Measuring Economic Rents in the Mutual Fund Industry*

Jonathan B. Berk
Stanford University

Jules H. van Binsbergen
Northwestern University and Stanford University

May 9, 2011

PRELIMINARY DRAFT
Please do not quote or circulate without the authors' permission.

Abstract

We demonstrate that the skill to pick stocks or time the market exists amongst mutual fund managers and that this skill is persistent. Using this skill, the average mutual fund manager adds between $\$\frac{1}{2}$ million and \$1 million per month. The top 10% of managers add about \$5 million per month. About $\frac{1}{3}$ of managers add value while $\frac{2}{3}$ destroy value. There is also evidence of persistence amongst managers that destroy value, suggesting that these managers do not know their own ability.

*We could not have conducted this research without the help of the following research assistants for which we are grateful: Maxine Holland, Christine Kang, Fon Kulalert, Ian Linford, Binying Liu, Jin Ngai, Michael Nolop, William Vijverberg and Christina Zhu.

One of the first principles a student of economics learns is that agents only earn economic rents if they have a skill in short supply. As central as this principle is to microeconomics, surprisingly little empirical work has addressed the question of whether or not talent is actually rewarded, or, perhaps more interestingly, whether people without talent can earn rents. One notable exception is the literature on mutual fund managers. There is a very long literature in financial economics that has studied the question of whether stock picking or market timing talent exists and the overall conclusion of this literature is that it does not. In a recent paper on the subject, Fama and French (2010) conclude that the average mutual fund manager has no talent. Given that mutual fund managers are amongst the highest paid members of society, this conclusion represents a clear violation of the important principal of microeconomics relating skill to rents. It implies that it is possible to make economic rents without possessing a skill in short supply.

Given the importance of the question, the objective of this paper is to determine whether or not the conclusion in the mutual fund literature is right. We will find that it is not. Not only is the average mutual fund manager talented, but he uses his talents to add a considerable amount of value. In addition, there is evidence of a superstar effect[1] — a small number of managers appear to have extremely high skill and thus generate, relative to other managers, an enormous amount of value. In short, the mutual fund industry looks much like any other industry — there is strong evidence that skill exists and managerial compensation reflects the economic rents this skill generates.

The idea that active mutual fund managers lack skill has its roots in the very early days of modern financial economics. Indeed, the original papers that introduced the Efficient Market Hypothesis (Fama (1965, 1970)) cite as evidence in favor of the hypothesis the empirical results that, as a group, investors in active mutual funds underperform the market, and, more importantly, mutual fund performance is unpredictable. Both facts have been verified countless times since then and are routinely cited as evidence that managers lack skill. Although both results are robust characteristics of the data, as Berk and Green (2004) show, neither result implies the conclusion — both results are consistent with a model in which managers have skill. Although the performance of mutual funds might appear at first blush to be a measure of the skill of the manager, Berk and Green's observation is that under the relatively mild assumption that eventually the industry features decreasing returns to scale, the expected return of a mutual fund is actually determined by competitive forces in the capital markets, not managerial skill.

In this paper we correctly specify a measure of managerial skill — the average value (in dollar terms) the manager generates in fees for himself and in superior (or inferior) performance for his investors. Using this measure we show that the average manager adds between $\frac{1}{2}$ and 1 million dollars a year. In addition, the top 10%, on average, generate over \$60 million dollars a year. We show that our measure of skill his highly persistent when it is positive. There is also

[1] See Rosen (1981)

evidence that it is persistent when it is negative, implying that managers who destroy value are unaware of it.

The paper is organized as follows. In the next section we briefly review the literature. In Section 2 we derive our measure of skill and in the following section explain how we estimate it. We describe the data in Section 4. Our results are summarized in Section 5.

1 Literature Review

There is an extensive literature that studies the performance of active mutual fund managers, beginning with Fama (1965) and Jensen (1968). Reviewing that literature is beyond the scope of this paper. Suffice it to say the conclusion is that as investment vehicles, active funds underperform passive, and, on average mutual fund returns before fees show no evidence of outperformance. As we have already mentioned this evidence is taken to imply that active managers do not have the skills required to beat the market, and so in Burton Malkiel's words: the "study of mutual funds does not provide any reason to abandon a belief that securities markets are remarkably efficient." (Malkiel, 1995, p. 571)

Recently, Fama and French (2010) do concede that while the average manager lacks skill, there is some evidence of talent in the tails of the distribution of managers. But they largely dismiss this evidence because, based on their estimate of gross alpha, they conclude that the amount of skill is small. However, without taking into account how much money these skilled managers control, it is not clear how one can judge the economic magnitude of the observed skill. As we will see, when economic value added is calculated by multiplying the gross alpha by assets under management, a completely different picture emerges — the top 10% of managers are able to use their skill to add over \$60 million a year on average.

Researchers have also studied persistence in mutual fund performance. Using the return the fund makes for its investors, a number of papers (see Gruber (1996), Carhart (1997), Zheng (1999) and Bollen and Busse (2001)) have documented that performance is largely unpredictable.[2] Although this evidence was widely interpreted as evidence against managerial skill, capital flows into and out of mutual funds are related to lagged measures of excess returns (see Chevalier and Ellison (1997) or Sirri and Tufano (1998)). Consequently, researchers found this evidence puzzling. If returns are unpredictable and managers lack skill, why does capital chase performance? The answer to this puzzle is provided in Berk and Green (2004) who argue that flow of funds is symptomatic of the competition in capital markets that drives investor returns to their competitive levels. Returns are unpredictable not because managers necessarily lack skill, but because capital markets are competitive.

Despite the widespread belief that managers lack skill, there is in fact a literature in financial

[2]Some evidence of persistence does exist in low liquidity sectors or at shorter horizons, see, for example, Bollen and Busse (2005), Mamaysky, Spiegel, and Zhang (2008) or Berk and Tonks (2007).

economics that finds evidence of skill. One of the earliest papers is Grinblatt and Titman (1989), which documents positive gross alphas for small funds and growth funds. In a followup paper, Grinblatt and Titman (1993), these authors show that at least for a subset of mutual fund managers, stocks perform better when they are held by the managers than when they are not. Wermers (2000) finds that the stocks mutual funds hold outperform broad market indices, and Chen, Jegadeesh, and Wermers (2000) find that stocks managers' buy outperform stocks that they sell. Kosowski, Timmermann, Wermers, and White (2006) use a bootstrap analysis and find evidence, using gross and net alphas, suggesting that 10% of managers have skill. Kacperczyk, Sialm, and Zheng (2008) compare the actual performance of funds to the performance of the funds' beginning of quarter holdings and find that for the average fund, performance is indistinguishable, suggesting superior performance gross of fees and thus implying that the average manager adds value during the quarter. Cremers and Petajisto (2009) show that the amount a fund deviates from its benchmark is associated with better performance, and that this superior performance is persistent. Finally Cohen, Polk, and Silli (2010) and Jiang, Verbeek, and Wang (2011) show that this performance results from overweighting stocks that subsequently outperform the stocks that are underweighted.

