The Impact of Mutual Fund Family Membership on Investor Risk

Edwin J. Elton

Martin J. Gruber

T. Clifton Green*

May 2004

Abstract

Many investors confine their mutual fund holdings to a single fund family, either for simplicity or through restrictions placed by their retirement savings plan. We find evidence that mutual fund returns are more closely correlated within fund families, which reduces the benefits of investor diversification. The increased correlation is due primarily to common stock holdings, but is also more generally related to families having similar exposures to economic sectors or industries. Fund families also show a propensity to focus on high risk or low risk strategies, which leads to a greater dispersion of risk across restricted investors.

* Elton, eelton@stern.nyu.edu, and Gruber, mgruber@stern.nyu.edu, are Nomura Professors of Finance at the Stern School of Business at New York University. Green, clifton.green@emory.edu, is Assistant Professor of Finance at Goizueta Business School, Emory University.

Individuals frequently place all of their mutual fund investments with one family of mutual funds. Employer sponsored retirement plans often necessitate this behavior by limiting fund choices to the offerings of a single family.[1] Load fees, which are typically not charged when moving switching funds within a family, also encourage family loyalty. On a more basic level, investors may restrict their attention to one family to help narrow the search process and simplify record keeping. In this article, we examine whether the propensity of investors to confine their investments to a single fund family influences the risk characteristics of their portfolios.

There are several reasons to expect that funds may be more similar inside than outside fund families. Portfolio managers within families are likely to have access to the same research analysis produced either by internal analysts or by a particular set of external research firms. Many families also have a prescribed investment style that influences the type of securities they hold. A common view on individual companies could lead to similar stock holdings across portfolios with even different objectives. In addition, a family's relationship with an investment brokerage firm could also lead to common holdings of new offerings.

Fund similarities within families may also arise from macro level influences. Portfolio managers may begin the security selection process with an economic forecast that is shared by other fund managers within the firm. For example, a family's portfolio managers may sit on a strategy committee that shares insights regarding the overall economy. A common family-wide economic outlook could result in similar exposures to various economic sectors. Commonalities related to both sectors and individual securities will be greater whenever one portfolio management team manages multiple funds within a family.

[1] For example, Elton, Gruber and Blake (2004) study over six hundred 401K plans and find that most restrict fund choices to one family.

In this study we study the size and determinants of increased fund return correlations within families. The influence that family membership can have on fund returns can be illustrated by considering the Eaton Vance Special Equities Fund and Eaton Vance Tax Managed Small Cap Growth Fund. The Investment Company Data Inc. (ICDI) classifies the special equity fund as Aggressive Growth and the small cap fund as Long-Term Growth. They have separate prospectuses and are categorized on the family's web site in different groups of funds. The prospectus of the Special Equities Fund states the objective as "Growth in Capital" whereas the Managed Small Cap Growth Fund's stated objective is "A diversified fund seeking long-term after-tax returns by investing in emerging growth companies."

An investor looking at standard materials from Eaton Vance would have little reason to believe the funds were similar. However, the correlation between their monthly returns over a five-year period is 0.995. Both funds employ the same portfolio manager. The rank order of the major holdings in the two funds is the same, with only small differences in percentages invested. The rank order of the amount invested in each economic sector is the same, with only small differences in the percentages invested in any sector. Thus, an investor hoping to diversify by buying shares of both funds would be disappointed with the resulting risk profile.

Our analysis suggests that investors who limit their investments to one fund family hold riskier portfolios than those who diversify across families. Both within and across objectives, fund return correlations are significantly higher inside than outside fund families. An examination of fund holdings, combined with a factor model to characterize fund returns, reveals that roughly two thirds of the increase in return correlation is related to common stock holdings with the rest attributable to similar exposures to broad economic factors. The extent of overlap in stock holdings is surprising. Depending on the objective group being considered, as much as 34% of total

net assets consists of stocks held in common, with an across-objective median of 17% inside the family compared to 8% outside the family.

