How Does Size Affect Mutual Fund Behavior?

JOSHUA M. POLLET and MUNGO WILSON*

Abstract

If actively managed mutual funds suffer from diminishing returns to scale, funds should alter investment behavior as assets under management increase. Although asset growth has little effect on the behavior of the typical fund, we find that large funds and small-cap funds diversify their portfolios in response to growth. Greater diversification, especially for small-cap funds, is associated with better performance. Fund family growth is related to the introduction of new funds that hold different stocks from their existing siblings. Funds with many siblings diversify less rapidly as they grow, suggesting that the fund family may influence a fund's portfolio strategy.

*Pollet is from the Department of Finance, University of Illinois at Urbana-Champaign. Wilson is from the Department of Finance, Hong Kong University of Science & Technology. We thank Keith Brown, Laurent Calvet, John Campbell, Kalok Chan, Randy Cohen, Joshua Coval, Rafael Di Tella, Andre Perold, Jeremy Stein, Luis Viceira, Eric Zitzewitz, and an anonymous referee as well as seminar participants at Chinese University of Hong Kong, Harvard University, Hong Kong University of Science & Technology, Singapore Management University, University of Illinois at Urbana-Champaign, and the 2006 Western Finance Association Annual Meeting for their comments.

The average equity mutual fund does not outperform the stock market and relatively few actively managed equity funds can persistently outperform passive investment strategies.[1] The absence of superior performance for the average fund combined with the lack of performance persistence appears to suggest a lack of managerial skill. In the absence of skill, why do actively managed funds manage so much money? Berk and Green (2004) indicates that diminishing returns to scale can reconcile the lack of average outperformance and performance persistence with the existence of managerial skill. In their model, money flows to a fund until the marginal dollar can no longer be invested advantageously. In this paper, we investigate the effect of asset growth on aspects of fund investment *behavior*, to identify more precisely the constraints acting on funds as they grow. Regardless of whether diminishing returns to scale should affect fund performance in equilibrium, fund behavior should respond to constraints imposed by growth.

How should a mutual fund invest new money? Should it research a larger universe of investment ideas, hiring new staff and expanding its research capabilities, or should it continue to invest, as far as feasible, in a given set of stocks? Our first set of results documents that funds overwhelmingly respond to asset growth by increasing their ownership shares rather than by increasing the number of investments in their portfolio.

In the year 2000, a typical large fund held fewer than twice the number of stocks held by a fund less than one-hundredth its size. In the panel, a doubling of fund size is associated with an increase in the number of stocks of just under 10%, but this rate of increase declines very rapidly as the number of stocks held by the fund grows. Doubling the number of stocks already held by the fund reduces this rate of increase to zero. Thus, funds appear to be very reluctant to diversify in response to growth but instead tend to acquire ever larger ownership shares in the companies they already own. Ownership shares above 5% are common in our sample for large funds. These results appear to identify limits to the scalability of fund portfolios, such as price impact or liquidity constraints, as the proximate cause of the diminishing returns to scale assumed by Berk and Green. We often refer to such limits to scalability as ownership costs.

Our second set of results provides evidence that diversification is associated with higher monthly risk-adjusted fund returns. Funds that invest in the small-cap sector benefit the most from diversification controlling for fund size and fund family size. These results are complementary to the findings of Chen et al. (2004), which presents evidence that smaller funds outperform larger ones in the small-cap sector. Both our results and those of Chen et al. support liquidity constraints as an explanation for why large-cap funds diversify more slowly in response to growth in assets under management.[2]

These findings are consistent with at least two ways in which liquidity constraints can affect fund performance. In the first case, managers have no ability to generate additional investment ideas when existing opportunities have been fully exploited. All they can do is "go down their list" to the next-best investment opportunity. Managers diversify only because they are prevented by their size from increasing their existing holdings without incurring prohibitive ownership costs. If some managers are able to add superior stocks with greater ease because they have a better list, liquidity constraints will not lower returns as much for these managers. In this situation, managers diversify optimally and the level of diversification reveals an aspect of managerial skill. Thus, diversification will be associated with better fund performance, controlling for size, particularly when liquidity constraints are severe. In the second case, some managers are overconfident about their ability to select superior stocks or underestimate transaction costs. Again, diversification will be positively associated with fund performance, particularly when liquidity constraints are important. However, in this case the overconfident managers are not diversifying optimally.[3]

In either case, funds severely constrained by high ownership costs, for example, small-cap funds, will display a positive association between diversification and subsequent performance, controlling for fund size. By contrast, funds less constrained by ownership costs, for example, small funds, large-cap funds, or possibly funds in large families that benefit from an improved trading environment, will exhibit a weaker relationship.

This evidence regarding fund performance may have implications for the theory of the financial firm. A mutual fund is essentially a firm whose two inputs are financial and human capital and whose output is a set of investments. The results suggest that there are limits to the human capital that can be productively added to a fund. Which factors constitute the sources of these limits, the underlying causes of scaling and lack of diversification, remains an open question.

Our third set of results examines how fund families, rather than individual funds, respond to growth in assets under management. Fund family growth in assets is associated with large increases in the number of funds in the family, especially for the families whose constituent funds already manage a large combined quantity of assets. Moreover, the portfolios in these family funds appear to be different from one another, since the number of different stocks in the family "fund of funds" grows as rapidly as, or more rapidly than, the number of funds as family size increases. Hence, family growth, unlike individual fund growth, appears to be strongly associated with the generation of additional investment ideas and these ideas are produced through the creation of new funds rather than within existing funds. This effect is most pronounced for large families, which dominate the industry in terms of market share. Khorana and Servaes (1999) identifies a cross-sectional relationship whereby large families are more likely to set up a new fund. While our results are consistent with those in Khorana and Servaes, we show that the increase in the number of funds in a family is associated with an increase in family total net assets (TNA).

Finally, we show that the number of sibling funds in a fund family has an additional effect on the response of individual funds to asset growth. While the average fund diversifies slowly in response to growth, funds with many siblings diversify even more slowly. At the very least, fund families do not appear to boost their funds' capacity to generate additional investments within each fund. Indeed, fund families appear to influence individual fund investment behavior in the opposite direction by focusing funds on fewer stocks. Alternatively, families may play a role in alleviating liquidity constraints for individual funds by providing an environment in which the combined family holding in a given stock can be traded at lower cost.[4] These lower costs might also explain why funds in large families diversify more slowly.

The results for fund families are consistent with a world in which large fund families maintain market share through managing a broad range of different funds. Each individual portfolio in the family is kept distinct from its sibling funds even as the portfolio in question becomes extremely large. This family behavior could be interpreted as evidence of product proliferation within the fund family discussed in Massa (1998). Since Sirri and Tufano (1998) indicates that the fund flows respond to marketing and advertising, it is certainly possible that fund families will prefer to establish new funds rather than hire additional managers within an existing fund for marketing reasons.

The rest of the paper proceeds as follows. Section I describes our basic hypotheses and presents data, summary statistics, and evidence from the cross-section. Section II presents results of panel regressions. Section III analyzes the impact of diversification on fund returns. Section IV presents results on the effect of family size. Section V concludes.

I. How Do Fund Portfolios Change with Size?

A. Fund Portfolios with Ownership Costs

What is the effect of growth in total net assets (TNA) on fund behavior? One possible answer is that TNA growth has no effect on behavior: a manager of a $1 billion fund will select stocks in the same way as he or she would managing only $10 million. The manager's chosen portfolio weights for the fund's investments are independent of fund size. We refer to this null hypothesis as "perfect scaling," or "scaling" for short. Of course, we do not expect to observe funds scaling perfectly. It may not even be feasible to invest $1 billion at the same portfolio weights as $10 million. More likely, the increased costs of investing $1 billion in such a manner make this option undesirable. The economically interesting question is not whether funds scale perfectly, but how and to what extent they deviate from scaling.

Berk and Green suggest that diminishing returns to scale in the mutual fund industry can reconcile the lack of persistence in fund return performance with the presence of managerial skill at picking stocks. If money flows to the point at which investors are indifferent between funds, skilled managers will manage larger funds than inept managers, but in expectation no fund will outperform any other. In their model, managers are assumed to face costs that are positive, increasing, and convex in fund TNA. These assumptions are intended to capture the idea that "with a sufficiently large fund, a manager will spread his information gathering activities too thin or that large trades will be associated with a larger price impact and higher execution costs."

We emphasize that if the acquisition of a large holding does not increase price impact then there is no need for a particular manager to alter information gathering activities at all. The manager can simply scale up his or her few best investment ideas. The price impact costs of large holdings are the necessary seed of diminishing returns to scale, although there may certainly be interesting auxiliary sources of diminishing returns that may begin to act in the presence of price impact. Price impact requires managers to deviate from perfect scaling by increasing the number of distinct holdings as fund TNA grows.

We consider two basic propositions. First, in the presence of liquidity costs, managers will slowly increase the number of distinct holdings in their portfolios in response to flows of new money. This response will be greater when liquidity costs are more severe. Second, managers will increase ownership shares in response to new flows at a declining rate as the fund grows.

The apparently limited ability of fund managers to generate additional (equally good) investment ideas given the imperfect scalability of their fund's portfolio is particularly important. Otherwise, why not invest in these additional ideas and avoid price impact altogether? One possibility is that it is suboptimal to hire additional money managers or research staff to augment the number of investment ideas. Both the costs of organizational diseconomies described by Chen et al.[5] and the benefits of product proliferation as part of marketing strategies for fund families described by Massa would be consistent with this explanation.

Another possible response to liquidity constraints is to close the fund to new investors. Bris et al. (2006) investigate fund closures in detail. They find that funds usually close in response to inflows of new money and that the majority of such funds report small company growth as their investment objective. These results are entirely consistent with the hypothesis that closure is primarily a response to higher liquidity costs. Since the largest number of closures in any year of the study's sample is 24, a tiny fraction of the mutual fund population, we do not consider fund closure separately as a response to liquidity costs.

We measure the extent to which funds scale and the extent to which they diversify in response to growth in TNA. To the extent that funds scale, fund ownership shares should increase with TNA. If diversifying forces such as price impact are at work, a higher level of ownership should slow the rate at which ownership increases with TNA and force funds to add new stocks to their portfolios. We start by discussing the cross-section before turning to the results of panel estimates of scaling versus increased diversification.