There is also evidence suggesting where this skill eminates from. Coval and Moskowitz (2001) find that geography is important — funds that invest a greater proportion of their assets locally do better. Kacperczyk, Sialm, and Zheng (2005) find that funds that concentrate in industries do better than funds that don't. Baker, Litov, Wachter, and Wurgler (2010) show that, around earnings announcement dates, stocks that active managers purchase outperform stocks they sell and Shumway, Szefler, and Yuan (2009) produce evidence that superior performance is associated with beliefs that more closely predict future performance. Cohen, Frazzini, and Malloy (2007) find that portfolio managers place larger bets on firms they are connected to through their social network, and perform significantly better on these holdings relative to their non-connected holdings. These studies suggest that the superior performance documented in other studies in this literature is likely due to specialized knowledge and information.

Why the literature documenting the existence of managerial skill has not been more influential is not clear. At least part of the explanation is most likely attributable to the lack of any convincing evidence of the existence of significant rents attributable to this skill. Put succinctly, if this skill exists, where are the accompanying rents? Our object for the rest of this article is to provide this evidence.

2 Theory

In this section we use the theoretical contributions of Berk and Green (2004) to motivate our methodological approach to test the hypothesis that managers lack skill. Although widely used historically, the *net alpha* — the alpha earned by investors — does not measure managerial skill because it is determined by competition in the capital market. To understand why recall the intuition that Eugene Fama used to motivate the Efficient Market Hypothesis — just as the expected return of a firm does not reflect the quality of its management neither does the expected return of a mutual fund. Instead what the net alpha measures is the rationality and competitiveness of capital markets. A zero net alpha indicates that markets are competitive and investors rational. A positive net alpha implies that capital markets are not competitive, that the supply of capital is insufficient to compete away all the economic rents. A negative net alpha implies that investors are committing too much capital to active management — it is evidence of sub-optimality on the part of at least some investors. So the inference that should be drawn from the widely documented and robust empirical result that the average actively managed mutual fund underperforms the market is that some investors' behavior departs from the rational paradigm.[3]

Many have argued that the *gross alpha* — the alpha earned by the fund before management expenses are deducted — should be used to measure managerial skill. But this measure is also flawed. Just as the internal rate of return cannot be used to measure the value of an investment opportunity, the gross alpha cannot be used to measure the value of a manager. It measures the return the manager makes, not the value she adds. This point can be seen most starkly in the Berk and Green model. Because managers act optimally, they always commit the optimal amount of capital to active management and index the rest. This means that the amount of money under management does not affect the value added, which in turn implies that the fee the manager charges is irrelevant. A manager can either manage a large fund for a small fee or a small fund with a large fee — the fee merely determines the amount of money he chooses to index. In the limit the manager can set an infinitesimal fee, which would imply that the gross return and net return were indistinguishable. But we already argued that the net return is not a good measure of managerial talent, hence, neither is the gross return.

To correctly measure the skill of the manager, one has to measure the dollar value of what the manager adds. The expected excess return (that is, the expected return in excess of the risk free rate) generated by the i'th manager, $E_t[R_{it+1}]$, can be written as

$$E_t[R_{it+1}] = \alpha_{it} + f_{it} + E_t[\tilde{R}_{it+1}] \tag{1}$$

where $E_t[\tilde{R}_{it+1}]$ is the risk adjustment, that is, the expected return in excess of the risk free rate that the fund is expected to make given its level of risk (or the expected return on a well diversified portfolio of equivalent risk), α_{it} is what the manager adds or subtracts, and f_{it} are the fees the manager or mutual fund company charges. Because the manager's expected return

[3]Glode, Hollifield, Kacperczyk, and Kogan (2009) document that this irrationality varies by market conditions. Why this irrational behavior persists in the mutual fund industry is a largely unexplored empirical question. To our mind this empirical regularity is amoungst strongest evidence the profession has uncovered against the fully rational paradigm.

is unobservable, we must rewrite (1) as a function of the managers realized return:

$$R_{it+1} = \alpha_{it} + f_{it} + E_t[\tilde{R}_{it+1}] + \epsilon_{it+1} \tag{2}$$

with $E_t[\epsilon_{it+1}] = 0$. If the i'th manager at time t has total assets under management of q_{it}, her total value added (or destroyed), V_{it} is

$$V_{it} \equiv q_{it} f_{it} + q_{it}\alpha_{it} = q_{it}(\alpha_{it} + f_{it}) = \text{Size} \times \text{Gross Alpha}. \tag{3}$$

The value added consists of two parts — the part the manager takes home with him as compensation (the dollar value of all fees charged), which is necessarily positive, plus any value he provides (or extracts from) investors, which can be either positive or negative.

Before we turn to how we actually estimate V_{it} it is worth first considering what the main hypotheses in the literature imply about the the value of this measure of skill. The most widely accepted hypothesis and the one considered in Fama and French (2010) is that managers have no skill. In this case the fees they charge come out of investors pockets, that is, $\alpha_{it} = -f_{it}$, implying that $E_t[V_{it}] = 0$. We will test two forms of this hypothesis: (1) what we term the *strong form*, originally put forward by Fama in his Efficient Market papers, that no manager has skill, that is, $V_{it} = 0$ for any i and t and (2) the *weak form*, that, on average, managers do not have skill, $\sum_i E_t[V_{it}]/N_t = 0$, where N_t is the total number of managers at time t.

The second hypothesis we consider is that managers have skill (which is in short supply). Because of competition in capital markets, investors do not benefit from this skill and so managers derive the full benefit of the economic rents they generate from their skill. If investors are fully rational and can observe skill perfectly, then these assumptions imply that $\alpha_{it} = 0$ and because $f_{it} > 0$, (3) implies $V_{it} > 0$, for every i and t. When investors cannot observe skill perfectly the extent to which an individual manager actually adds value depends on the ability of investors to differentiate talented managers from charlatans. If we recognize that managerial skill is difficult to measure, then we would expect unskilled managers to take advantage of this uncertainty, and so one would expect to observe the presence of charlatans — managers who charge a fee but have no skill. Thus when skill cannot be perfectly observed, it is possible that for some managers $V_{it} = 0$. However, even when skill is not perfectly observable, because investors are rational, every manager must still add value in expectation, $E_t[V_{it}] > 0$. Hence, under this hypothesis $\sum_i V_{it}/N_t > 0$ — the average manager must add value.

A third possibility is a combination of the above two hypotheses — some managers add value and others exploit dumb investors. If managers know their own ability, they would never choose to destroy value because they can always charge their fee and index their portfolios. This observation implies is that $-f_{it} \le \alpha_{it} \le 0$, which implies that $V_{it} \ge 0$. If at least some managers add value, then the average manager must add value: $\sum_i V_{it}/N_t > 0$. What differentiates this hypothesis from the previous one is the prediction on alpha. The prior hypothesis predicts that

$E_t[\alpha_{it}] = 0$ for all i and t, whereas this hypothesis predicts that if managers are able to exploit dumb investors, then for some i, $E_t[\alpha_{it}] < 0$. Thus, this hypothesis also predicts $\sum_i \alpha_{it}/N_t < 0$.