We also find evidence that fund families show a propensity to focus on either high risk or low risk strategies. Within each objective, families are significantly more likely to have funds with standard deviations higher or lower than the median standard deviation for that classification. While this phenomenon does not increase the risk of an average investor's portfolio, it does increase the distribution of risk across investors. The increased dispersion raises the probability of having very high risk by investing exclusively in one family. Taken together, our results indicate that confining mutual fund investments to one family has a detrimental effect on investor risk that is statistically and economically meaningful. Portfolios of funds within families result in greater overall risk and greater risk clustering than similar portfolios created from funds across families.

Massa (2003) and Mamaysky and Spiegel (2001) develop models of the mutual fund industry in which investors' heterogeneous tastes provide incentives for families to compete through product differentiation. These studies, as well as Khorana, and Servaes (2003), offer empirical evidence that families increase market share by adding funds with different objectives.[2] Our findings indicate that fund proliferation within families does not provide the same level of diversification benefits as combining funds across families.

The paper is organized as follows: Section I discusses the sample and the data sources we use in the study. Section II examines the correlation between fund returns within and between fund families, where funds are grouped according to standard objective classifications. In Section III we begin our examination of the determinants of the increased correlation by applying factor models to remove the impact of various economic sectors. This allows us to examine return correlation due to sector and security

[2] In other research on fund families, Ivkovich (2001) and Nanda, Wang, and Zheng (2003) find evidence that outperforming funds have spillover effects on family cash inflows. Das and Sundaram (2002) study the impact of different fee structures on investor welfare.

bets as opposed to macro market bets. Section IV examines the actual security holdings of funds to examine the extent to which higher correlation within a fund family is due to holding the same securities. In Section V we study the propensity of families to engage in high or low risk strategies by studying similarities in fund standard deviations. Section VI offers conclusions and implications for investors.

1. The Data

The principal source of our data is the CRSP Mutual Fund Database.[3] Our initial sample is all fund families that existed in January 1998 and the mutual funds that were part of these families. We use objective classifications from Investment Company Data Inc. (ICDI) to categorize funds and eliminate the following: specialized funds, index funds, international funds, money market funds, single state municipal funds, precious metal funds, sector, and utility funds. This left us with funds in eleven ICDI objective categories. From this list we eliminate duplicate funds (versions of the fund that differ only in the expenses charged), which are of two varieties. First, many funds have multiple share classes related to different fee structures and we eliminate all but the class with the longest history. Second, many families offer the same fund to institutional investors or to financial planners under different names, and we eliminate the duplicate funds. Finally, we perform a detailed examination of pairs of funds that were highly correlated to ensure that duplicate funds are removed from the sample.

After removing duplicate funds, we then eliminate all fund families with a single remaining fund. For this remaining set of families we draw a final sample by randomly selecting one in three fund families while maintaining the same distribution of families in terms of the number of mutual funds offered. The resulting sample consists of 988 unique

[3] CRSP database suffers from omission bias, a form of survivorship bias. See Elton, Gruber and Blake (2001). Survivorship bias is not important for this study because we are looking at diversification at a point in time and we have data on all funds at that time.

funds from 100 different families. Table 1 shows the number of funds within four broad objective categories (stock, combination, high-yield, and bond), and eleven subcategories.[4] The vast majority of families offer a stock fund, either aggressive and/or long-term growth. Most families also offer at least one combination fund and one bond fund. Across objectives, the median number of funds per family is six. The largest family in the sample had 85 distinctly different funds.

For each fund we draw monthly returns for up to five years starting January 1998. In what follows, if a fund ceased to exist we calculated correlations over the common period (none had less than a year's data). In addition to data on fund families and returns, we also collect data on fund stock holdings from Thompson Financial Services' Mutual Fund Holdings database. We obtain portfolio holdings for funds that report their holdings for December of 2000.