B. Data and Summary Statistics

We use mutual fund data from two sources. The mutual fund database from CRSP contains fund assets under management (TNA) at the end of the year, objective codes, management company, and assets allocated to equities for funds since 1961. The mutual fund database from CDA (now owned by Thomson Financial) has fund equity holdings by stock, objective codes, management company, and another measure of TNA for most equity mutual funds in the CRSP data set from 1975. We use the matched sample from 1975 to 2000, rather than just the CDA data, because of the higher quality of the CRSP data on fund returns, TNA, and management objective codes. In addition, CRSP gives each fund a unique identifier, whereas funds in the CDA database can change identifier when their name changes, making it difficult to track all funds through their entire history. Finally, foreign funds investing in equities listed in the United States are excluded from CRSP but not from CDA.

We match these databases by fund name, TNA, and, when available, NASDAQ ticker.[6] Starting with the CRSP data, and using objective codes and a keyword search of fund names, we exclude balanced funds, bond funds, commodity funds, index funds, sector funds, and foreign equity funds. Funds missing monthly returns or TNA data for all months in a given year are excluded for that year, as are funds with less than 50% of their TNA allocated to equities throughout their history. The remaining sample is matched to CDA. We use CDA data for the last report issued during the year. Next, we exclude funds with fewer than 10 different equity holdings.[7] To supplement the equity holdings information from CDA, we link portfolio holdings to CRSP stock data with prices and shares outstanding. We treat funds with the same management company identifier in CDA as belonging to the same family of funds.

Table I gives summary statistics for the matched sample for every fifth year since 1975. Column 2 gives the number of funds, column 3 the number of fund families, column 4 the average fund TNA, column 5 the combined TNA managed by these funds, column 6 the combined TNA as a percentage of CRSP total market value (a measure of the sample's market share), and column 7 the average value-weighted return, after expenses, earned by this group. Column 8 gives the CRSP total market return. The number of funds in our sample differs from Carhart (1997) because we aggregate share classes of the same fund into one observation for each year and some funds in CRSP do not have a matching record in CDA.

Approximate location for Table I

Column 2 shows steady growth in the number of mutual funds in the sample, from 253 in 1975 to 1,421 in 2000, with the number nearly tripling in the 1990s. The ownership share of the funds in our sample in the market as a whole grew from less than 5% of the market capitalization of all stocks in CRSP in 1980 to approximately 13% in 2000. In the last year of the sample the average fund managed $1.44 billion dollars and the sample as a whole managed approximately $2 trillion. From the point of view of growth in market share, the industry has been extremely successful. Since we exclude many kinds of funds that hold equities listed in the United States, this calculation is a lower bound for the total market share of the actively managed equity fund industry. The last two columns show that investors in actively managed equity mutual funds have earned high average returns, although the average returns for these funds are not as high as those on the aggregate market. An aggregate market index would have outperformed a typical mutual fund investment, but not by very much.[8]

C. The Cross-section of Funds

For each year we sort all funds in the sample into quintiles by fund TNA. We report results in Table II for every tenth year starting in 1980. These years are representative of the full sample period. Quintile 1 contains the smallest funds. Column 3 reports the number of funds in each quintile. In addition to reporting statistics for each fund size quintile, we also include attributes of the CRSP stock price database. Column 4 shows the percentage of the sample's combined TNA for a given year managed by the funds in each quintile, giving a measure of the relative size of each quintile. The share of the largest quintile has grown over the sample period from 74% in 1980 to 86% in 2000. In 1980 the largest 40% of funds managed 89% of total industry TNA, rising to over 95% in 2000. Column 5 reports the mean TNA of funds in each quintile. While the size of funds in the bottom quintile increased by less than a factor of five from 1980 until 2000, the size of funds in the top quintile increased by more than a factor of 10 during the same time period. This is consistent with a pattern of rising stock prices and entry by relatively small new funds.

Approximate location for Table II

Column 6 presents the main result of the table. Although the average number of different stocks held by a fund in a given quintile increases with TNA, it does so very slowly. The number for the largest quintile is never more than three times the number for the smallest. However, in 1980 funds in the largest quintile were about 100 times as large as those in the smallest and in 2000 funds in the largest quintile were approximately 300 times as large, holding fewer than twice as many stocks. The ratio of the average number of stocks held by the largest versus the smallest quintiles actually declined over this period even though the spread in TNA widened. The bottom quintile may be misleading because of the exclusion of funds with very few stocks, but the differences between the middle quintiles are in some ways even more remarkable. In 2000, a fund managing $6.2 billion hardly had any more stocks, on average, than a fund managing \$650 million (144 versus 137). Relatively large mutual funds do not behave as if they have many more good investment ideas nor as if they have a great deal more difficulty investing their money compared to their smaller counterparts. The row labeled 'CRSP' in column 6 reports the total number of stocks listed in the United States, excluding ADRs, closed-end investment funds, REITs, and certain other kinds of companies.

The average number of stocks held by a fund has increased over time, irrespective of TNA. Campbell et al. (2001) shows that the average idiosyncratic risk of stocks increased during this period, so that the number of randomly chosen stocks required to reduce risk below a given level has increased. This finding might suggest that funds choose a minimal level of diversification to reduce risk. Alternatively, the number of firms with a relatively small market capitalization has increased over the sample period. As a result, the average fund may have increased the number of its holdings precisely as an optimal response to rising ownership costs associated with the shrinking market capitalization of a typical firm. These two explanations are not mutually exclusive. Indeed, they may be closely related because Brown and Kapadia (2006) indicates that all of the increase in idiosyncratic risk noted by Campbell et al. is due to new listings. Among these new listings, small firms are disproportionately represented.

We define a fund's ownership share in a company as the number of shares held divided by the number of shares outstanding. Column 8 reports the mean (equally weighted across funds) maximum ownership share in each TNA quintile. If ownership costs are the main constraint preventing perfect scaling, then the maximum ownership share is associated with a fund's most expensive stock pick. The fund's largest ownership share increases strongly with fund TNA to above 4% for highest-TNA funds in all years. Figure 1 plots, for every fifth year in the sample, average maximum ownership share for each TNA-sorted decile against that decile's share of total market value, an increasing function of average TNA. Broadly speaking, the relationship is increasing but concave, with the curves flattening out well before the legal upper limit of 10%. The last column also reports the cross-sectional mean of the average ownership share within the fund. This measure also increases monotonically with fund TNA in every year in the sample. Figure 2 plots mean ownership share against market share by decile and the relationship is also increasing and concave.

Approximate location for Figure 1 and Figure 2

Column 7 of Table II shows a tendency for funds with higher TNA to hold stocks in companies with larger market capitalizations. This "style" measure is defined as the weighted average market capitalization of companies owned by the fund, using the fund's portfolio weights. Thus, for fund i, stocks j with market capitalizations $mcap_{jt}$, and portfolio weights w_{ijt}, the fund's style at time t is given by

$$Style_{it} = \sum_j w_{ijt} mcap_{jt}. \tag{1}$$

We also report this measure for the CRSP index in billions of dollars. All quintiles have an average style less than that of the market, suggesting that most funds do not have relatively high weights on stocks with extremely large market capitalizations.[9] Fund style has a discernible tendency to increase with TNA in most sample years. One way for a manager to mitigate increasing ownership costs without diversifying as a fund grows is to migrate to stocks with larger market capitalizations. However, the cross-sectional relationship may also be due to market segmentation, with greater demand for larger-cap funds.

We also sort funds into quintiles based on our style measure and report the same fund attributes by quintile in Table III that we report in Table II. Column 4 of Table III shows that assets are fairly evenly distributed across the different styles. The smallest-cap funds account for less market share and tend to have lower TNA but the other quintiles are similar to each other. The specialist large-cap funds are not always the largest funds or the largest market segment. The average number of stocks is also similar across funds but tends to be higher for the quintiles of funds with low market capitalization styles. The number of stocks held by a fund shows the same upward trend over time documented in Table II. Column 7 shows that the funds in the highest style quintile have an average style that is higher than the market style in all selected years. Maximum and mean ownership shares are generally declining with style as we would expect, suggesting that liquidity constraints bind most on small-cap funds.

Approximate location for Table III

II. How Much Do Funds Scale? Panel Evidence

A. Regression Specifications

The cross-sectional evidence provides a natural starting-point to examine how size relates to fund portfolio characteristics. However, to establish the effect of growth in TNA on portfolio choice, it is necessary to use a panel specification to follow funds over time. Funds that choose stocks well and do not rebalance their portfolios will have high TNA and a portfolio of large-cap stocks in the absence of any interesting economic relationships between size and style. Such funds would have high TNA growth and no change in the number of stocks held. Therefore, we use the log of the flow of money into the fund instead of total growth in fund size. We define log flow as the change in log TNA not attributable to the portfolio return, or

$$\log Flow_{it} = \log \frac{TNA_{it}}{TNA_{it-1}} - \log(1 + R_{it}). \tag{2}$$

A log flow measure of 1% corresponds to a net increase in fund TNA of approximately 1% if returns over the same period were zero. Our results, in terms of the effect of log TNA growth on our dependent variables, are very similar if we include log TNA growth and lagged returns in place of log flow.

We focus on two aspects of portfolio selection, namely, diversification, as proxied by the number of different stocks in the portfolio, and scaling, as measured by the portfolio-weighted average log ownership share in stocks held. These are respectively measured by the log number of stocks $\log S_{it}$ and

$$\log Own_{it} = \sum_{j=1}^{J} w_{ijt} \log c_{ijt}. \tag{3}$$

We define c_{ijt} as the number of shares held by fund i in firm j divided by the number of shares outstanding at time t. Our dependent variables are the first differences, annually, of these quantities. For $\Delta \log Own_{it}$, we also consider various permutations including holding weights constant at either old or new portfolio weights or using hypothetical weights that change solely as a result of returns. None of these alternatives affect our results to any significant degree.

We estimate two regressions:

$$\begin{aligned}\Delta \log S_{it} &= \alpha_1 + \delta_{1t} + \beta_{11} \log Flow_{it} + \beta_{12} \log S_{it-1} + \beta_{13} \log Own_{it-1} \\ &\quad + \beta_{14} X_{it-1} + \beta_{15} X_{it-1} * \log Flow_{it} + \varepsilon_{1it},\end{aligned} \tag{4}$$

and

$$\begin{aligned}\Delta \log Own_{it} &= \alpha_2 + \delta_{2t} + \beta_{21} \log Flow_{it} + \beta_{22} \log S_{it-1} + \beta_{23} \log Own_{it-1} \\ &\quad + \beta_{24} X_{it-1} + \beta_{25} X_{it-1} * \log Flow_{it} + \varepsilon_{2it}.\end{aligned} \tag{5}$$

The variables $\log S_{it-1}$ and $\log Own_{it-1}$ are included as proxies for the extent to which liquidity constraints are already important at $t-1$. We also use two different indicator variables for X_{it-1}, separately and together. First, $High\text{-}TNA_{it-1}$ equals one if fund i had above median TNA at the end of year $t-1$. Second, $Large\text{-}Cap_{it-1}$ equals one if fund i's style was greater than the median at the end of year $t-1$.