Finally, a fourth possibility exists — managers might not know their own ability. In this case some managers might actually destroy value, by for example, actively trading with better informed agents. For these kinds of managers $\alpha_{it} < -f_{it}$, so $V_{it} < 0$. Whether such managers make up the majority of managers depends on investor rationality. If investors are perfectly rational, then $E_t[\alpha_{it}] = 0$, which by the argument outlined above necessarily implies that $\sum_i V_{it}/N_t > 0$. But if investors also make mistakes, it is possible for the whole industry to destroy value in agreggate. To sum up, to observe the condition that the average manager destroys value, $\sum_i V_{it}/N_t < 0$, we need two conditions: managers cannot know their own ability and some fraction of investors must be irrational.

It is worth noting that some have claimed, based on Sharpe (1991), that in a fully rational general equilibrium it is impossible for the average manager to add value, that is, $\sum_i E_t[V_{it}]/N_t > 0$ is impossible. In fact this argument has two flaws. To understand the flaws, it is worth quickly reviewing Sharpe's original argument. Sharpe argued as follows. He divided all investors into two sets: people who hold the market portfolio, who he called "passive" investors and the rest, who he called "active" investors. Because market clearing requires that the sum of active and passive investors' portfolios is the market portfolio, the sum of just active investors' portfolios must also be the market portfolio. This observation immediately implies that the gross alpha of the average active investor must be zero, or $\sum_i V_{it}/N_t = 0$. As convincing as the argument appears to be, it cannot be used to conclude that the average active mutual fund manager cannot add value. In his definition of "active" investors, Sharpe includes *any* investor not holding the market, not just active mutual fund managers. If active individual investors exist, then as a group active mutual fund managers could provide an positive alpha by making trading profits from individual investors who make a negative alpha. Of course, as a group individual investors are better off investing in the market, which leaves open the question of why these individuals are actively trading.

Perhaps more surprisingly to some, Sharpe's argument does not rule out the possibility that the average active manager adds value even if all investors are assumed to be fully rational, so that all active investors can expect to outperform passive investors. What Sharpe's argument ignores is that even a passive investor must trade twice, once to get into the passive position and once to get out of the position. If we assume that active investors are better informed than passive, then whenever these liquidity trades are made with an active investor, in expectation, the passive investor must lose and the active must gain. Hence, the expected return to active investors must exceed the return to passive investors, which necessarily implies that it is possible for the average active investor to have a positive gross alpha. Thus Sharpe's insight, by itself, does not rule out the possibility that the second and third hypotheses, that the average manager has skill, holds.

3 Econometrics

We measure the value added by a manager in two steps. We first identify a risk adjustment, that is, a well diversified portfolio of equivalent risk. The realized excess return on this portfolio is

$$\tilde{R}_{it+1} \equiv E_t[\tilde{R}_{it+1}] + \nu_{it+1} \tag{4}$$

where $E_t[\nu_{it+1}] = 0$. Substituting (4) into (2) provides

$$\begin{aligned} R_{it+1} &= \alpha_{it} + f_{it} + \tilde{R}_{it+1} - \nu_{it+1} + \epsilon_{it+1} && (5) \\ &= \alpha_{it} + f_{it} + \tilde{R}_{it+1} + \xi_{it+1} && (6) \end{aligned}$$

with $E_t[\xi_{it+1}] = E_t[\epsilon_{it+1}] - E_t[\nu_{it+1}] = 0$. Rearranging terms provides us with our empirical measure of gross alpha, $\hat{\alpha}^G_{it}$:

$$\hat{\alpha}^G_{it} \equiv \alpha_{it} + f_{it} + \xi_{it+1} = R_{it+1} - \tilde{R}_{it+1} \tag{7}$$

We will use two approaches to identify the risk adjustment. The first approach is the standard measure of risk used in the mutual fund literature — the Fama-French-Carhart (FFC) factor specification. In this case

$$\tilde{R}_{it+1} = \beta^{mkt}_i \text{MKT}_t + \beta^{sml}_i \text{SML}_t + \beta^{hml}_i \text{HML}_t + \beta^{umd}_i \text{UMD}_t$$

where $\text{MKT}_t, \text{SML}_t, \text{HML}_t$ and UMD_t are the realizations of the four factors and $\beta^{\cdot}_i$ is the sensitivity to the factor of the i'th mutual fund. Although standard practice, this approach has a the drawback that no theoretical reason exists for why these factors should measure systematic risk in the economy. Fama and French (2010) recognize this limitation but argue that one can interpret the factors as returns on passive benchmark portfolios. There are two concerns with this interpretation. First, to interpret the portfolios this way, it is important that investors have the opportunity to invest in these portfolios. But because these portfolios only became widely known and available to investors in the mid-90's and our sample starts in the early 60's, for a significant fraction of the sample period it would have been impossible for most investors to invest in these porfolios. Second, even for investors who could invest in the portfolios, they would have had to pay transactions costs to do so. By interpreting these portfolios as benchmark portfolios we are essentially benchmarking a return that includes transaction costs to a return that does not.

In light of the issues with the FFC specification, we will also use another approach motivated by the interpretation of the FFC factor portfolios as benchmark portfolios. But instead of using the FCC portfolios, we use the set of index funds offered by Vanguard. This approach has a number of advantages over the FFC specification. First, by using the Vanguard funds as a benchmark, we can be certain the investors had the opportunity to actually invest in the funds at the time. Secondly, the returns of these funds necessarily include transaction costs. If Δ_t^j is the excess return of the j'th Vangaurd index fund at time t, then in this case

$$\tilde{R}_{it+1} = \sum_{j=1}^{n(t)} \beta_i^j \Delta_t^j$$

where $n(t)$ is the total number of index funds offered by Vanguard at time t and β_i^j is the sensitivity of the i'th mutual fund to the j'th Vanguard index fund. In principle, any investor can replicate $\tilde{R}_{it+1}$ by borrowing $\sum_{j=1}^{n(t)} \beta_i^j$ at time t and investing β_i^j in each of Vanguard's $n(t)$ funds. Notice that there is no requirement that each index fund be in existence throughout the life of the i'th actively managed mutual fund.

The advantage of the second methodology is that as the market for asset management services evolves, strategies that were once only available through active management become widely known and are offered as passive alternatives. In the early days when such strategies were not widely know, our measure gives the active manager credit for the skill of finding and using these strategies. Later, once the strategies become widely known, so that no skill is required to use them, the manager is no longer given credit for using these strategies. Thus, by using Vanguard index funds as benchmarks, we naturally take into account the dynamic evolution of active management.