II. Correlation Within and Between Fund Families

The first attribute of fund family risk we explore is fund return correlation. We calculate correlations for each pair-wise combination of fund objectives. Specifically, for each fund within a fund family we compute the correlations with all other funds in the family with a given objective and the correlation with funds outside the family with the same objective. For example, when calculating the average correlation within the family between aggressive growth and long-term growth funds, we also calculate the average correlation between an aggressive growth (long-term growth) fund within a family and long-term growth (aggressive growth) fund from outside the family. We then average these results across all families. We calculate statistical significance using two methods: a

[4] Subsequently, the aggregate classification "stock" will refer to funds with aggressive growth or long-term growth ICDI objectives; "combination" will refer to funds with both stocks and bonds in their portfolios as designated by the objectives Total Return, Growth and Income, Balanced, and Income; and "bond" will refer to Ginnie Mae, High Quality Bond, High Quality Municipal Bond, or Government Securities funds.

two-sample t-test of difference in mean correlations, and a one-sided binomial test that the proportion of families with greater within-family correlations is greater than 0.5.

The results are presented in Table 2. The table documents a pattern of increased correlation with families. For example, consider combining a stock fund with a combination fund. The fifth row of Table 2 shows that 78 fund families offered at least one stock fund and one combination fund. The average correlation between stock funds and combination funds is 0.757 if they are inside a family and 0.709 if they are from two separate families. For 67% of the families, selecting stock and combination funds from inside the family results in a higher correlation than selecting from two different families.

For each of the broad objective pairs shown in Table 2, within-family correlations are higher than between-family correlations.[5] Combining funds into stock-stock, combination-combination, and stock-combination pairs results in statistically significant higher return correlations within families than outside families using both t-tests and the binomial test. The influence of fund families on return correlations is weaker among bond funds. None of the correlations involving bond funds are statistically significant according to t-tests.

When funds are grouped according to more narrowly defined objectives, we find fourteen of the correlation differences are statistically higher within families at the one percent level using the t-test. Using the binomial test, eleven are statistically significant at the one percent level and fifteen are significant at the five percent level. The results for bond funds remain weak after partitioning funds into the more narrowly defined objective categories.

The correlations reported in Table 2 are generally reasonable in magnitude. The correlation between two stocks funds is higher than the correlation between a stock and a

[5] In the interest of space we omit several groupings with a smaller likelihood of commonality, such as Ginnie Mae–Aggressive Growth. In general the omitted objective pairs show higher correlation inside than outside families, although the differences tend to be small and none are statistically different from zero.

combination fund, which in turn is higher than the correlation between a stock and a high yield bond fund, which is higher than the correlation between a stock and a bond fund. The high correlation between combination funds is somewhat surprising. We show later that combination funds hold the highest percentage of stocks in common both for funds inside and outside the family, which may reflect a similar equity objective (stability of return). The high correlation among high yield bond funds is also intuitive, due to the relative homogeneity of strategy across funds. The correlation between bond and stock funds and bond funds and combination funds is negative, reflecting the correlation between stocks and bonds during the sample period.

As a robustness check, we examine whether the results are sensitive to the method used to classify objectives. Brown and Goetzmann (1997) find evidence that funds classify their objective in a strategic way that reduces the accuracy of reported classifications. Although this is less of a concern for our broad objective measures, we also group funds into eleven style categories based on cluster analysis approach similar to Brown and Goetzmann (1997). Although some of the funds are reshuffled into different categories, the difference in correlations within and between groups are very similar to the results shown in Tables 2 and are not reported for brevity. We also examine whether the correlation differences are sensitive to the size of the family. Grouping families into categories based on the number of funds in the family results in positive correlation differences for each group, with no significant differences between them.

In order to evaluate the economic significance of the higher correlation within families, we consider the following exercise. Assume an investor holds a fund with a particular objective, and she is considering adding one or two new funds to her portfolio.[6] The investor can add these funds from inside or outside the fund family. For each new outside fund, we calculate the number of new inside funds that would need to be added to

[6] The Investment Company Institute reports that the median (mean) number of stock funds held by individual investors in 2002 is 3 (5).

arrive at the same level of portfolio risk. We make two simplifying assumptions. We first assume that equal amounts are invested in each fund. This $1/N$ approach has empirical support in the studies of Benartzi and Thaler (2001) and Liang and Weisbenner (2002). Our second simplifying assumption is that all funds within an objective classification have the same variance, which we measure as the average across all funds with that objective. We then apply the standard formula for portfolio variance:

$$\sigma_p^2 = \sum_{i=1}^{N}\left(\frac{1}{N}\right)^2 \sigma_i^2 + \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N}\left(\frac{1}{N}\right)^2 \sigma_{ij}\,, \tag{1}$$

where N is the number of funds, σ_p^2 is the variance of the portfolio, σ_i^2 is the average fund variance, and σ_{ij} is the average covariance between funds.