We assume that the residuals are not autocorrelated and calculate standard errors that adjust for heteroskedasticity and contemporaneous correlation across funds by clustering the residuals by year. Our concern is that the error term in these regressions exhibits cross-sectional correlation of unknown form because funds with similar unobserved characteristics may have correlated residuals. This problem is appropriately addressed using clustering by year. However, it is also possible that residuals are correlated across years for a given fund. We examine this possibility by clustering by fund in our panel regressions; we find that clustering by year yields much more conservative standard errors.

While we do not explicitly control for differences in the investment opportunity sets of different funds, the dependent variable $\Delta \log S_{it}$ internally adjusts for these differences. We could have used the change in the log of the number of holdings for each fund divided by all available holdings within the fund's style universe as an alternative measure that explicitly benchmarks the number of holdings against the investment opportunity set. For example, assume that a small-cap fund holds 200 out of a possible 5,000 small-cap stocks and a large-cap fund holds 100 out of a possible 500 large-cap stocks. The investment opportunity set of the small-cap fund is greater than the investment opportunity set of the large-cap fund. Suppose each fund receives the same proportional flow of new money between this period and next period and each doubles the number of holdings in its portfolio, so that the small-cap fund now holds 400 stocks and the large-cap fund now holds 200. The change in the log number of stocks is log(2) for each fund and our regressions using $\Delta \log S_{it}$ would then report that each fund diversified at the same rate in response to the inflow of money. The alternative measure is also the same for each fund. This measure equals log(200/500)-log(100/500) or log(2) for the large-cap fund and log(400/5000)-log(200/5000) or log(2) for the small-cap fund. Even if the number of available holdings in all styles is not constant, but instead grows at the same rate for all styles for a particular year (e.g., as required by any style measure built using a quantile-based sorting methodology), the year fixed effects in our specifications fully control for the change in the investment opportunity set.

B. The Effect of Flows on Diversification and Ownership Shares

Table IV presents estimates of equation (4). Standard errors are reported in parentheses. The estimated coefficient on log flow is between 10.9% and 14.6% for the different specifications and is always highly significant. The coefficient on log number of stocks is opposite in sign and almost equal in magnitude - about 10% in all specifications - and also statistically significant. Thus, although a 1% increase in size not due to previous returns is associated with a 0.1% increase in the number of stocks held, this is reduced to almost zero if the existing number of stocks is 1% higher. These results support our hypotheses that funds add stocks at the margin as they grow, but at a rate that declines with the ability to avoid high ownership costs by investing new money in a greater number of existing holdings.[10]

Approximate location for Table IV

The coefficient on average log ownership is positive and significant in all but the last specification. Other things equal, higher existing ownership shares are associated with mildly more rapid diversification as funds grow. This evidence suggests that funds facing higher ownership costs diversify more rapidly. However, the last specification suggests that existing ownership share has no effect on diversification except to the extent that it proxies for existing fund size and capitalization style.

Funds with higher TNA diversify more rapidly in response to fund asset growth, as hypothesized, controlling for the number of stocks already owned and existing ownership shares. The coefficient on the interaction of $High\text{-}TNA_{it-1}$ and $\log Flow_{it}$ is positive, stable across specifications, and highly significant. In column 8, a 1% increase in TNA due to flows of new money raises the number of stocks by an additional 0.04% for the larger funds. As predicted, large-cap funds diversify more slowly in response to growth. The effect of style is quite large in the relevant specifications. Controlling for the existing number of stocks, a fund in the large-cap half of the sample diversifies about 50% more slowly than a fund in the small-cap half, adding 0.05% fewer stocks in response to a 1% increase in assets due to inflows.

Table V reports estimates of equation (5). The main result in this table is that funds appear to be quite close to scaling perfectly. In column 5, a flow of 1% is associated with an increase in average ownership share of approximately 0.72%. This estimate is not very different in the other specifications. Funds do not appear to alter their investment strategies in response to new flows of money to any great extent, from year to year.

Approximate location for Table V

On the margin, growth in log ownership share is related to fund size. First, the more diversified the fund's existing portfolio, the more rapidly ownership share grows as new funds flow in. This is shown by the positive coefficient on the log number of stocks. Second, a higher current ownership share reduces scaling by between 0.04% and 0.06% for a 1% increase in TNA due to flows of new money. This acts as a drag on ownership share growth in future years. High-TNA funds scale less rapidly and large-cap funds scale more rapidly in response to flows but the coefficients are not statistically significant.

III. Diversification and Returns

The preceding section presents evidence that funds diversify extremely slowly as their assets under management grow. At the same time, funds' ownership shares in the firms they invest in increase with inflows of new money. Holding the manager's information set constant, this result suggests that, absent liquidity constraints, funds invest new money in the same proportions as their existing portfolios. Where liquidity constraints can reasonably be expected to be more onerous - in the small-cap sector and among large funds - we find funds diversify faster as they grow and scale less as they receive new money.

In this section we examine the relationship between diversification and subsequent performance controlling for other determinants of fund performance that have been proposed in the literature. Our measure of diversification is the reciprocal of the number of stocks $1/S_{it-1}$ held by fund i at the end of year $t-1$. By construction, this variable is the average weight for the fund. If the average weight is low, then the fund is better diversified. This measure declines nonlinearly with the number of stocks in a way that intuitively should capture any relationship between returns and diversification: an increase from 30 to 50 stocks is a much more significant change if transaction costs are convex compared to changing from 3,000 to 5,000 stocks. The latter example is essentially a comparison of two slightly different stock indices. This variable has good properties in the case of outliers and closet indexers.

For each month in the sample we estimate cross-sectional regressions of risk-adjusted fund returns (after fees and expenses) on a constant and fund characteristics from December of the previous calendar year, and then average the coefficients across months, using the procedure of Fama and MacBeth (1973). We follow Fama and MacBeth rather than use panel regressions to ensure comparability between our results and those of Chen et al.[11] For the same reason we also report t-statistics, rather than standard errors, in parentheses in Table VI and Table VII. The risk-adjusted monthly return is the fund return minus each fund factor loading multiplied by its respective factor return. Fund factor loadings are estimated from the regression of the monthly portfolio return of the stocks held in the fund at the end of year $t-1$ (using portfolio weights from CDA) on the Fama-French-Carhart four factors for the previous 36 months. The fund characteristics are log family TNA (one plus the combined TNA of all other funds in a given fund's family), expense ratio, total load, log of fund age, and industry concentration index.[12] We group funds into families based on the management company abbreviation from CDA.

Using the capitalization style measure (the weighted average market capitalization of firms whose stocks are held by the fund), we sort funds into quintiles and create indicator variables $cap1_{it-1}$ through $cap5_{it-1}$, where $cap5_{it-1}$ contains large-cap funds. We include $\log TNA_{it-1}$ and $1/S_{it-1}$ together with indicator variables for style quintiles and their interactions with $\log TNA_{it-1}$ and $1/S_{it-1}$ in our analysis of fund performance. We estimate regressions both for the full matched sample in Table VI and for the sample excluding all funds with TNA under \$100 million in Table VII. As can be seen from Table II, the excluded funds accounted for less than 2% of the total TNA of all funds in 2000 and approximately 10% of all TNA in 1980. The economic importance of large funds is much greater so we are interested mainly in results that hold for these larger funds.

In column 1 of Table VI, we estimate the simple relationship between the risk-adjusted fund return and $1/S_{it-1}$. The interpretation of the coefficient depends on the initial number of distinct holdings in the portfolio. For example, two otherwise identical funds with respectively 20 and 100 holdings have $1/S$ of 0.05 and 0.01, respectively. Hence, the monthly return for the more diversified fund is higher by -0.0305 times -0.04, or by about 1.5% in annualized terms. In column 2 we add fund TNA and in column 3 we add both fund TNA and family TNA. As Chen et al. finds, higher fund TNA is associated with lower returns, while higher family TNA is associated with higher returns. However, diversification has a larger impact, controlling for the size of the fund and the size of the family. Column 4 includes controls for other fund characteristics. We find higher expenses and total load to be significantly negatively associated with returns and higher industry concentration to be positively associated with marginal significance. Including these other variables does not change the qualitative impact of diversification and the magnitude of the coefficient on $1/S_{it-1}$ increases further.

Approximate location for Table VI

In column 5 we divide the funds into style quintiles and estimate the marginal effect (which may be highly nonlinear) of a fund's market capitalization style. The coefficients on $1/S_{it-1}$ and $\log TNA_{it-1}$ are interpreted as the coefficients for the smallest fund style quintile. The diversification and size coefficients for the funds in the other quintiles are those on the relevant interaction terms plus the coefficient on the smallest style quintile. The impact of $1/S_{it-1}$ on fund returns grows monotonically from the smallest (it is twice as large in magnitude compared to column 5) to the fourth largest style quintile, but declines slightly for the largest-cap style quintile. Essentially, undiversified funds have lower returns if they invest in small-cap stocks because the coefficient on $1/S_{it-1}$ is significantly negative for small-cap funds. The coefficients on the interaction of $1/S_{it-1}$ with the style quintiles suggest that the impact of diversification is significantly different in the other style quintiles compared to the bottom quintile. We conclude that diversification is associated with significantly higher returns for small-cap funds but this relationship is much weaker for the other style quintiles and is actually associated with lower returns for style quintile 4 (although not significantly so).

The impact of fund TNA on fund returns increases monotonically with fund style. This is consistent with the results of Chen et al., which finds that the negative relationship between fund size and return is most pronounced for small-cap funds. Our results are similar in this respect because the relationship between returns and fund TNA is significantly negative for funds in the bottom style quintile. We also find that the relationship for funds in style quintiles 2 through 5 is significantly different from the relationship for the lowest style quintile. Indeed, the differences are so large that there is actually a positive (although not significant) relationship between fund returns and fund size for funds in the top two style quintiles.

In Table VII we exclude all funds with TNA of less than \$100 million and repeat our analysis of fund returns. Removing small funds mitigates any residual survivorship bias and in any case these funds account for a very small proportion of all assets under management. In addition, equilibrium theories of fund and investor behavior are more likely to be valid among larger funds. The industry concentration index ceases to be statistically significant in this subsample. The results in Table VII for fund size and family size are largely similar to the results in Table VI. However, fund size is no longer beneficial for funds in the largest capitalization style quintiles. The coefficient estimates for $1/S_{it-1}$ double in magnitude compared to those in

Table VI in all specifications. If money moves between funds to alleviate diseconomies of scale to any extent (e.g., in Berk and Green subsequent performance is independent of size because flows fully reflect liquidity constraints), then the true impact of liquidity constraints on fund performance is larger than that estimated in Table VI and Table VII.