The next step uses our estimate of gross alpha to construct an estimate of the value added at time t, $\hat{V}_{it}$, by multiplying the size of the fund q_{it} by the estimate of gross alpha:

$$\hat{V}_{it} \equiv q_{it} \hat{\alpha}_{it}^G.$$

Note that this estimate is an unbaised measure of V_{it} because (using (7)),

$$E_t[\hat{V}_{it}] = q_{it}(\alpha_{it} + f_{it} + E_t[\xi_{it+1}]) = q_{it}(\alpha_{it} + f_{it}) = V_{it}.$$

The average value added for a fund that exists for T_i periods is then

$$\bar{V}_i = \sum_{t=1}^{T_i} \frac{\hat{V}_{it}}{T_i}.$$

4 Data

Our main source of data is the CRSP survivorship bias free database of mutual fund data first compiled in Carhart (1997). The dataset spans the period from January 1962 to December 2010. Although this dataset has been used extensively, it still has a number of important shortcomings that we needed to address in order to complete our study. As a result we undertook an extensive project to address these shortcomings, the details of which are described in a 17 page appendix to this paper. Here we will briefly summarize what we did.

Even a casual perusal of the returns on CRSP is enough to reveal that some of the reported returns are suspect. Because part of our objective is to identify highly skilled managers, misreported returns, even if random, are of concern. Hence we procured an additional data from Morningstar. Each month Morningstar sends a complete updated database to its clients. The monthly update is intended to completely replace the previous update. We purchased every update from January 1995 through December 2010, and constructed a single database by combining all the updates. One major advantage of this database is that it is guaranteed to be survivorship free. Morningstar adds a new fund or removes an old fund in each new monthly update. By definition, it cannot change an old update because its clients already have that data. So we are guaranteed that in each month whatever data we have was the actual data available to Morningstar's clients at that time.

We then compared the returns reported on CRSP to what was reported on Morningstar. Somewhat surprisingly, 3.3% of return observations differed. Even if we restrict attention to returns that differ by more than 10 b.p., 1.3% of the data is inconsistent. To determine which database is correct we used dividend and net asset value (NAV) information reported on the two databases to compute the return. In cases in which on one database the reported return is inconsistent with the computed return but the other database was consistent, we used the consistent database return. If both databases were internally consistent, but differed from each other, but within 6 months one data base was internally inconsistent, we used the database that was internally consistent throughout. Finally, we manually checked all remaining unresolved discrepancies that differed by more than 20 b.p. on Bloomberg and used that to resolve the discrepancy. All told we were able to correct about two thirds of the inconsistent returns. In all remaining cases we used the return reported on CRSP.

The discrepancies between what Morningstar and CRSP report for total assets under management (AUM) are even worse than for the return data. Even allowing for rounding errors (by ignoring AUMs that differ by less that 2 b.p.), fully 16% of the data differs across the two databases. But casual observation reveals that much of this discrepancy appears to derive from Morningstar often lagging CRSP in updating AUM. When both database reported numbers we decided to use the numbers reported on CRSP with one important exception. If the number reported on CRSP changed by more than $8\times$ (we observed a number of cases where the CRSP number is off by a fixed number of decimal places) and within a few months the change was reversed in order of magnitude, and, in addition, this change was not observed on Morningstar, we used the value reported on Morningstar. Unfortunately both databases contained significant numbers of missing AUM observations. Even after we used both databases as a source of information, 17.2% of the data was missing. In these cases we filled any missing observation by the most recent observation in the past. Finally we adjusted all AUM numbers by inflation by expressing all numbers in January 1, 2000 dollars.

For our purposes the amount of missing data is a major problem. To compute the gross return, expense rations are needed and over 40% of expense ratios are missing. Because expense ratios are actually reported annually by funds, we were able to fill in about 70% of these missing values by extending any reported observation during a year to the entire fiscal year of the fund and combining the information reported on Morningstar and CRSP. We then manually looked up the remaining missing values on EDGAR, the SEC web site. At the end of this process we were left with only 1.6% of observations missing, which we elected to drop.

Both databases report data for active and passively managed funds. CRSP does not provide any way to discriminate between the funds. Morningstar provides this information, but the accuracy of their classification is suspect and we only have the information after 1995 when the Morningstar database begins. We therefore used the following algorithm to identify the passively managed funds. We first generate a list of common phrases that appear in fund names identified by Morningstar as index funds. We then compile a list of funds with these common phrases but not labelled as index funds by Morningstar and compile a second list of common phrases from these funds. We then manually checked the original prospectuses of any fund that contained a word from the first list but was not identified as an index fund at any point in its life by Morningstar (including the pre 1995 period) or was identified as an index fund at some point in its life by Morningstar but nevertheless contained a phrase in second list. Funds that were not tracked by Morningstar (e.g., only existed prior to 1995) that contained a word from the first list were also manually checked. Finally, we also manually checked cases in which fund names satisfied any of these criteria in some periods but not in others even when the Morningstar classification was consistent with our name classification to verify that indeed the fund had switched from active to passive or vis versa. We reclassified 14 funds using this algorithm.

It is important to identify subclasses of mutual funds because both databases report subclasses a separate fund. In most cases the only difference amongst subclasses are the amount of expenses charged to investors, so simply including them as separate funds would artificially increase the statistical significance of any identified effect. For funds that appear in the CRSP data base, identifying subclasses is a relatively easy process — CRSP provides a separator in the fund name (either a ":" or a "/"). Information after the separator denotes a subclass. Unfortunately, Morningstar does not provide this information, so for mutual funds that only appear on the Morningstar database we used the last word in the fund name to identify the subclass (the details of how we did this are in the appendix). Once identified we aggregated all subclasses into a single fund.

We dropped all money market funds[4] and any fund observations before the fund's (inflation adjusted) AUM reached $5 million. In the end we were left with 4464 equity funds.[5] This sample

[4]We identified a money market fund is a fund that on average held more than 20% of assets in cash.

[5]We classed a fund as an equity fund if, on average, it held more than 50% of assets in stocks and the rest we

Fund Name	Ticker	Asset Class	Inception Date
500 Index	VFINX	Large-Cap Blend	08/31/1976
Growth Index	VIGRX	Large-Cap Growth	11/02/1992
Small-Cap Index	NAESX	Small-Cap Blend	10/03/1960
Small-Cap Value Index	VISVX	Small-Cap Value	05/21/1998

Table 1: **Benchmark Vanguard Index Funds**

is considerably larger than comparable samples used by other researchers. There are a number of reasons for this. Firstly, we do not restrict attention to U.S. equity funds. Although this is standard practice in the mutual fund literature we could not justify imposing this selection criteria. Clearly, managerial skill, if it exists, could potentially be used to pick non-U.S. stocks. Second, the Morningstar database contains funds not reported on CRSP. Third, we use the longest possible sample length available.

We picked four Vanguard index funds to use as benchmark funds – see Table 1. Because we explicitly want to let the benchmark number of funds change as the mutual fund space matures, we are limited in the number of funds we can employ because each time a new fund enters the database the number of regressors increases by the number of index benchmarks already in existence. We picked these four funds because they were the funds with the longest history that also span the same space as the FFC factors (small vrs. large and growth vrs. value).[6] To limit the number of regressors, we start the sample in 1976 with two benchmark funds (S&P500 and Small-Cap) increase the number of benchmark funds in 1992 to three (when Vanguard introduced the Growth Index fund) and to four in 1998 (when the Small-Cap Value index fund was introduced).