Let N_1 be the number of funds currently held and N_2 be the number of funds to be added. Let ρ_F be the correlation between the funds that are currently within the portfolio and ρ_N be the correlation for the new funds, which takes on a different value depending on whether the new funds are from inside or outside the family (taken from Table 2). If the new funds are added from the same objective, Equation (1) simplifies to:

$$\sigma_p^2 = \sigma_i^2\left[\frac{1}{N_1+N_2} + \left(\frac{1}{N_1+N_2}\right)^2\left[N_1\left(N_1-1\right)\rho_F + 2N_1N_2\rho_N + N_2\left(N_2-1\right)\rho_N\right]\right] \tag{2}$$

If the funds added are in different classifications (e.g., combination being added to stock), then no informative simplified formula is available and equation (1) is used directly.

The results of this analysis are presented in Table 3. The table makes it clear that diversification within a fund family is less advantageous than diversification across families. If an investor owned one stock fund and was considering adding two more funds outside the family, she would have to add four internal stock funds to ensure the portfolio was not more risky. The results for the other groupings have the same pattern. When adding two combination funds to an existing combination fund, the investor would

have to add 17 funds inside the family to achieve the same level of risk as adding two funds outside the family. Taken together, the evidence suggests that investors or savings plans that limit investment to one fund family is giving up a significant amount of the benefits of diversification.

III. What explains the higher correlation?

In this section we use a number of diagnostic approaches to examine the portfolio management activities that lead to the increased fund return correlation within families. We begin with a macro level approach. If portfolio managers within a family begin the security selection process with a shared economic forecast, we may expect similar exposures to different economic factors. We examine this hypothesis with a number of multi-index models, beginning with a two-factor model.

A. Two-Index Model – Sensitivity to Bonds and Stocks

Combination funds own both bonds and stocks, stock funds frequently own some bonds, and bond funds often contain some securities with stock-like attributes. Thus, we begin with a two-index model where stock returns are measured using the value weighted CRSP index and bond returns are measured using the Merrill Lynch aggregate U.S. Corp/Gov/Mortgage bond index. For each fund in our sample we estimate a least squares regression on five years of monthly data to estimate the following relationship:[7]

$$R_i - R_F = \alpha_i + B_{is}\left(R_s - R_F\right) + B_{iB}\left(R_B - R_F\right) + e_i \qquad (3)$$

Where R_i is the return of fund *i*, R_F is the riskless rate, R_M is the return on the stock index, R_B is the return on the bond index, B_{is} and B_{iB} are the sensitivity of fund returns to the stock and bond index, α_i is the non-market return, and e_i is a random error. Under the two-index model, the correlation between two funds, *i* and *j* is given by:

[7] For the regressions we require the fund to have at least 36 monthly return observations.

$$\frac{\operatorname{cov}\left(R_i R_j\right)}{\sigma_i \sigma_j} = \frac{B_{is} B_{js} \sigma_s^2 + B_{iB} B_{jS} \operatorname{cov}(SB) + B_{iS} B_{jB} \operatorname{cov}(SB) + B_{iB} B_{jB} \sigma_B^2 + E\left(e_i e_j\right)}{\sigma_i \sigma_j} \quad (4)$$

where σ_S^2 and σ_B^2 are the variance of the stock and bond indexes and cov(*SB*) is the covariance between indexes, σ_i^2 and σ_j^2 are the standard deviation of funds *i* and *j*, and $E\left(e_i e_j\right)$ is the covariance of the fund return residuals.

The above expression separates the correlation between funds into two parts, the correlation due to systematic movements and the part due to residual movements. This decomposition allows us to examine how much of the higher correlation within a family is due to systematic market effects and how much is due to residual effects. Residual correlation can come about because two funds hold the same securities or because they are sensitive to similar factors not captured by the two-factor model. For example, a family may employ similar style choices such as emphasizing small stocks or large stocks or have a similar sensitivity to a particular industry factor such as technology stocks.

