthly percentage change in the fund’s closing price (ACP) adjusted for dividends and splits. Monthly returns on the relevant benchmarks ($R_{mt}$) were calculated in the same manner. A fund’s risk premium in any month is the difference between the fund’s return and the monthly yield on 3-month Treasury Bill ($R_{ft}$) (as proxy for the risk-free rate) for the same period: $RP_{it} = R_{it} - R_{ft}$. Similarly, a benchmark’s risk premium is its excess return over risk-free rate: $RP_{mt} = R_{mt} - R_{ft}$. Using this notation, Jensen’s ‘alpha’ (based on the CAPM model) is expressed as

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{mt} - R_{ft}) + t_i,$$

where $\alpha_i$ is the fund’s beta as previously discussed, ($R_{mt} - R_{ft}$) is the market risk premium, $R_{it} - R_{ft}$ is the fund’s risk premium, and $i$ is the fund’s ‘alpha’. Based on CAPM, when a fund’s risk premium is regressed on the benchmark’s risk premium, the intercept coefficient (Jensen’s ‘alpha’) should not be significantly different from zero. If alpha is significantly positive (negative), the fund manager has outperformed (underperformed) the market.

However, there may be additional factors affecting a fund’s performance other than manager skill as demonstrated by researchers since the 1970s (e.g. Basu, 1977; Reinganum, 1981; Jegadeesh and Titman, 1992). Questioning the adequacy of a single index model in which it is not clear whether a positive alpha is the result of manager skill or other factors at play, Fama and French (1993) added two more factors based on size.
and book-to-market ratios, to develop their three factor model (FF3F) with the regression equation

$$R_{it} - R_{ft} = \alpha_i + \beta_{0i}(R_{mt} - R_{ft}) + \beta_{1i}\text{SMB} + \beta_{2i}\text{HML} + \varepsilon_{it} \quad (5)$$

where $R_{it} - R_{ft}$ is fund $i$’s risk premium, $R_{(mt)} - R_{(ft)}$ is the market risk premium at time $t$, SMB$_t$ is the difference in return between a small cap portfolio and a large cap portfolio at time $t$, and HML$_t$ is the difference in return at time $t$ between a portfolio containing 'value' stocks (with a high BM) and one consisting of 'growth' stocks (with low BM). In an effort to improve on FF3F, and capture Jegadeesh and Titman’s (1992) momentum anomaly and explain the “hot hands” effect, Carhart (1997) developed his four factor model (C4F) which can be modeled as

$$R_{it} - R_{ft} = \alpha_i + \beta_{0i}(R_{mt} - R_{ft}) + \beta_{1i}\text{SMB} + \beta_{2i}\text{HML} + \beta_{3i}\text{MOM} + \varepsilon_{it} \quad (6)$$

where all terms except MOM term are the same as FF3F. The MOM$_t$ factor intended to explain the “hot hands” phenomenon is the difference in return between a portfolio of past 12 months winners and a portfolio of past 12 month losers. The original notation for this term by Carhart (1997) was PRIYR$_t$, and represented one-year momentum in stocks at time $t$. Others have used the notation WML$_t$ (winners minus losers) to make it consistent with FF3F’s SMB$_t$ and HML$_t$ factors. Equations (5) and (6) are the regression equations used in the study for the Fama-French three-factor and Carhart four factor models, respectively. Detailed explanation of the factors and information on sources of data are provided in the Web Appendix. By using both CAPM and multi-factor models for fund performance evaluation, this study aims to find out (1) if any difference in performance uncovered by CAPM exists when compared with the results from the multifactor models and (2) if the multi-factor models do indeed offer better explanation of performance behavior than the single-factor CAPM based model.
RESULTS AND CONCLUSIONS

Institutional funds cater to a fast growing market of retirement plans and other institutional investors. A growing class of funds, they differ from retail funds in important aspects that could affect their relative performance such as close monitoring by the relatively more sophisticated institutional investors, discrete and larger fund flows, and much higher initial and subsequent investments. It is increasingly recognized that institutional investors (the principals) have the power to reduce the agency problem associated with investing through fund managers (the agents) and improve the governance of these companies (see e.g. Lucian et al, 2017). According to Fama and Jensen (1983), the ability of investors to redeem shares represents market governance that they argue reduces the need for other forms of governance in mutual funds. Their close monitoring and ability to withdraw funds more easily than retail investors allows the presumably more sophisticated institutional investors to exercise market governance and punish poorly performing managers by withdrawing assets under management (e.g., Del Guercio and Tkac, 2002); and Goyal and Wahal, 2008).