5 Results

We begin by estimating $\bar{V}_i$ for every equity fund in our sample. Table 2 provides the results. Notice that we can reject the Null Hypothesis (both weak form and strong form). Using the FCC factor specification as the risk adjustment, we estimate that the average manager adds about $480,000 per month. The standard error of this average is just 0.12, implying at t-statistic of 4, so we can easily reject the hypothesis that the mean is zero at the 99% confidence level. There is also enormous variation across funds. The least skilled manager amongst the top 1% of managers is able to generate $13.36 million per month, or $160 million annually on average.

classed as bond funds.

[6]The complete list of all Vanguard's Index funds can be found here:
https://personal.vanguard.com/us/funds/vanguard/all?reset=true&mgmt=i.

Even the least skilled manager amongst the top 10% of managers generates over a million dollars a month on average. But there is also evidence a significant number of managers who actually destroy value, that is, we see negative values of $\bar{V}_i$. The bottom 1% destroy at least \$6 millon per month and 10th percentile managers destroy just under a million dollars per month. This is evidence in favor of the fourth hypothesis we put forward — a subset of managers appear to not know their own ability. Of course, better managers manage larger funds, which explains why even though the median manager actually destroys value, the average manager adds value.

Mean	0.48	0.97
Standard Error of the Mean	0.12	0.14
Standard Deviation	8.21	9.61
1st Percentile	-5.90	-3.88
5th Percentile	-1.88	-1.04
10th Percentile	-0.91	-0.52
50th Percentile	-0.02	0.02
90th Percentile	1.17	2.01
95th Percentile	3.22	4.32
99th Percentile	13.36	18.62
No of Funds	4464	4406

Table 2: **Summary Statistics of value added $\bar{V}_i$ for Equity Mutual fund:** The numbers are reported in \$ Millions per month.

When we use the Vanguard index funds as benchmark portfolios the estimates are higher. In this case we estimate that the average manager adds almost a million dollars a month, implying at t-statistic of 6.7. Measured this way, the 99th percentile manager adds more than \$200 million per year. There are three reasons to expect the Vanguard based estimates to be higher than the FFC based estimates. First, Vanguard index funds include transaction costs and so the returns of these funds will be lower than non-marketed benchmarks like the FFC factors. Second,in the earlier years in our sample managers are not benchmarked against strategies that were not widely known about at the time. Finally, it is possible that the Vanguard funds do not span risk factors that the FFC factors do.

Although not the focus of this paper, it is interesting to note that the estimate of the average *net alpha* is quite sensitive to the measure of risk. When the FFC factor specification is used to adjust for risk then, as is evidenced in Table 3, we reproduce that result that active funds underperform passive to the tune of 9 b.p./month. However, when we instead use the Vanguard funds as benchmark portfolios then we find that the underperformance drops to barely

	FFC Risk Measure	Vanguard Benchmark
Mean	-9.12	-1.79
Standard Error of the Mean	0.67	0.68
Weighted Mean	6.53	13.8
Standard Error of the Weighted Mean	0.45	0.46

Table 3: **Net Alpha (in b.p./month):** We first estimate the net alpha of each fund by recording the intercept term of net excess returns regressed on the risk adjustment (indicated in the column head). The upper panel reports the average net alpha across all fund and the lower panel reports the average weighted by the average size of the fund.

2 b.p./month. More interestingly, in both cases, when we compute the average weighted by the average size of the fund, active funds appear to outperform passive. So if there is any evidence of investor irrationality, it is in the small funds.

One thing the distribution of summary statistics in Table 2 hides is that the $\bar{V}_i$'s are estimate with very different precision. Better funds are likely to be around much longer, and hence the estimate of their performance is likely to be more precise than newer funds or funds that did not last as long, both of which are associated with lower levels of skill. To address this issue we measure the precision of the estimate of value added by calculating its time series t-statistic ($\bar{V}_i$ divided by its time series standard error). We then sort funds by this t-statistic into 11 quantiles and plot the average value added in each quantile in Figure 1.

The extreme quantiles are the portfolios that were have the most confidence in our estimate of value. As the figure shows, skill is concentrated in the largest funds — the funds for which we have the most confidence that value is added have well over a billion dollars under management on average. They, add, on average, about \$5.5 million (\$6 million) dollars per month under the FCC factor specification (Vanguard benchmark). This is more than twice the value added of the next lowest quantile, indeed this pattern continues for all quantiles with positive V estimates for the FCC specification — each quantile is more than twice the previous one (see Table 4). There is also a large asymmetry. The funds that we have the most confidence that value is being destroyed destroy just over \$1 million, less than a fifth of what the funds in the top quantile add. This pattern continues for all the extreme quantiles, in each case the value added by the upper quantile is significantly larger than the value destroyed in the corresponding lower quantile. The main difference between the FCC and Vanguard risk adjustment is that, for the reasons outlined above, the numbers are larger using the Vanguard benchmark. Overall, Figure 1 tells the following story — most managers destroy a little bit, but about 1/3 of managers add considerable value, and because they control much larger funds (as evidenced by average AUM in Table 4), on average, managers add value.

	Low	2	3	4	5	6	7	8	9	10	High
					Panel A: FCC Risk Adjustment						
Value Added	-1.09	-1.21	-0.78	-0.64	-0.63	-0.32	0.13	0.62	1.43	2.31	5.42
Net Alpha	-57	-31	-21	-16	-6	-5	2	4	17	20	36
AUM	174	363	279	347	533	910	532	853	689	973	1390
Age	56	87	109	125	152	142	156	139	121	99	68
					Panel B: Vanguard Benchmark						
Value Added	-1.12	-0.88	-0.67	-0.33	-0.11	0.33	0.83	1.24	1.89	3.47	6.01
Net Alpha	-48	-29	-19	-7	-1	2	8	18	23	25	34
AUM	211	237	292	360	407	580	750	721	703	1265	1629
Age	51	84	106	124	132	144	148	133	121	103	74

Table 4: **Characteristics of 11 Quantiles Sorted on t-statistics:** The table reports average values for the funds in each quantile. *Value Added* is average $\bar{V}_i$ (in \$ millions/month), *Compensation* is the average of $\sum_{t=\tau_i}^{T_i} q_{it} f_{it}/T_i$, (in \$ millions/month), *Fees* (in %/annum) is the average $\sum_{t=\tau_i}^{T_i} f_{it}/T_i$, *Net Alpha* (in b.p./month) is the value weighted average intercept term of net returns regressed on the risk measure, AUM is average assets under management (in \$ millions) expressed in 2000 dollars, *Age* is the average number of monthly observations.

A central tenant of the literature on skill in mutual fund management is that if indeed managers have skill, then performance should be persistent. In the past researchers have tested this hypothesis using net alpha, and because they found little evidence of persistence, concluded that skill did not exist. As we have already argued, such a conclusion is premature because net alpha does not measure skill. Yet, quite clearly, the idea is right — if managers have skill, and skill is measured correctly, then the measure should be persistent.