An increase in systematic correlation would come about if funds in the same family have similar portfolio sensitivities to bonds and stocks. For example, if the average combination fund is equally invested in stocks and bonds, but a particular family chooses to hold 70% in bonds, we would expect to observe higher systematic within-family correlation. The average difference in within-family correlation compared to between-family correlation due to residual correlation is the difference in the value of $E\left(e_i e_j\right)/\sigma_i \sigma_j$ for the two groups.

In Table 4 we examine the within- and between-family correlation due to residual commonality for the pairs of objectives where within-family correlation is higher than between-family correlation at a statistically significant level. We start by examining the aggregate groups from Table 2, for the two-index model. As shown in Table 4, Panel A the contribution to overall correlation from residual correlation is higher for two funds in the same family than when funds are in the two different families. The differences in

residual correlation are significant for the same three broad objective cases (stock-stock, stock-combination, and combination-combination) where difference in overall correlation were significant in Table 2. If one compares the differences in correlation in Table 4 with the differences in correlation in Table 2, then it is apparent that the higher overall within-family correlation is almost completely due to higher residual correlation. For the three aggregate pairs where the differences are significant the percentage of the overall differences in correlation due to differences in residual correlation are 110% (stock-stock), 100% (stock-combination, 104.3% (combination-combination).

A similar pattern exists for the more narrowly defined objective categories. The residual correlation accounts for more than 80% of the difference in within-family and between-family correlation except for aggressive growth with growth and income (where it accounts for 66%), long-term growth with growth and income (75%), and balanced with balanced (48%). Note that the only pairings where systematic influences have an important influence on correlation differences are pairings involving combination funds. This implies that one of the reasons these funds have higher within-family correlation is that they make similar choices concerning the split between stocks and bonds.

B. Multi-Index Models

Panel B of Table 4 shows the results for a six-factor model, which adds the Fama-French size and value factors (Fama and French, 1992), and decomposes the bond factor into three separate bond indexes (government, mortgage-backed, and high yield). The table indicates that higher residual correlation within a family is still an important component of the overall increase in correlation, but its relative importance falls. For the four cases shown in Table 2 where within-family correlation was significantly higher than between-family correlation, the percentage of the overall difference due to residual correlation from a six-index model was 78% (stock-stock), 48.3% (stock-combination), 58.7% (combination-combination), or an average of 62%. Comparing Panels A and B,

about 41% of the difference in the percentage of overall difference in the residual correlation between within- and between-family funds is explained by common factors beyond market factors.

The same general pattern exists when we use the more narrowly defined ICDI classifications. When pairing aggressive growth with aggressive growth, 81% of the additional correlation within families is due to residual correlation. When grouping aggressive growth with growth and income, 36% is due to residual correlation, and when grouping balanced with balanced, 33% is due to residual correlation. For the remaining categories, roughly 50% of the difference in correlation is due to residual risk. While some of the increased correlation is due to a common sensitivity to non-market factors, the residual is still an important component.[8]

If we have successfully captured all of the relevant factors, then the remaining correlation in residuals is due to common holdings. In addition, some of the effect of common holdings may be captured in the loadings to non-market factors. Thus, it is worthwhile to examine the effect of common holdings directly.

IV. Common Holdings

We now examine the extent to which common holdings of individual stocks translate into increased return correlations within fund families. We first document the amount of common holdings and then relate this to fund return correlations.

A. Difference in Common Holdings

The first question to examine is whether funds in the same family hold more securities in common than funds in different families. The simplest measure of common

[8] We also fit an eight-factor model that uses five industry portfolios and the three bond indexes and find similar results.

holdings for two funds is to sum the minimum fraction of the portfolio held in any stock *i* between the two funds or:

$$\mathrm{COM}(A,B) = \sum_i \min\left(X_{Ai}, X_{Bi}\right). \tag{5}$$

where X_{Ai} is the fraction of fund *A*'s portfolio invested in stock i, and X_{Bi} is the fraction of fund *B*'s portfolio invested in stock *i*.