The market governance thus exercised by the sophisticated institutional investors should theoretically result in better performance by institutional funds than retail funds as suggested by Grossman and Stiglitz’s (1980) model of market efficiency, which predicts that informed investors earn higher returns than uninformed investors. To test this proposition, relative performances of institutional funds and retail funds were examined and compared first as two fund classes, then by separating each class into three categories by market capitalization (large-cap, midcap, small-cap) and three categories by book-to-market ratio (value, blend, growth) based on Morningstar style box. The models used for the evaluation were CAPM and two dominant multifactor models, Fama and French (1993) three factor model and Carhart’s (1997) four factor model, which are generally accepted to offer better explanation of performance behavior than the traditional, single-factor CAPM based models. All of the funds included
in the study invest primarily in U.S. domestic equities, have at least 10 years of historical prices available as of June 30, 2017, and institutional funds require a minimum initial investment of at least $100,000 (most over $1 million) and are designated as “institutional” by Morningstar. The set of hypotheses to test whether institutional funds outperform retail funds were:

\[ H_0: \alpha_i \leq \alpha_r \] (Institutional funds do not outperform retail funds).

\[ H_1: \alpha_i > \alpha_r \] (Institutional funds outperform retail funds).

where \( \alpha_i \) and \( \alpha_r \) are alphas (representing manager skill) for institutional and retail funds, respectively, in each of the models used. Web Appendix-A provides the results and the details for each model, including the single factor CAPM for comparison.

First, overall performance of fund groups was tested using all three models with results shown in Table A-1 in the Web Appendix. Testing with CAPM revealed that institutional funds as a class outperformed not only their benchmarks (“positive alpha”), but also retail funds (which also outperformed the benchmark) with the difference being statistically significant at the 10% level. However, the differences in alpha were not significant with the FF3F and C4F models where institutional funds outperformed retail funds, albeit with smaller margins. The slightly higher adjusted \( R^2 \) relative to the CAPM model confirms the expectation that multifactor models better explain mutual fund returns. A first glance at the market factor loading shows both types of funds have similar, positive and significant exposure in the CAPM model where fund coefficients are both very close to “1.0”. However, retail funds have less market exposure (as shown by the coefficient) in the multifactor models, probably because the omission of the size, BM and momentum factors in CAPM distorted market exposure. It is worth noting the reduction in the market risk was smaller for the institutional funds, with the coefficient still around “1.0”. The coefficient for market exposure was statistically significant at the 1% level for both classes in all models, but the difference was not significant in
any of the models. The size factor (SMB) produced positive and significant coefficients for both types in both multifactor models, although the loadings were smaller for institutional funds indicating they are relatively less exposed to small caps (or more exposed to large caps). The differences were significant in both models, indicating different exposures to market capitalization might explain the difference in performance rather than alpha (manager skill) the difference in which was significant in CAPM, but not in the multifactor models.

The book-to-market ratio factor (HML) adds less explanatory power as its loadings are not significant in the FF3M model, and only significant for the institutional funds in the C4F model. Although the differences were not significant, institutional funds had higher loadings indicating that they follow a relatively more value-oriented style. In fact, the coefficients were negative for retail funds in both models, indicating they pursue a more growth-oriented strategy. The examination of the momentum factor (MOM) in the C4F model reveals significant negative loadings for both groups, indicating contrarian investment strategies rather than riding the momentum. Overall, while CAPM results showed a statistically significant difference in alphas (manager skill) between the two fund groups, after controlling for market risk, size, book-to-market ratio and momentum, the difference was not significant. As such, we cannot reject the null hypothesis when tested with the multifactor models.