To test for persistence in our skill measure, in each period t we sort managers into 11 quantiles based on the t-statistic of the estimate of $\bar{V}_i$ using data available up to that point. Because of the difficulty of getting a reliable estimate of α_{it}^G we only include a fund after it has at least 2 years of data. We then compute the average value of V_{it+1} for all funds in the quantile. There two ways to compute V_{it+1} depending on how $\tilde{R}_{it}$ is determined. If adopt the view that $\tilde{R}_{it}$ is an asset pricing model (that is, the interpretation of the FFC factors) then we should use all information up to and including information at $t+1$ to estimate the model. Thus, in this interpretation, to calculate $\alpha_{it+1}^G = R_{it+1} - \tilde{R}_{it+1}$ we re-estimate the beta estimates using the returns up to and including the return in $t+1$ to compute $\tilde{R}_{it+1}$. However, if we adopt the view that $\tilde{R}_{it}$ simply reflects the return of a passive investment alternative of similar risk (the justification for the Vanguard benchmark funds), then to invest in this passive alternative, investors cannot use information available at $t+1$. Hence in this case we do not re-estimate the betas, but compute $\tilde{R}_{it+1}$ using the regression estimates at time t. We repeat this process for all dates in the sample

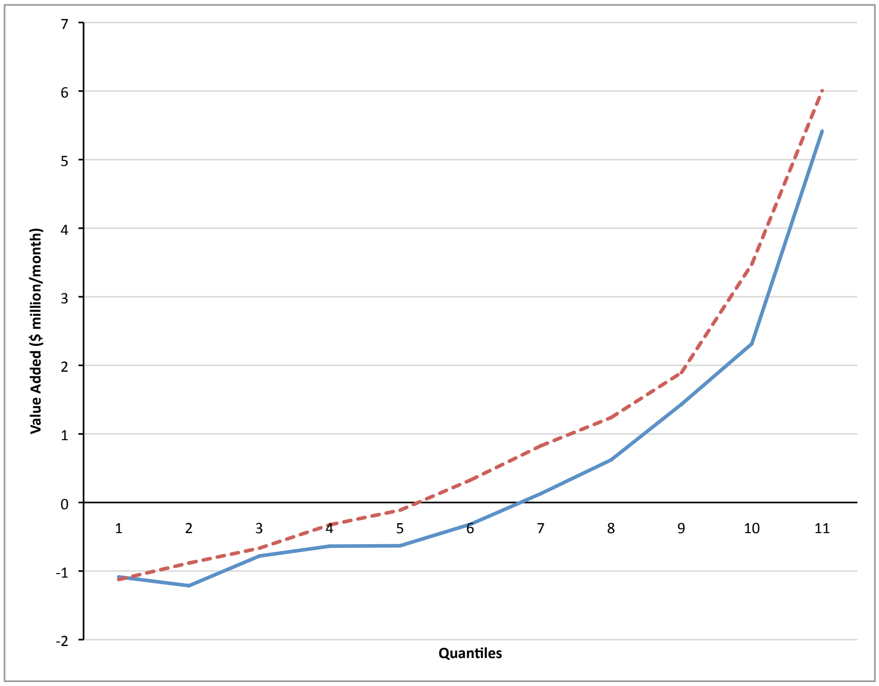

Figure 1: Average value added for 11 quantiles sorted on t-statistics.
The graph displays the average value added for 11 quantiles. Funds are sorted into 11 bins based on the t-statistic of the intercept of the return adjustment regression. The graphs plots for each bin the average value added $\bar{V}_i$ of all funds within that bin. The solid line uses the FFC factor specification as the risk adjustment, and the dashed line uses Vanguard index funds as benchmark portfolios.

and Figure 2 plots the average for each quantile.

Figure 2 shows that when managerial skill exists, it is persistent. The funds for which we have the most confidence that they add value (the top quantiles), continue to add value. For example, the probability under the Null that we would observe the top 3 quantiles ordered correctly (as we do in Figure 2) is $\frac{1}{11} \times \frac{1}{10} \times \frac{1}{9} = 0.1\%$, so the result has a p-value of 0.1%. Even more powerful tests that use the ordering in the time series reject at much higher levels of significance. For example, the number of times the top three quantiles are subsequently the top three quantiles (32 when the expected number is 3.3) or the subsequent top 3 quantiles are ordered correctly (7 when the expected number is 0.55) produce p-values of essentially zero.

There is also weaker evidence in favor of the fourth hypothesis, that managers do not know their own ability. Note that funds that we have the most confidence that they destroy value, appear to continue to destroy value.[7] Interestingly, funds in the bottom two quantiles go on to destroy less value than funds in the third quantile, which is likely due to the fact that the lowest

[7]The probability under the Null of observing the 6 quantiles with negative value added in Table 4 having negative value added in Figure 2 (as we do) is $\frac{1}{2^6} = 1.5\%$.

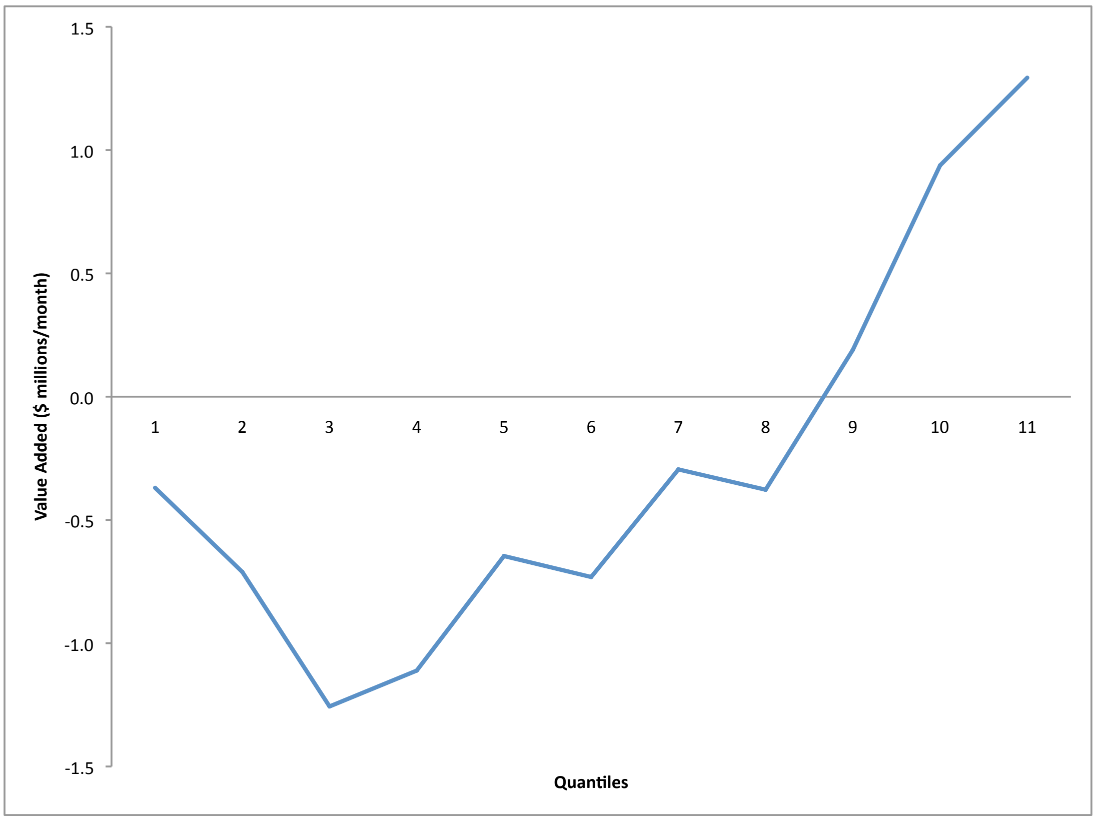

Figure 2: Persistence in Value Added: The graph displays the average one month out of sample performance for funds sorted on the t-statistic of value added, $\bar{V}_i$ computed up to that point using the FFC factor specification.