Approximate location for Table VII

Given the results of Table IV, we conclude that the lack of scalability of small-cap portfolios, together with high TNA, forces funds to diversify or suffer erosion of returns. However, unless forced on funds through liquidity constraints, diversification is not associated with subsequent fund performance. The positive association between fund performance and diversification for liquidity-constrained funds is consistent with sub-optimal diversification by fund managers, optimal diversification in the presence of differential managerial skill, or both. Thus, we find that some funds are unable to generate many additional successful investment ideas and no fund need try to do so except as a response to liquidity constraints.

IV. Fund Families and Asset Growth

A. Larger Funds or More Funds?

Funds are often members of larger fund families - groups of funds with the same management company, such as Fidelity and Vanguard. These companies can offer a variety of mutual funds to the public, just as a food company can offer more than one food product. Since a family is a group of funds, it can also be treated as an investment portfolio, with a family TNA equal to the combined TNA of its constituent funds, a diversification measure equal to the number of distinct equity holdings across all of its funds, and so forth. In this section we present evidence on the relationship between growth in assets at the family level and other aspects of fund families.

Table VIII presents statistics on fund families for the years 1980, 1990, and 2000. Families are sorted into quintiles by family TNA. According to column 3, the number of fund families has more than tripled over the sample period. Like mutual funds, most of the increase in the population of mutual fund families occurred in the 1990s. Column 5 indicates that the average TNA of a family in the smallest quintile has only risen by a factor of three since 1980. During the same period, the average TNA of a family in the largest quintile has increased by nearly 20 times. Consequently, the market share (in column 4) of the largest family quintile has increased from about 80% in the first 10 years of the sample to more than 90% in 2000. The TNA distribution of fund families is even more skewed than that of funds themselves, and from the point of view of market share, the smallest 40% of fund families are irrelevant. Thus, the growth in market share of large funds since 1980 shown in Table II has been outpaced by the growth in market share of large families.

Approximate location for Table VIII

Column 6 shows that the smallest quintile of families often consists of single-fund families, or "only-children." The number of siblings grows with family TNA. During the sample period, the number of funds in the largest families has increased, from an average of 4.32 in 1980 to 7.23 in 2000 and from a maximum of 14 in 1980 to 38 in 2000. The last column shows that family diversification as a function of TNA is roughly in proportion to the number of sibling funds and is more rapidly increasing than the diversification of individual funds (with the possible exception of 2000, given the enormous average size of the estimate for the highest-TNA quintile).

We estimate the association of family asset growth with the change in the log number of funds in the family ($\Delta \log F_{ft}$) and the change in the log number of different stocks ($\Delta \log S_{ft}$) using the panel regression approach described in Section III. We use the lagged levels $\log F_{ft-1}$ and $\log S_{ft-1}$ to proxy for the effects of prevailing circumstances before realization of TNA growth. We sort all families into TNA quintiles at the beginning of each year and perform the following panel regression for each quintile, for $x_{ft} = \log F_{ft}$ and $\log S_{ft}$:

$$\Delta x_{ft} = \alpha + \delta_t + \beta_1 \Delta \log FAMTNA_{ft} + \beta_2 x_{ft-1} + \varepsilon_{ft} \tag{6}$$

The results are reported in Table IX with standard errors clustered by year in parentheses. For all quintiles, the number of fund siblings increases in response to family asset growth. The existing number of funds retards the rate at which new funds are added. All estimates are highly statistically significantly different from zero. While the retarding effect of existing numbers of siblings is roughly the same across family TNA quintiles, the rate at which funds are added as the family asset base expands, controlling for existing number of funds, rises monotonically across quintiles from 8% for the smallest families to 48% for the highest TNA families.

Approximate location for Table IX

The $\Delta \log S_{ft}$ regressions show that the number of different stocks in a fund family grows at roughly the same rate as the number of funds across quintiles (slightly more rapidly for low TNA families and slightly more slowly for high TNA families). Again, the number of new stocks added with TNA growth is retarded by the existing number of stocks at a similar rate across quintiles. A comparison with Table IV shows that most families' combined portfolios diversify three to four times more rapidly than individual fund portfolios in response to growth.

Our interpretation of Table IX is that the new funds have quite different portfolios from existing funds. New funds appear to be created for marketing purposes in order to attract new flows and preserve market share, a strategy pursued most aggressively by the families with the highest market share. In principle, this strategy could explain why there is such a high number of mutual funds in our sample and the proliferation of funds within a few large families. Our results are consistent with those of Khorana and Servaes (1999) and Khorana and Servaes (2006). The former paper shows that high TNA families are more likely to open new funds and that the probability of opening a new equity fund with a given objective is negatively related to the percentage of family TNA previously invested in the same objective. The latter paper shows that fund families that open a greater number of new funds have a higher market share. We show that increases in family TNA are associated with increases in the number of funds and that this association is much stronger for high TNA families. Furthermore, the new funds are genuinely different from the existing funds, consistent with product proliferation. Massa indicates that the mutual fund industry could have the characteristics that support an equilibrium with product proliferation.

B. Does Organization Matter?

The final set of regressions is motivated by the following thought experiment: how would a fund manager in a large organization behave differently from the same person managing the same fund in a smaller organization? The highest TNA funds diversify in response to growth. If the size of an organization is important, funds of equal TNA should diversify at different rates in different organizations. For example, funds with many siblings may be constrained, for marketing or other purposes, from straying into their sibling funds' areas of competence when looking for new ideas. In general, the more funds in a family, the more complicated the decision-making procedures of a family may be in order to avoid conflicts of interest, satisfy marketing objectives, and verify that managers are carrying out their assigned tasks. Given two talented fund managers, it may be more productive in a large organization to give each of them a fund of their own rather than to give them joint control of a single large fund. In addition, families may provide a better trading environment for their funds and thereby relieve constraints on scalability.

Our measure of organizational complexity, the number of funds in a family, is designed to capture these possibilities but is unable to distinguish between them. Our prediction is that funds in more complex organizations are less likely to diversify in response to fund TNA growth. We consider the following regression to test this possibility:

$$\begin{aligned}
\Delta \log S_{it} \quad = \quad & \alpha + \delta_t + \beta_1 \log Flow_{it} + \beta_2 Big\text{-}Family_{it-1} + \beta_3 High\text{-}TNA_{it-1} \\
& + \beta_4 Big\text{-}Family_{it-1} * High\text{-}TNA_{it-1} \\
& + \beta_5 Big\text{-}Family_{it-1} * \log Flow_{it} \\
& + \beta_6 High\text{-}TNA_{it-1} * \log Flow_{it} \\
& + \beta_7 Big\text{-}Family_{it-1} * High\text{-}TNA_{it-1} * \log Flow_{it} + \varepsilon_{it}.
\end{aligned} \tag{7}$$

To construct the variable $Big\text{-}Family_{it-1}$, we sort all funds each year into quintiles by fund TNA and identify the family in each size quintile with the largest number of funds (the most populous family); $Big\text{-}Family_{it-1}$ is an indicator variable that equals one if fund i is in the most populous family in its TNA quintile in year t and zero otherwise. We interact this indicator variable with $\log Flow_{it}$.

Table X reports our results with standard errors clustered by year in parentheses. For the full sample period, membership in a populous family does not significantly change the rate at which a fund diversifies in response to flows. The results in the first two specifications are not qualitatively different from those in Table IV. Note the important role played by the number of stocks $\log S_{it-1}$ in the portfolio – the R^2 of the regression doubles when this variable is included.

Approximate location for Table X

We also consider the second half of the sample period separately because large fund families are not prevalent in the 1970s or early 1980s. In this more recent period, the triple interaction term for $High\text{-}TNA_{it-1}$, $Big\text{-}Family_{it-1}$ (many siblings), and $\log Flow_{it}$ has a large negative sign and is statistically significant (only at the 10% level if lagged number of stocks is included). Large funds diversify much more slowly, other things equal, if they have many sibling funds. Although this result is statistically weaker than many of the results discussed in the other sections, it is economically large. Moreover, summing the four relevant coefficients, we would fail to reject the hypothesis that large funds in populous families do not diversify at all in response to flows. The results also suggests a role for the family both in enforcing focus on a smaller universe of stocks, perhaps for marketing reasons, and preventing such focus from impairing fund performance, perhaps by enhancing liquidity.

V. Conclusion

We present new results on ways in which fund portfolios are affected by growth in TNA. Primarily, funds increase ownership shares in the companies they own. This evidence suggests that managers seek to scale up their existing investments as the fund grows, and are not interested in generating additional investment ideas except to compensate for liquidity constraints. Managers appear to be trying to remain focused on their few best bets, following Mark Twain's investment advice: "Put all your eggs in one basket, and watch that basket."

Second, we show that the number of stocks in the portfolio increases at a slow rate in response to flows. Thus, managers act as if they internalize slowly growing costs of ownership. This diversification in response to growth is less pronounced for funds with large numbers of siblings and for larger-cap funds.

Our results provide new evidence that the proximate cause of diminishing returns to scale for mutual funds is the inability to scale an investment strategy as the fund becomes large. Funds diversify and scale less as they grow and small-cap funds, large funds, and less diversified funds display these responses more strongly, consistent with the limits to scalability being related to liquidity constraints. In contrast to the previous literature, we document a response in fund behavior to size growth, rather than just linking fund size (and other characteristics) to returns.

We document a positive relationship between diversification and subsequent performance, controlling for fund size and fund family size. This relationship is stronger for small-cap funds presumably because they are more constrained. Either managers diversify optimally and the level of diversification reveals their skill or some managers diversify suboptimally.

We find that family growth, especially for the dominant large families, is mainly associated with the addition of new funds instead of an expanded scope of activities for existing funds. Our results are consistent with those in Khorana and Servaes (1999), but we show that the increase in the number of funds is associated with an increase in family TNA, rather than it merely being the case that larger families set up more funds. Furthermore, we provide more direct evidence that the new funds are different from the existing funds because of the rapid increase in diversification at the family level, in contrast to the slow increase at the fund level. This finding is monotonically increasing in fund family size, indicating that marketing considerations may be particularly important for large families. Product proliferation is a well-documented phenomenon in industrial organization and is widely understood as a strategy aimed at preserving market share. Consistent with this possibility, we present new evidence that funds in populous families diversify less than similar funds as they grow.