For most mutual funds included in the Thompson Financial database, the aggregate amount invested in stocks does not equal 100% of total net assets. The principal reason for this is that most mutual funds hold some cash. Certain mutual funds such as balanced funds may hold a large fraction of their assets in bonds and the Thompson database only includes stock holdings. A second possible reason is that some small stock holdings may not be included in the Thompson database. Equation (5) can understate the impact of common holdings for it assumes that there is no impact from common stocks omitted from the Thompson database and there is no impact (extra correlation) due to bonds held in common. The effect of these omissions on return correlations should be small, both because the Thompson database contains a large fraction of common stock holdings and because the correlation between pairs of bonds is so high that common holdings does not cause much of an increase in correlation. The reason for this will be clear shortly when we examine how common holdings affect correlation.

Nevertheless, in order to clarify the extent of common holdings, we formulate a second measure that expresses the holdings as a fraction of the total identifiable amount of common stock held in the portfolio so the percentages add to 100% as follows:

$$\mathrm{COM}'\left(A,B\right) = \sum_i \min\left(\frac{X_{Ai}}{\sum X_{Ai}}, \frac{X_{Bi}}{\sum X_{Bi}}\right) \tag{6}$$

Which of these measures is more accurate depends on the proportion of holdings, whether bonds or stocks, not shown in the Thompson database that are held in common. If the only securities held in common are those listed in the Thompson database, then the first measure is accurate. If the portion of securities held in common for those omitted from the Thompson database is the same as the portion of stocks held in common for those included in the Thompson database, then the second measure is an accurate measure of common holdings. Both of the measures can be calculated for funds inside the family and funds outside the family.

The results are shown in Table 5. The table shows the common holdings for all stock and combination funds combined and for each of the subcategories.[9] Examining Panel A of Table 5 reveals (even under our conservative measure of common holdings) a surprisingly high level of common holdings and a larger increase in common holdings when one compares within-family funds with outside funds. Starting with the aggregate comparison, we see that within families the grouping stock-stock has 13.3% of the portfolio in common, for stock-combination groupings the overlap is 14.9%, and for combination-combination it is 27.4%. Furthermore, all of these percentages are more than twice as large as the percentages of common holdings in the same category when a fund inside the family is compared to a fund outside the family, and all of the differences are statistically significant at the 1% level.

When we examine the more narrowly defined objectives we see similar overall results. In all cases, funds in the same family hold more stocks in common than funds outside the family. Nine of the 14 combinations are statistically significant.. Panel B documents the same pattern of results. The logical question to ask is how much of the

[9] We lose some observations due to an insufficient match of TNA/fund name between CRSP and Thomson. The table omits the Balance-Balance and Total Return-Total Return groupings due to insufficient number of observations.

increased holdings correlation between funds in the same family is due to this phenomenon.

B. Impact of Common Holdings on Correlation

If we assume for the moment that the variance of all stocks can be reasonably represented by a single number *CVAR* and that the covariance between pairs of stocks can be represented by a single number *CCOV*, then we can express the covariance between two funds which hold only stocks but hold some securities in common as:

$$\mathrm{COV}\left(R_A, R_B\right) = \sum_{i \in S} X_{Ai} X_{Bi} CVAR + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} X_{Ai} X_{Bi} CCOV \tag{7}$$

Where *S* is the set of all stocks held in common. Note that the first term shows the impact on covariance when two stocks are held in common, and the second term when they are different stocks.[10] For all stocks not held in common either X_{Ai} or X_{Bi} must be equal to zero. Thus we can write the equation as:

$$\frac{\mathrm{COV}\left(R_A, R_B\right)}{\sigma_A \sigma_B} = \sum_{i \in S} \frac{X_{Ai} X_{Bi}}{\sigma_A \sigma_B} \left[CVAR - CCOV\right] + \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{X_{Ai} X_{Bi}}{\sigma_A \sigma_B} CCOV \tag{8}$$

Note that this equation holds assuming that the correlation between fund *A* and fund *B* is only due to correlation between known common holdings. Since we only have information on stock holdings, we underestimate the effect on common holdings for funds which have bonds in their portfolio or where small holdings of stocks are left out. However, since pairs of bonds should be highly correlated, the difference between variance and covariance is small and examining the first term of equation (8) shows the

[10] The covariance between two stocks is $\rho_{ij} \sigma_i \sigma_j$. If they are held in common, $\rho_{ij} = 1$ and $\sigma_i = \sigma_j$ and the covariance becomes the variance.

impact on portfolio correlation should be small. Thus, we will only present data assuming we have all the common holdings.