	Low	2	3	4	5	6	7	8	9	10	High
				Panel A: FCC Risk Adjustment							
Compensation	0.17	0.32	0.26	0.33	0.48	0.71	0.48	0.64	0.57	0.77	1.00
Fees	1.39	1.38	1.36	1.38	1.25	1.26	1.32	1.31	1.34	1.26	1.25
				Panel B: Vanguard Benchmark							
Compensation	0.21	0.24	0.28	0.34	0.38	0.50	0.66	0.60	0.57	0.90	1.15
Fees	1.46	1.46	1.41	1.30	1.30	1.28	1.32	1.29	1.27	1.26	1.22

Table 5: **Characteristics of 11 Quantiles Sorted on t-statistics:** The table reports average values for the funds in each quantile. *Value Added* is average $\bar{V}_i$ (in \$ millions/month), *Compensation* is the average of $\sum_{t=\tau_i}^{T_i} q_{it} f_{it}/T_i$, (in \$ millions/month), *Fees* (in %/annum) is the average $\sum_{t=\tau_i}^{T_i} f_{it}/T_i$, *Net Alpha* (in b.p./month) is the value weighted average intercept term of net returns regressed on the risk measure, AUM is average assets under management (in \$ millions) expressed in 2000 dollars, *Age* is the average number of monthly observations.

quantile funds are more likely to be shut down and hence drop from the sample. Interestingly, the persistence of value destruction in the third quantile rivals the value added of the top quantile. Presumably this quantile contains managers who destroy value but this destruction is not as obvious as it is in the quartiles with more negative t-statistics and so investors do not withdraw their funds as readily.

In light of the evidence supporting the hypothesis that managers have skill, we can address the question posed in the introduction — do the managers capture the rents associated with this skill? Figure 3 plots (and Table 5 reports) the average compensation earned by managers in each quantile in Figure 1. The results are striking. There is close to a monotonic relation between skill and compensation, especially in the extreme quantiles where we have the most confidence of our estimates of value. Better managers earn more. Even more interesting is how the compensation is determined. Figure 4 plots (and Table 5 reports) the average percentage fee charged in each quantile. There is very little variation, if anything, better managers charge slightly *lower* fees. Thus compensation differences are not determined by the manager or the mutual fund company. They are determined by the amount of money under management — that is, investors determine the managerial compensation differential, confirming a central insight in Berk and Green (2004).

Although there is a strong relation between skill and pay, it is interesting that in the upper quantiles, compensation estimates are considerably lower than the estimates of value added. Taken at face value, this suggests that the best managers leave some rents to investors. However, some degree of caution is in order. Because of how we form the quantiles, in Figure 1 we are overstating the value added in the top quartile and understating it in the bottom quartile.

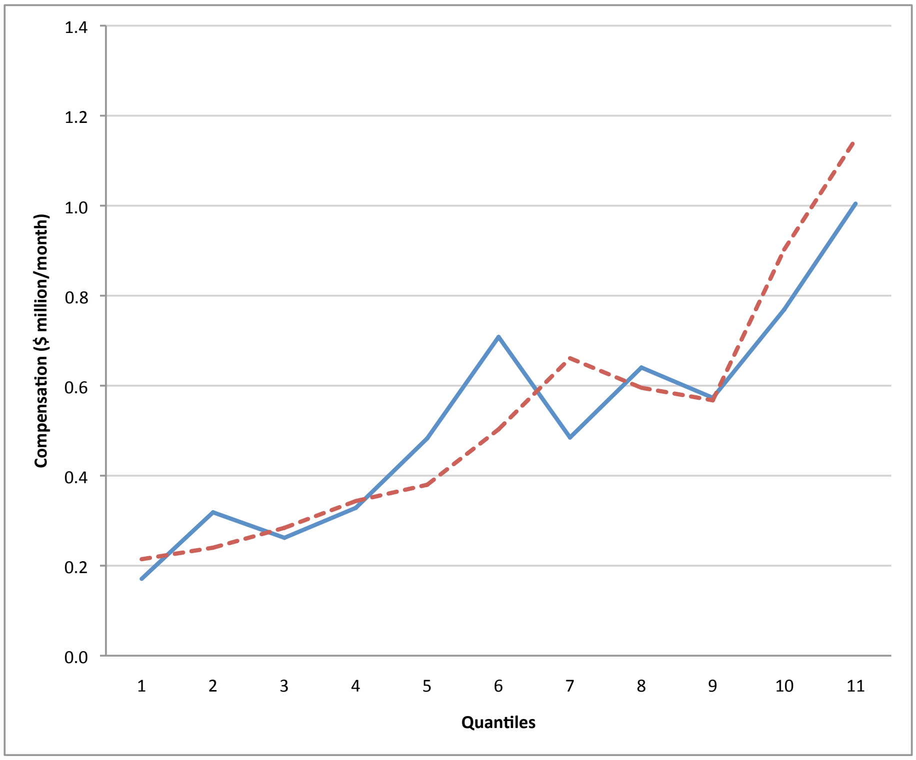

Figure 3: Average Managerial Compensation:
The graph displays average managerial compensation, that is: $q_{it} f_{it}$ in each quantile. Funds are sorted into 11 bins based on t-statistics of the estimate of value added. The graphs plots for each bin the average monthly compensation $(q_{it} f_{it})$ of all funds within that bin. The solid line uses the FFC factor specification as the risk adjustment, and the dashed line uses Vanguard index funds as benchmark portfolios.

The results in Table 4 suggests that like other industries, there is a superstar effect in the mutual fund industry — highly skilled managers make ten times what low skilled managers make. But the relatively high pay of the highest skilled managers appears to be due to their skill difference, rather than a matching problem as in Rosen (1981). That is, our estimates of the skill differential between talented and untalented managers are much larger than the estimates of the compensation difference, although, as we have already noted, the skill estimates are upwardly biased and so this result could very well be due to the bias.