An alternative but not inconsistent explanation for these results is advanced by Chen et al. Organizational diseconomies in the form of hierarchy costs at large funds rapidly reduce the marginal product of additional human capital to the point where extra managers contribute no useful additional investment ideas. While potentially very interesting, this leaves open questions about why funds cannot contract to set up internal sub-funds that are managed independently but marketed to the public as one investment product.

REFERENCES

[1] Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269-1295.

[2] Bris, Arturo, Huseyin Gulen, Padma Kadiyala, and P. Raghavendra Rau, 2007, Good stewards, cheap talkers or family men? The impact of mutual fund closures on fund managers, flows, fees and performance, *Review of Financial Studies* 20, 953-982.

[3] Brown, Gregory, and Nishad Kapadia, 2007, Firm-specific risk and equity market development, *Journal of Financial Economics* 84, 358-388..

[4] Campbell, John Y., Martin Lettau, Burton G. Malkiel, and Yexiao Xu, 2001, Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk, *Journal of Finance* 56, 1-43.

[5] Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

[6] Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276-1302.

[7] Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristics-based benchmarks, *Journal of Finance* 52, 1035-1058.

[8] Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 71, 607-636.

[9] Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 51, 783-810.

[10] Jensen, Michael C., 1968, The performance of mutual gunds in the period 1945-1968, *Journal of Finance* 23, 389-416.

[11] Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed mutual funds, *Journal of Finance* 60, 1983-2011.

[12] Khorana, Ajay and Henri Servaes, 1999, The determinants of mutual fund starts, *Review of Financial Studies* 12, 1043-1074.

[13] Khorana, Ajay and Henri Servaes, 2006, Conflicts of interest and competition in the mutual fund industry, working paper, London Business School.

[14] Malkiel, Burton G., 1995, Returns from investing in equity mutual funds, 1971-1991, *Journal of Finance* 50, 549-572.

[15] Massa, Massimo, 1998, Are there too many mutual funds? Mutual fund families, market segmentation and financial performance, working paper, INSEAD.

[16] Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589-1622.

[17] Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs and expenses, *Journal of Finance* 55, 1655-1695.

Endnotes

[1]These empirical regularities have been documented by a large number of studies including Carhart (1997), Gruber (1996), Jensen (1968), and Malkiel (1995). Please see Berk and Green (2004) for a more complete survey.

[2]Fund return predictability is not actually consistent with the model of Berk and Green (2004). In addition to diminishing returns to scale, their model assumes that risk-adjusted expected returns are equal across funds of different sizes in equilibrium. However, our findings do suggest that there are diminishing returns to scale in the mutual fund industry.

[3]Other factors, such as marketing considerations, may also affect fund behavior.

[4]This benefit is presumably independent of how the combined holding is divided between funds in the family.

[5]Theoretically, fund families could avoid organizational diseconomies within a fund by setting up internal sub-funds that are managed independently and then marketing a combination of these sub-funds to the public as one investment product.

[6]Our matching procedure is similiar to the approach described in Wermers (2000).

[7]The Investment Company Act, 1940, section 5(b)1 defines a fund as diversified if no more than 5% of its assets is invested in any one company's securities and it holds no more than 10% of the voting shares in any one company. Thus, funds with fewer than 10 equity holdings, if diversified, must have less than half of assets under management allocated to equities.

[8]The apparent outperformance of the funds in the sample during recessions can be explained by the cash reserves maintained by mutual funds.

[9]Daniel et al. (1997) calculate fund styles from the same database, assigning all stocks to a size quintile, from 1 (small) to 5, then calculate the portfolio-weighted average quintile of stocks held by each fund. Averaging this style measure over all funds and all years gives an aggregate style measure of 3.97. The median stock will have a style measure of three, but the market-weighted average style is greater than three under this measure. The market-weighted style is the appropriate measure to use for comparison purposes when portfolio weights, not equal weights, are used to measure fund style.

[10]Lagged fund returns do not have any explanatory power and nearly identical results are obtained if $\log Flow_{it}$ is replaced with $\Delta \log TNA_{it}$.

[11]In unreported analogous panel regressions the coefficient estimates and the patterns of statistical significance are qualitatively similar.

[12]Kacperczyk, Sialm, and Zheng (2005) measure industry concentration as the sum, over industries, of the squared difference between a fund's portfolio weight on a given industry and that industry's market weight for 10 industries. They present evidence that greater industry concentration is associated with higher fund returns.

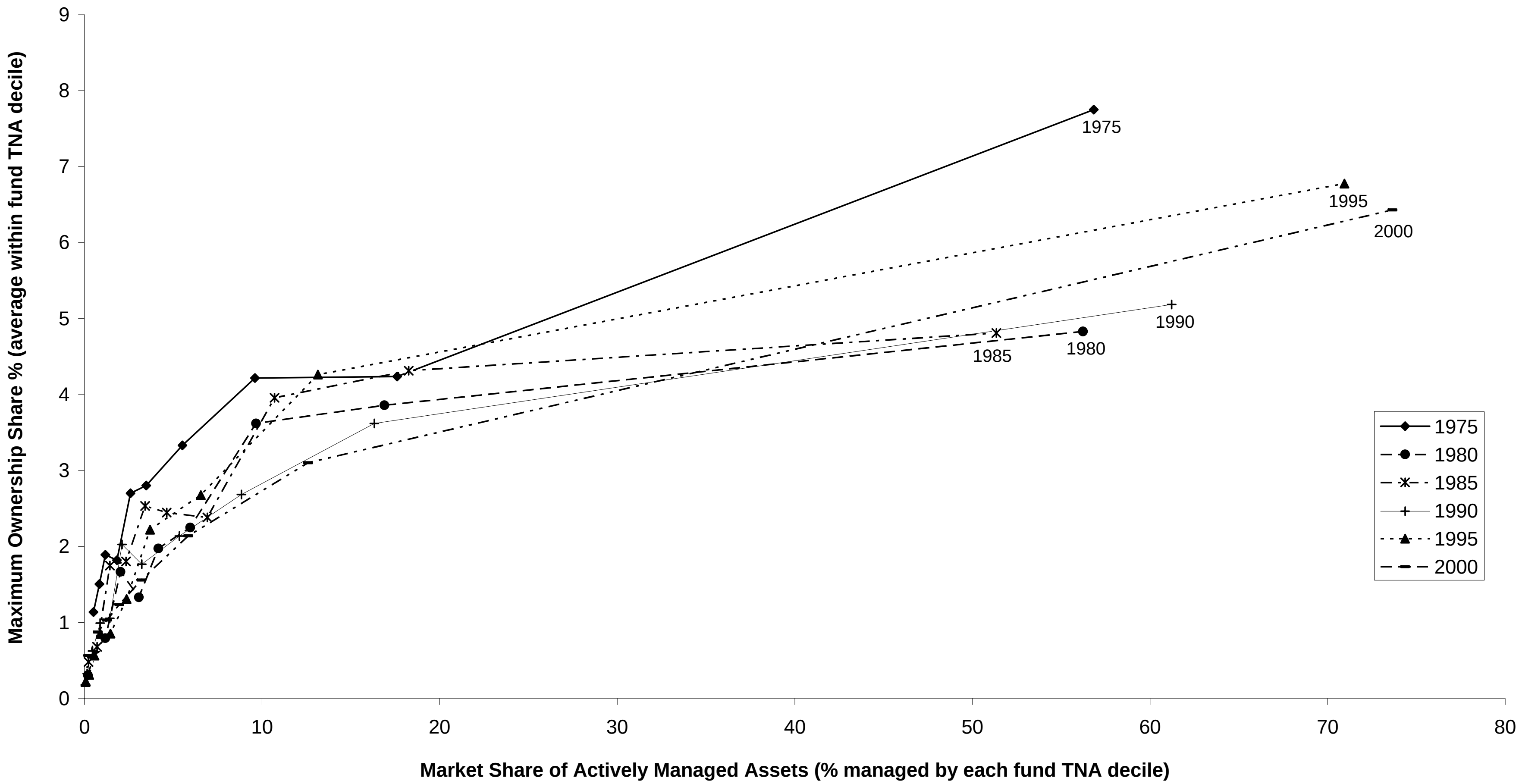

Figure 1. Maximum ownership share and market share. The figure plots maximum ownership share against market share for each total net assets (TNA) decile. Funds are sorted into deciles by TNA and the total TNA of each decile as a proportion of total TNA for all deciles is defined as the decile market share. Ownership share is defined to be the number of shares in a given firm owned by a fund divided by the number of shares outstanding. We plot the average maximum ownership share (equal-weighted across all funds in the same TNA decile) against decile market share for every fifth year in the sample.