One critique of the FF3F and C4F models is that resulting alphas for value (growth) funds are systematically biased downward (upward) (Huij & Verbeek, 2009). Comparing institutional and retail fund performances within categories avoids any such biases. For this reason, and to test whether relative fund performance may be affected by investment strategies within each group, performance differences were examined within categories constructed based on Morningstar style box and the size factor (SMB) and the BM factor (HML) in the multifactor models used, resulting in Large Cap, Mid-Cap, Small Cap, Value, Blend and Growth categories. Estimation with CAPM revealed that both fund
groups outperformed their benchmarks in all size categories, but only institutional funds had a positive and significant alpha and only within the large cap category. The difference in alphas was also significant indicating that institutional funds outperformed retail funds in this category. Institutional funds slightly outperformed (underperformed) retail funds in small cap (midcap) categories, but the differences were not significant. Interestingly, both groups had significantly low market exposure in the small cap category while institutional funds had smaller market factor loadings than retail funds in all three of the size categories. Market factor loadings were significant albeit at different levels, and the difference was significant in the large cap category. Testing with FF3F and C4F models confirmed the significant differences in both the alphas and market exposure in the large cap category.

A glance at other factor loadings shows that SMB had less explanatory power within categories than overall evaluation, but the difference was significant in the large cap category in the C4F model, where the factor loading for institutional funds was negative indicating these funds had more exposure to large cap equities than the benchmark and the retail funds, which had a positive loading. The discrepancy in the sign and significance of the SMB loading between model for the large cap category can be explained by the explanatory power of momentum (MOM) for which the difference was significant. Overall, it can be argued that institutional funds significantly outperform retail funds in the large cap category, and the difference may stem from their relative exposure to small cap funds and the degree of their contrarian investment strategies.

Examination of the results within the book-to-market value categories revealed superior performance by both fund groups relative to their benchmark in the value category across all models. Although institutional funds underperformed across models in this category, the difference was not significant in any of them. Institutional fund performance was also significantly superior to its benchmark within the blend category across models, but the difference with retail fund performance was not
significant in any of the models. Finally, institutional funds outperformed (underperformed) the benchmark (retail funds) in the growth category in all models, but the differences were not significant.

This paper studied performance of institutional funds relative to appropriate benchmarks and to retail fund. The significantly different characteristics between the fund groups, most notably the close monitoring by the more informed institutional investors who exercise “market governance” through the use of sophisticated, risk adjusted performance measures to evaluate performance and determine fund flows, should theoretically result in better performance by institutional funds. This proposition was tested using the traditional CAPM model and two multifactor models, the Fama-French Three Factor model (FF3F) and the Carhart Four Factor (C4F) model. When compared as overall groups without separating into categories, institutional funds outperformed retail funds with a statistically significant difference in the CAPM model, but not in the multifactor models, suggesting that the difference may have stemmed not from different manager skill, but from their different exposures to market capitalization as evidenced by the statistically significant difference in their SMB loadings in both models. Consequently, the null hypothesis that institutional funds do not outperform retail funds could not be rejected. However, extending the comparison to within categories did reveal that institutional fund performance was superior only in the large cap category with a statistically significant difference. Fund performance was not significantly different in any other category although both groups outperformed their benchmarks in every size and BM category. It is worth noting that when considering the typically lower fees and expenses charged by institutional funds not considered in the study, the net returns of these funds are expected to be considerably higher than those of their retail counterparts.
Limitations and Further Research

One limitation of this study is that it focused on mutual funds investing primarily in U.S. equities. Further studies that include funds that invest in international equities could be conducted to see if these results hold for those funds. Another limitation has to do with the data: the Morningstar sourced data used in the study have survivorship bias. As pointed out by Brown et al. (1992), using a dataset that has survivorship bias leads to an overestimation of average performance. Although this should affect the performance of both groups, institutional funds and the retail funds, the size of the effect is unlikely to be equal. This study could be repeated with the alternative (CRSP) database to see if the same results would be obtained. However, CRSP database has omission bias that has been shown to have very similar effects on performance evaluation as omission bias (Elton et al, 2001). If a data free of survivorship bias and omission bias can be obtained in the future, the study should be repeated with the superior data. Additionally, this study compared performance evaluation using the two of the most commonly used multifactor models developed in the 1990s; it could be repeated with performance measures that have been developed since then, such as the Fama and French (2015) five-factor model.

REFERENCES


https://doi.org/10.1111/0022-1082.00410


https://doi.org/10.1016/S0167-2681(98)00093-6


Citation Information


Web Appendix

A web appendix for this paper is available at: https://dx.doi.org/10.1523/j.brcacadjb.2020.10.01.wa01