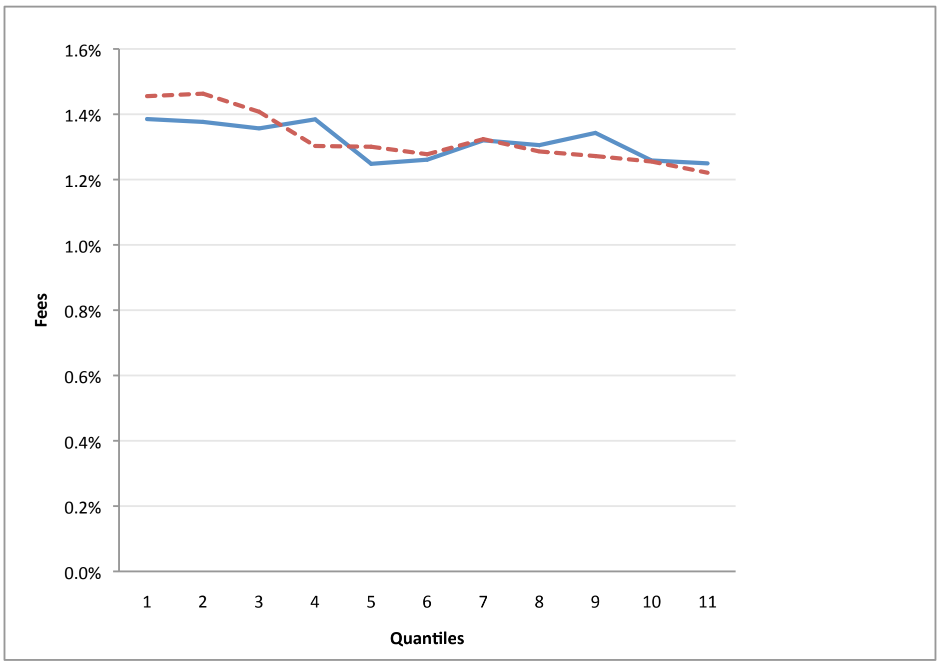

Figure 4: Average Percentage Fee.
Funds are sorted into 11 bins based on t-statistics of the estimate of value added. The graphs plots, for each bin, the average fee f_{it} (in %/annum) of all funds within that bin. The solid line uses the FFC factor specification as the risk adjustment, and the dashed line uses Vanguard index funds as benchmark portfolios.

References

Baker, M., L. Litov, J. A. Wachter, and J. Wurgler (2010): "Can Mutual Fund Managers Pick Stocks? Evidence from Their Trades Prior to Earnings Announcements," *Journal of Financial and Quantitative Analysis*, 45(05), 1111–1131.

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

Berk, J. B., and I. Tonks (2007): "Return Persistence and Fund Flows in the Worst Performing Mutual Funds," Working Paper 13042, National Bureau of Economic Research.

Bollen, N. P. B., and J. A. Busse (2001): "On the Timing Ability of Mutual Fund Managers," *The Journal of Finance*, 56(3), 1075–1094.

——— (2005): "Short-Term Persistence in Mutual Fund Performance," *Review of Financial Studies*, 18(2), 569–597.

Carhart, M. M. (1997): "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 57–82.

Chen, H.-L., N. Jegadeesh, and R. Wermers (2000): "The Value of Active Mutual Fund Management: An Examination of the Stockholdings and Trades of Fund Managers," *Journal of Financial and Quantitative Analysis*, 35, 343–368.

Chevalier, J., and G. Ellison (1997): "Risk Taking by Mututal Funds as a Response to Incentives," *Journal of Political Economy*, 106, 1167–1200.

Cohen, L., A. Frazzini, and C. Malloy (2007): "The Small World of Investing: Board Connections and Mutual Fund Returns," Working Paper 13121, National Bureau of Economic Research.

Cohen, R. B., C. K. Polk, and B. Silli (2010): "Best Ideas," *SSRN eLibrary*.

Coval, J. D., and T. J. Moskowitz (2001): "The Geography of Investment: Informed Trading and Asset Prices," *The Journal of Political Economy*, 109(4), pp. 811–841.

Cremers, K. J. M., and A. Petajisto (2009): "How Active Is Your Fund Manager? A New Measure That Predicts Performance," *Review of Financial Studies*, 22(9), 3329–3365.

Fama, E. F. (1965): "The Behavior of Stock-Market Prices," *Journal of Business*, 38(1), 34–105.

——— (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25(2), 383–417.

Fama, E. F., and K. R. French (2010): "Luck versus Skill in the Cross-Section of Mutual Fund Returns," *The Journal of Finance*, 65(5), 1915–1947.

Glode, V., B. Hollifield, M. Kacperczyk, and S. Kogan (2009): "Is Investor Rationality Time Varying? Evidence from the Mutual Fund Industry," Working Paper 15038, National Bureau of Economic Research.

Grinblatt, M., and S. Titman (1989): "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings," *The Journal of Business*, 62(3), pp. 393–416.

——— (1993): "Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns," *The Journal of Business*, 66(1), pp. 47–68.

Gruber, M. J. (1996): "Another Puzzle: The Growth in Actively Managed Mutual Funds," *Journal of Finance*, 51, 783–810.

Jensen, M. C. (1968): "The Performance of Mutual Funds in the Period 1945–1964," *Journal of Finance*, 23, 389–416.

Jiang, H., M. Verbeek, and Y. Wang (2011): "Information Content When Mutual Funds Deviate from Benchmarks," *SSRN eLibrary*.

Kacperczyk, M., C. Sialm, and L. Zheng (2005): "On the Industry Concentration of Actively Managed Equity Mutual Funds," *The Journal of Finance*, 60(4), pp. 1983–2011.

——— (2008): "Unobserved Actions of Mutual Funds," *Review of Financial Studies*, 21(6), 2379–2416.

Kosowski, R., A. Timmermann, R. Wermers, and H. White (2006): "Can Mutual Fund 'Stars' Really Pick Stocks? New Evidence from a Bootstrap Analysis.," *Journal of Finance*, 61(6), 2551 – 2595.

Malkiel, B. G. (1995): "Returns from Investing in Equity Mutual Funds 1971 to 1991," *The Journal of Finance*, 50(2), pp. 549–572.

Mamaysky, H., M. Spiegel, and H. Zhang (2008): "Estimating the Dynamics of Mutual Fund Alphas and Betas," *Review of Financial Studies*, 21(1), 233–264.

Rosen, S. (1981): "The Economics of Superstars," *The American Economic Review*, 71(5), pp. 845–858.

Sharpe, W. F. (1991): "The Arithmetic of Active Management," *Financial Analysts Journal*, 47(1), pp. 7–9.

SHUMWAY, T., M. B. SZEFLER, AND K. YUAN (2009): "The Information Content of Revealed Beliefs in Portfolio Holdings," .

SIRRI, E. R., AND P. TUFANO (1998): "Costly Search and Mutual Fund Flows," *The Journal of Finance*, 53(5), 1589–1622.

WERMERS, R. (2000): "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses," *The Journal of Finance*, 55(4), 1655–1703.

ZHENG, J. X. (1999): "Testing Heteroskedasticity in Nonlinear and Nonparametric Regressions with an Application to Interest Rate Volatility," Working paper, University of Texas at Austin.