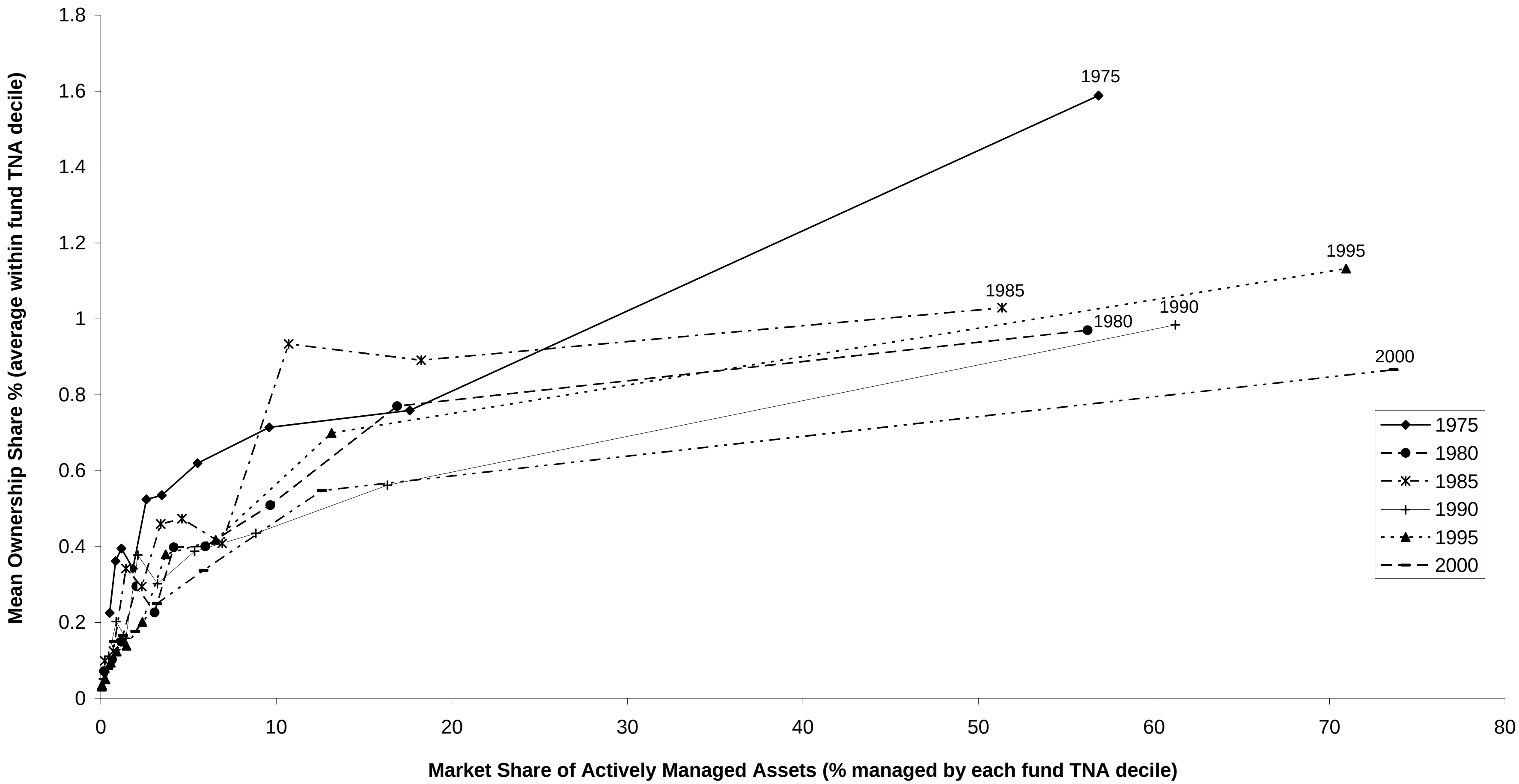

Figure 2. Mean ownership share and market share. The figure plots mean ownership share against market share for each total net assets (TNA) decile. Funds are sorted into deciles by TNA and the total TNA of each decile as a proportion of total TNA for all deciles is defined as the decile market share. Ownership share is defined to be the number of shares in a given firm owned by a fund divided by the number of shares outstanding. We plot the average mean ownership share (equal-weighted across all funds in the same TNA decile) against decile market share for every fifth year in the sample.

Table I
Summary Statistics for the Matched CDA-CRSP Sample

Column 1 is the year associated with the fund records. Column 2 reports the number of diversified equity funds in the CRSP mutual fund database that match with records of equity holdings in the CDA/Spectrum mutual fund database given the selection criteria discussed in the text. Column 3 reports the number of fund families classified by management company abbreviations in CDA/Spectrum. Column 4 reports the average size (average TNA in CRSP mutual fund database). Column 5 reports the combined assets under management for funds in the sample (column 2 multiplied by column 4). Column 6 reports assets under management for funds in the sample relative to the size of the stock market (Column 5 divided by the market capitalization of all stocks in CRSP). Column 7 reports the annual TNA-weighted mutual fund return after expenses. Column 8 reports the annual CRSP value-weighted stock market return.

Year	Number of matched funds	Number of fund families	Mean fund TNA ($mn)	Combined TNA ($mn)	Percentage of US stock	TNA weighted average return	CRSP market return
1975	253	138	149	37,671	4.69%	31.53%	37.35%
1980	314	153	145	45,638	3.09%	30.51%	33.23%
1985	358	146	282	100,919	4.25%	28.23%	31.46%
1990	553	203	344	190,035	5.87%	-5.82%	-6.03%
1995	1,170	418	671	784,924	11.03%	31.73%	35.73%
2000	1,421	478	1,438	2,043,522	12.70%	-5.12%	-10.97%
					Average	15.49%	16.78%

Table II

Basic Characteristics of Funds by Fund Size Quintile (selected years)

Table II presents statistics for funds sorted into quintiles using total net assets (TNA) under management. Column 1 is the selected year. Column 2 is the quintile (low TNA funds are in quintile 1). Column 3 reports the number of funds in each quintile. Column 4 reports the percentage of total TNA of all funds in the sample managed by funds in the specific quintile. Column 5 reports the mean TNA (in millions of US$) managed by funds in each quintile. Column 6 reports the mean number of distinct investments for funds in each quintile. Column 7 reports the mean of the average market capitalization (in billions of US$) of stocks held by a fund using portfolio weights for funds in each quintile. Column 8 reports the mean of the largest ownership share of each fund for funds in each quintile. Column 9 reports the mean of the average ownership share using equal weights within each fund for funds in each quintile.The CRSP row for each year reports the total number of stocks in the CRSP index and the weighted average market capitalization of all stocks in CRSP using market weights. Standard deviations are in parentheses.

Year	Size quintile	Number of funds	Percentage of all assets	Mean TNA ($mn)	Mean number of stocks	Mean w-avg. mkt. cap. ($bn)	Mean maximum share (%)	Mean average share (%)
1980	1	62	0.82	6.04	29.58	3.77	0.48	0.09
				(3.81)	(14.82)	(3.09)	(0.79)	(0.17)
	2	63	3.17	22.96	40.91	3.10	1.23	0.22
				(7.46)	(18.34)	(2.95)	(1.24)	(0.21)
	3	63	7.17	51.92	51.43	3.58	1.65	0.31
				(9.49)	(25.75)	(2.64)	(1.43)	(0.35)
	4	63	15.58	112.86	58.97	4.10	2.95	0.46
				(32.16)	(42.28)	(3.01)	(2.20)	(0.37)
	5	63	73.26	530.73	74.29	4.80	4.35	0.87
				(417.05)	(32.17)	(2.67)	(2.42)	(0.64)
	CRSP				4933	6.85		
1990	1	110	0.60	10.35	43.54	7.10	0.48	0.08
				(6.12)	(39.35)	(6.05)	(0.82)	(0.16)
	2	111	2.31	39.58	46.93	8.35	1.03	0.18
				(11.10)	(22.51)	(6.02)	(1.64)	(0.27)
	3	110	5.29	91.31	59.79	8.03	1.90	0.34
				(23.01)	(67.21)	(5.82)	(1.98)	(0.45)
	4	111	14.11	241.50	81.94	8.68	2.41	0.41
				(71.25)	(71.86)	(5.54)	(2.44)	(0.52)
	5	111	77.70	1330.21	121.49	9.65	4.41	0.77
				(1562.57)	(156.45)	(4.89)	(3.95)	(0.85)
	CRSP				6305	14.24		
2000	1	284	0.27	19.75	73.01	52.07	0.37	0.05
				(12.51)	(149.36)	(44.62)	(1.09)	(0.12)
	2	284	1.18	84.64	90.31	57.79	0.71	0.11
				(27.20)	(107.21)	(46.91)	(1.57)	(0.31)
	3	284	3.22	231.72	97.26	56.05	1.14	0.17
				(57.92)	(84.29)	(50.06)	(1.74)	(0.26)
	4	284	9.01	648.19	137.06	60.72	1.85	0.29
				(228.81)	(270.00)	(47.42)	(3.49)	(0.61)
	5	285	86.32	6189.40	143.88	67.49	4.77	0.71
				(9612.00)	(203.18)	(43.37)	(6.27)	(1.05)
	CRSP				7119	96.31		

Table III

Basic Characteristics of Funds by Fund Style Quintile (selected years)

Table III presents statistics for funds sorted into quintiles by style (the average market capitalization of stocks held by a fund using portfolio weights). Column 1 is the selected year. Column 2 is the style quintile (small-cap funds are in quintile 1). Column 3 reports the number of funds in each quintile. Column 4 reports the percentage of total TNA of all funds in the sample managed by funds in the specific quintile. Column 5 reports the mean TNA (in millions of US$) managed by funds in each quintile. Column 6 reports the mean number of distinct investments for funds in each quintile. Column 7 reports the mean of the average market capitalization (in billions of US$) of stocks held by a fund using portfolio weights for funds in each quintile. Column 8 reports the mean of the largest ownership share of each fund for funds in each quintile. Column 9 reports the mean of the average ownership share using equal weights within each fund for funds in each quintile.The CRSP row for each year reports the total number of stocks in the CRSP index and the weighted average market capitalization of all stocks in CRSP using market weights. Standard deviations are in parentheses.

Year	Style quintile	Number of funds	Percentage of all assets	Mean TNA ($mn)	Mean number of stocks	Mean w-avg. mkt. cap. ($bn)	Mean maximum share (%)	Mean average share (%)
1980	1	62	10.42	76.69	52.94	0.60	2.87	0.62
				(150.03)	(36.76)	(0.29)	(2.47)	(0.64)
	2	63	12.44	90.13	47.59	1.65	2.37	0.47
				(134.41)	(29.14)	(0.44)	(2.54)	(0.55)
	3	63	18.91	136.99	52.84	3.54	1.96	0.34
				(266.38)	(34.13)	(0.61)	(1.72)	(0.33)
	4	63	29.73	215.35	55.56	5.22	2.06	0.32
				(355.13)	(33.77)	(0.44)	(2.28)	(0.37)
	5	63	28.50	206.47	46.62	8.29	1.44	0.21
				(341.56)	(25.94)	(2.15)	(1.64)	(0.24)
	CRSP				4933	6.85		
1990	1	110	7.52	130.00	84.79	0.79	3.61	0.74
				(208.34)	(147.32)	(0.61)	(4.01)	(0.80)
	2	111	24.37	417.19	74.54	4.35	2.60	0.48
				(943.54)	(68.19)	(1.18)	(2.69)	(0.67)
	3	110	23.01	397.58	68.46	8.35	1.83	0.27
				(1238.40)	(65.46)	(1.12)	(2.09)	(0.38)
	4	111	22.30	381.71	57.61	11.69	1.47	0.19
				(732.46)	(39.40)	(1.00)	(2.15)	(0.26)
	5	111	22.80	390.31	68.72	16.57	0.74	0.11
				(813.01)	(89.26)	(2.75)	(1.09)	(0.17)
	CRSP				6305	14.24		
2000	1	284	5.59	402.45	149.77	1.46	3.23	0.58
				(784.76)	(282.62)	(0.68)	(4.99)	(0.93)
	2	284	14.07	1012.74	80.39	18.27	2.73	0.44
				(2255.14)	(73.92)	(11.83)	(4.47)	(0.81)
	3	284	35.02	2519.98	97.78	57.66	1.32	0.16
				(7474.13)	(112.61)	(10.25)	(3.08)	(0.35)
	4	284	19.79	1424.16	113.30	91.63	0.99	0.09
				(3219.56)	(200.46)	(9.86)	(2.72)	(0.21)
	5	285	25.52	1829.74	100.44	124.89	0.59	0.07
				(6838.15)	(134.40)	(14.01)	(1.96)	(0.15)
	CRSP				7119	96.31		

Table IV
Asset Growth and the Number of Holdings

Columns 1 through 8 report OLS coefficient estimates from panel regressions of the annual log growth rate for the number of stocks on various explanatory variables from 1976 until 2000. Each specification includes year fixed effects. For fund i at the end of year t, log $Flow_{it}$ is the log flow of new funds and is defined as the difference between the log growth rate for TNA and the log return for the fund between $t-1$ and t, TNA_{it} is the fund's total net assets under management, S_{it} is the number of distinct equity holdings in the fund's portfolio, log Own_{it} is the portfolio weighted average log ownership fraction in firms whose stocks are held by the fund, $High\text{-}TNA_{it-1}$ is an indicator variable that equals one if the fund's TNA is greater than the sample median at the end of year $t-1$, and $Large\text{-}cap_{it-1}$ is an indicator variable that equals one if the fund's style (the portfolio-weighted mean market capitalization of companies held) is greater than the median at the end of year $t-1$.

Δ is the difference operator and the dependent variable $\Delta\log S_{it}$ is the log growth rate of the number of stocks from $t-1$ to t. Because year $t-1$ information is needed for each observation the sample period is from 1976 to 2000 rather than 1975 until 2000. Each specification includes year fixed-effects. Standard errors are clustered by year and reported in parentheses. The OLS R^2 is reported in the last row.

Dependent variable: $\Delta\log S_{it}$ (change in log number of stocks from year $t-1$ to t for fund i)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log $Flow_{it}$	0.115	0.109	0.142	0.137	0.120	0.111	0.146	0.137
	(0.014)	(0.015)	(0.013)	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)
log S_{it-1}	-0.096	-0.106	-0.097	-0.108	-0.102	-0.107	-0.101	-0.108
	(0.005)	(0.006)	(0.005)	(0.006)	(0.002)	(0.007)	(0.006)	(0.006)
log Own_{it-1}					0.009	0.005	0.008	0.001
					(0.002)	(0.003)	(0.002)	(0.003)
$High\text{-}TNA_{it-1}$		0.036		0.038		0.026		0.037
		(0.006)		(0.006)		(0.008)		(0.009)
$High\text{-}TNA_{it-1}$ *log $Flow_{it}$		0.041		0.044		0.040		0.044
		(0.017)		(0.018)		(0.017)		(0.017)
$Large\text{-}cap_{it-1}$			-0.017	-0.021			-0.005	-0.019
			(0.007)	(0.007)			(0.007)	(0.008)
$Large\text{-}cap_{it-1}$ *log $Flow_{it}$			-0.055	-0.056			-0.052	-0.056
			(0.018)	(0.018)			(0.018)	(0.018)
R^2	0.091	0.095	0.093	0.098	0.094	0.096	0.095	0.098

Table V

Table V: Asset Growth and Average Ownership Shares

Columns 1 through 8 report OLS coefficient estimates from panel regressions of the annual change in the portfolio-weighted log ownership share on various explanatory variables from 1976 until 2000. For fund i at the end of year t, log Own_{it} is the portfolio-weighted average log ownership fraction in firms whose stocks are held by the fund, log $Flow_{it}$ is the log flow of new funds and is defined as the difference between the log growth rate for TNA and the log return for the fund between t-1 and t, TNA_{it} is the fund's total net assets under management, S_{it} is the number of distinct equity holdings in the fund's portfolio, $High\text{-}TNA_{it-1}$ is an indicator variable that equals one if a fund's TNA is greater than the sample median at the end of year t-1, and $Large\text{-}cap_{it-1}$ is an indicator variable that equals one if the fund's style (the portfolio-weighted mean market capitalization of companies held) is greater than the median at the end of year t-1.

Δ is the difference operator and the dependent variable Δlog Own_{it} is the change of the portfolio-weighted average log ownership share from t-1 to t. Because year t-1 information is needed for each observation the sample period is from 1976 to 2000 rather than 1975 until 2000. Each specification includes year fixed effects. Standard errors are clustered by year and reported in parentheses. The OLS R^2 is reported in the last row.

Dependent Variable: Δlog Own_{it} (change in portfolio-weighted average log ownership between year t-1 and t for fund i)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log $Flow_{it}$	0.718	0.733	0.704	0.716	0.716	0.732	0.702	0.715
	(0.043)	(0.042)	(0.042)	(0.042)	(0.042)	(0.043)	(0.042)	(0.043)
log S_{it-1}					0.045	0.026	0.044	0.025
					(0.010)	(0.011)	(0.010)	(0.011)
log Own_{it-1}	-0.041	-0.058	-0.039	-0.064	-0.045	-0.058	-0.043	-0.064
	(0.006)	(0.008)	(0.008)	(0.011)	(0.006)	(0.008)	(0.008)	(0.011)
$High\text{-}TNA_{it-1}$		0.115		0.130		0.102		0.117
		(0.016)		(0.025)		(0.017)		(0.026)
$High\text{-}TNA_{it-1}$ *log $Flow_{it}$		-0.028		-0.029		-0.029		-0.030
		(0.037)		(0.037)		(0.036)		(0.036)
$Large\text{-}cap_{it-1}$			0.014	-0.037			0.010	-0.034
			(0.024)	(0.030)			(0.023)	(0.030)
$Large\text{-}cap_{it-1}$ *log $Flow_{it}$			0.030	0.027			0.030	0.027
			(0.042)	(0.043)			(0.043)	(0.043)
R^2	0.260	0.266	0.261	0.267	0.263	0.267	0.263	0.267

Table VI

Effect of Fund Size and Diversification on Fund Returns

Columns 1 through 6 present the time-series averages of the OLS coefficients from cross-sectional regressions of risk-adjusted monthly fund returns (after expenses) from 1976 until 2001 on fund characteristics lagged one year. Beneath each point estimate we report *t*-statistics based on time-series variation in parentheses.

The risk-adjusted monthly return is the fund return minus each fund factor loading multiplied by its respective factor return. Fund factor loadings are estimated from the regression of the monthly portfolio return of the stocks held in the fund at the end of year *t-1* (using portfolio weights from CDA) on the Fama-French-Carhart four factors for the previous thirty-six months. log TNA_{it-1} is log total net assets. log Age_{it-1} is the log of the age of the fund according to CRSP. log $FAMTNA_{it-1}$ is the log of one plus total TNA held by all the other funds in the fund's family, excluding fund *i*. $Expense_{it-1}$ is the expense ratio for the fund. $Totload_{it-1}$ is the combination of front-end and back-end loads for the fund. ICI_{it-1} is the industry concentration index for the fund's portfolio from Kacperczyk et al. (2005) calculated at the end of year *t-1*. $1/S_{it-1}$ is the reciprocal of the number of stocks in the portfolio at the end of year *t-1*. The variables $cap2_{it-1}$ through $cap5_{it-1}$ are four indicator variables for the quintiles of market capitalization style (style is the portfolio-weighted average market capitalization of stocks held in the fund portfolio at the end of year t-1). The indicator $cap5_{it-1}$ is for large cap funds.

Dependent Variable: Risk-adjusted monthly return after expenses for each month in year *t* for fund *i*

	(1)	(2)	(3)	(4)	(5)	(6)
log $FAMTNA_{it-1}$			0.0001	0.0001	0.0001	0.0001
			(4.21)	(4.07)	(4.37)	(4.42)
$Expense_{it-1}$				-0.0662		-0.0688
				(2.68)		(3.06)
$Totload_{it-1}$				-0.0001		-0.0001
				(2.18)		(2.50)
log Age_{it-1}				0.0000		-0.0001
				(0.26)		(1.50)
ICI_{it-1}				0.0060		0.0058
				(1.62)		(1.80)
log TNA_{it-1}		-0.0001	-0.0002	-0.0003	-0.0006	-0.0006
		(1.33)	(2.68)	(3.42)	(3.29)	(3.39)
log $TNA_{it-1}*cap2_{it-1}$					0.0003	0.0001
					(1.44)	(0.71)
log $TNA_{it-1}*cap3_{it-1}$					0.0003	0.0004
					(1.66)	(1.89)
log $TNA_{it-1}*cap4_{it-1}$					0.0006	0.0007
					(3.40)	(3.64)
log $TNA_{it-1}*cap5_{it-1}$					0.0006	0.0007
					(3.10)	(3.64)
$1/S_{it-1}$	-0.0305	-0.0363	-0.0323	-0.0421	-0.0633	-0.0766
	(3.62)	(3.49)	(3.10)	(3.14)	(3.00)	(3.41)
$(1/S_{it-1})*cap2_{it-1}$					0.0252	0.0288
					(1.03)	(1.20)
$(1/S_{it-1})*cap3_{it-1}$					0.0329	0.0483
					(1.29)	(1.88)
$(1/S_{it-1})*cap4_{it-1}$					0.0641	0.0750
					(2.40)	(2.74)
$(1/S_{it-1})*cap5_{it-1}$					0.0550	0.0545
					(2.07)	(1.98)
$cap2_{it-1}$					-0.0013	-0.0007
					(1.16)	(0.56)
$cap3_{it-1}$					-0.0023	-0.0026
					(1.64)	(1.92)
$cap4_{it-1}$					-0.0039	-0.0042
					(2.96)	(3.22)
$cap5_{it-1}$					-0.0038	-0.0031
					(2.64)	(2.13)

Table VII

Effect of Fund Size and Diversification on Fund Returns (excluding small funds)

Columns 1 through 6 present the time-series averages of the OLS coefficients from cross-sectional regressions of risk-adjusted monthly fund returns (after expenses) from 1976 until 2001 on fund characteristics lagged one year. Beneath each point estimate we report *t*-statistics based on time-series variation in parentheses. The regression specifications are identical to the specifications in Table VI, but the sample only includes funds with TNA of more than 100 million dollars at the end of year *t-1*.

The risk-adjusted monthly return is the fund return minus each fund factor loading multiplied by its respective factor return. Fund factor loadings are estimated from the regression of the monthly portfolio return of the stocks held in the fund at the end of year *t-1* (using portfolio weights from CDA) on the Fama-French-Carhart four factors for the previous thirty-six months. log TNA_{it-1} is log total net assets. log Age_{it-1} is the log of the age of the fund according to CRSP. log $FAMTNA_{it-1}$ is the log of one plus total TNA held by all the other funds in the fund's family, excluding fund *i*. $Expense_{it-1}$ is the expense ratio for the fund. $Totload_{it-1}$ is the combination of front-end and back-end loads for the fund. ICI_{it-1} is the industry concentration index for the fund's portfolio from Kacperczyk et al. (2005) calculated at the end of year *t-1*. $1/S_{it-1}$ is the reciprocal of the number of stocks in the portfolio at the end of year *t-1*. The variables $cap2_{it-1}$ through $cap5_{it-1}$ are four indicator variables for the quintiles of market capitalization style (style is the portfolio-weighted average market capitalization of stocks held in the fund portfolio at the end of year t-1). The indicator $cap5_{it-1}$ is for large cap funds.

Dependent Variable: Risk-adjusted monthly return after expenses for each month in year *t* for fund *i*						
	(1)	(2)	(3)	(4)	(5)	(6)
log $FAMTNA_{it-1}$			0.0001	0.0001	0.0001	0.0002
			(2.72)	(2.99)	(2.67)	(3.03)
$Expense_{it-1}$				-0.1695		-0.1576
				(2.92)		(2.77)
$Totload_{it-1}$				-0.0001		-0.0001
				(2.10)		(2.01)
log Age_{it-1}				0.0001		0.0000
				(0.53)		(0.27)
ICI_{it-1}				0.0040		0.0063
				(0.82)		(1.25)
log TNA_{it-1}		-0.0001	-0.0002	-0.0004	-0.0007	-0.0011
		(1.22)	(2.01)	(3.57)	(1.92)	(2.86)
log TNA_{it-1} $*cap2_{it-1}$					0.0003	0.0005
					(0.71)	(1.06)
log TNA_{it-1} $*cap3_{it-1}$					0.0008	0.0012
					(1.96)	(2.67)
log TNA_{it-1} $*cap4_{it-1}$					0.0005	0.0008
					(1.21)	(1.90)
log TNA_{it-1} $*cap5_{it-1}$					0.0006	0.0007
					(1.29)	(1.62)
$1/S_{it-1}$	-0.0594	-0.0657	-0.0581	-0.0645	-0.1149	-0.1246
	(3.51)	(3.59)	(3.11)	(3.13)	(2.86)	(2.84)
$(1/S_{it-1})*cap2_{it-1}$					0.0471	0.0428
					(0.86)	(0.75)
$(1/S_{it-1})*cap3_{it-1}$					0.1060	0.1180
					(2.54)	(2.61)
$(1/S_{it-1})*cap4_{it-1}$					0.1250	0.1303
					(2.41)	(2.33)
$(1/S_{it-1})*cap5_{it-1}$					0.0992	0.0823
					(2.17)	(1.68)
$cap2_{it-1}$					-0.0021	-0.0031
					(0.71)	(0.99)
$cap3_{it-1}$					-0.0063	-0.0085
					(2.29)	(2.87)
$cap4_{it-1}$					-0.0045	-0.0065
					(1.55)	(2.16)
$cap5_{it-1}$					-0.0045	-0.0052
					(1.46)	(1.63)

Table VIII

Basic Characteristics of Families by Family Size Quintile (selected years)

Table VIII presents some basic statistics on families of funds in the sample sorted into quintiles by family size (the combined TNA of all funds with the same management company identifier in CDA). Column 1 is the selected year. Column 2 is the quintile (families with low combined TNA are in quintile 1). Column 3 reports the number of families in each size quintile. Column 4 reports the total TNA managed by all families in the quintile as a percentage of total sample TNA. Column 5 reports the average TNA managed by a family in each quintile. Column 6 reports the average number of funds in each family. Column 7 reports the maximum number of funds in a family for each quintile. Column 8 reports the average number of stocks held by all the funds in one family combined. Standard deviations are reported in parentheses.

Year	Family size quintile	Number of families	Percentage of all assets	Mean TNA ($mn)	Mean number of funds	Maximum number of funds	Mean number of stocks in family
1980	1	30	0.36	5.58	1.07	2	27.93
				(3.93)	(0.25)		(13.79)
	2	31	1.77	26.21	1.35	3	50.13
				(9.15)	(0.61)		(0.61)
	3	30	4.07	62.27	1.57	3	63.17
				(18.58)	(0.73)		(40.50)
	4	31	12.68	187.63	2.10	5	82.74
				(65.51)	(1.22)		(54.01)
	5	31	81.12	1200.47	4.32	14	215.71
				(1007.54)	(2.91)		(149.77)
1990	1	40	0.25	11.83	1.18	2	35.10
				(7.98)	(0.38)		(19.06)
	2	41	1.09	51.42	1.54	4	56.39
				(17.31)	(0.81)		(23.68)
	3	40	2.95	141.84	2.20	5	98.73
				(42.62)	(1.04)		(95.96)
	4	41	10.08	473.29	2.88	6	182.68
				(146.47)	(1.35)		(227.85)
	5	41	85.64	4022.24	5.93	18	313.59
				(5899.29)	(3.55)		(235.89)
2000	1	95	0.07	14.91	1.18	3	66.32
				(9.78)	(0.46)		(168.99)
	2	96	0.34	75.68	1.32	4	101.18
				(24.88)	(0.64)		(216.68)
	3	95	1.12	250.63	1.82	5	125.65
				(91.79)	(1.05)		(147.89)
	4	96	5.74	1273.14	3.30	9	278.49
				(616.51)	(1.83)		(304.97)
	5	96	92.74	20585.43	7.23	38	500.25
				(47665.81)	(5.41)		(518.77)

Table IX

Fund Family Behavior and Family TNA Growth

Table IX reports OLS coefficient estimates for the behavior of family TNA growth with fund family size and family diversification. Funds are defined as being in the same family if they have the same management company abbreviation in CDA. The dependent variables are $\Delta\log F_{ft}$, the log growth of the number of funds in family f, and $\Delta\log S_{ft}$, the log growth of number of distinct stock holdings in all funds in the same family (count a specific stock only once if it is held by different funds in the same family). Family TNA is the sum of fund TNA over all funds in the family. Families are sorted into family size quintiles at the end of each year *t-1* based on Family TNA. Independent variables are the log growth rate of Family TNA and the lagged log level variable of which the dependent variable is the current first difference. Because year *t-1* information is needed for each observation the sample period is from 1976 to 2000 rather than 1975 until 2000. Each specification includes year fixed effects. Standard errors are clustered by year and reported in parentheses. The OLS R^2 is reported in the last column.

Equation: $\Delta X_{ft} = \alpha + \delta_t + \beta_1 \Delta\log FAMTNA_{ft} + \beta_2 X_{ft-1} + \varepsilon_{ft}$ for family f, year t

Family size quintile	ΔX_{ft}	$\Delta\log FAMTNA_{ft}$	X_{ft-1}	R^2
Smallest	$\Delta\log F_{ft}$	0.084	-0.138	0.15
		(0.021)	(0.060)	
	$\Delta\log S_{ft}$	0.165	-0.224	0.22
		(0.036)	(0.028)	
2	$\Delta\log F_{ft}$	0.261	-0.160	0.33
		(0.038)	(0.038)	
	$\Delta\log S_{ft}$	0.290	-0.197	0.25
		(0.040)	(0.030)	
3	$\Delta\log F_{ft}$	0.346	-0.187	0.46
		(0.034)	(0.031)	
	$\Delta\log S_{ft}$	0.378	-0.114	0.39
		(0.023)	(0.019)	
4	$\Delta\log F_{ft}$	0.470	-0.152	0.53
		(0.032)	(0.032)	
	$\Delta\log S_{ft}$	0.435	-0.103	0.40
		(0.025)	(0.026)	
Largest	$\Delta\log F_{ft}$	0.477	-0.149	0.75
		(0.034)	(0.047)	
	$\Delta\log S_{ft}$	0.400	-0.122	0.61
		(0.041)	(0.032)	

Table X
Fund Diversification and Family Structure

Columns 1 through 4 report OLS coefficient estimates from panel regressions of the annual log growth rate for the number of stocks on family characteristics from 1976 until 2000. For fund *i* at the end of year *t*, log $Flow_{it}$ is the log flow of new funds and is defined as the difference between the log growth rate for TNA and the log return for the fund between *t-1* and *t*. TNA_{it} is the fund's total net assets under management. $High\text{-}TNA_{it}$ is an indicator variable equal to one if the fund's TNA is above the sample median for that year. $Large\text{-}family_{it}$ is an indicator variable equal to one if the fund's family has the most other funds of all the families with a fund in the fund's TNA quintile. log S_{it-1} is the lagged log number of stocks in the fund's portfolio. Δ is the difference operator and the dependent variable Δlog S_{it} is the log growth rate of the number of stocks from *t-1* to *t*.

Because year *t-1* information is needed for each observation the sample period for columns 1 and 2 is from 1976 to 2000 rather than 1975 until 2000. In columns 3 and 4, the sample period begins in 1988 because fund families are not really prevalent in the 1970s and early 1980s. Each specification includes year fixed effects. Standard errors are clustered by year and reported in parentheses. The OLS R^2 is reported in the last row.

Dependent variable: Δlog S_{it} (change in log number of stocks from year *t-1* to *t* for fund *i*)				
	1976-2000		1988-2000	
	(1)	(2)	(3)	(4)
log $Flow_{it}$	0.114	0.114	0.099	0.099
	(0.015)	(0.015)	(0.014)	(0.014)
$Big\text{-}Family_{it-1}$	-0.007	0.021	-0.014	0.017
	(0.014)	(0.014)	(0.022)	(0.022)
$High\text{-}TNA_{it-1}$	-0.026	0.027	-0.031	0.015
	(0.007)	(0.010)	(0.008)	(0.009)
$High\text{-}TNA_{it-1}$ *$Big\ Family_{it-1}$	-0.003	0.035	-0.005	0.034
	(0.022)	(0.021)	(0.027)	(0.026)
$Big\text{-}Family_{it-1}$ *log $Flow_{it}$	0.012	0.021	-0.002	0.008
	(0.041)	(0.038)	(0.045)	(0.043)
$High\text{-}TNA_{it-1}$ *log $Flow_{it}$	0.076	0.062	0.080	0.074
	(0.036)	(0.036)	(0.042)	(0.042)
$High\text{-}TNA_{it-1}$ *$Big\text{-}Family_{it-1}$ *log $Flow_{it}$	-0.152	-0.076	-0.237	-0.192
	(0.114)	(0.119)	(0.106)	(0.109)
log S_{it-1}		-0.103		-0.097
		(0.007)		(0.006)
R^2	0.041	0.093	0.040	0.090