The first term on the right-hand side of Equation (8) represents the contribution of common holdings to the correlation between fund *A* and fund *B*. We compute the average for this statistic when fund *A* and fund *B* are in the same family, and a second average for the case where they are in different families. We then take the difference in this ratio and divide it first by the total correlation difference and second by the residual correlation from the one-index model (similar to the number in Table 4).[11] The first represents the fraction of the difference in correlation between funds in the same family and funds not in the same family that is due to the difference in common holdings.

The results are reported in Table 6. Common stock holdings account for 44% of stock-stock, 43% of stock-combination, and 89% of combination-combination correlation differences when we use our conservative estimate of common holdings. The estimates using the narrowly defined objective groupings show a similar pattern. The smallest ratio of correlation difference related to common holdings is 27%. Overall, common holdings explain roughly 50% of the difference between the correlation in fund returns (as well as residual fund returns) for funds inside and outside the family.

V. Differences in Variance Across Fund Families

In addition to increased fund return correlation, limiting investments to one family may also result in a greater dispersion of risk across investors. If several funds within a family share a similar strategy, we might expect fund variances not to be randomly distributed across families. Thus, if investors are restricted to one family, then similar strategies within families would cause a greater dispersion of investor risk than if high- and low-variance funds were randomly dispersed across fund families.

[11] The correlation numbers are not identical to those in Table 4 due to the omission of funds without a sufficient match between CRSP and Thomson.

We measure the propensity of families to choose high risk or low risk strategies by constructing a binomial test.[12] We begin by calculating the median standard deviation for each of the ICDI categories. We then label all funds with standard deviations above (below) the median for that objective as HIGH (LOW). We then examine whether the distribution of HIGHs and LOWs within fund families is different from that expected by chance.

Define $Y_{g,h}$ as the number of HIGHs obtained for fund family *h* when there are *g* funds in family *h*, and let $T_{g.h} = \left(Y_{g,h} - 1/2g\right)^2$. If the assignment of high and low variance funds to a fund family is random, then we would expect on average for a fund family to have $g/2$ HIGHs. Thus, $T_{g,h}$ becomes large if high risk funds are concentration in some families while low risk funds are concentrated in others. If the distribution of HIGHs and LOWs within a family is random, then under the null then $Y_{g,h}$ are independent binomial (g, $\frac{1}{2}$) random variables and the test statistic $T_{g,h}$ has the following moments:

$$
\begin{aligned}
E\left[T_{g,h}\right] &= E\left[\left(Y_{g,h} - 1/2g\right)^2\right] = \operatorname{var}\left(Y_{g,h}\right) = \frac{g}{4} \\
\operatorname{var}\left[T_{g,h}\right] &= \operatorname{var}\left(Y_{g,h} - 1/2g\right)^2 \\
&= \left\{E\left[\left(Y_{g,h} - \tfrac{1}{2}g\right)^4\right] - \left[E\left(Y_{g,h} - \tfrac{1}{2}g\right)^2\right]^2\right\} \qquad (9) \\
&= \left\{\left(3g^2\left(\tfrac{1}{2}\times\tfrac{1}{2}\right)^2 + g\left(\tfrac{1}{2}\times\tfrac{1}{2}\right)\left(1 - 6\left(\tfrac{1}{2}\times\tfrac{1}{2}\right)\right)\right) - \left[g\left(\tfrac{1}{2}\times\tfrac{1}{2}\right)\right]^2\right\} \\
&= \frac{g(g-1)}{8}
\end{aligned}
$$

Letting *H* be the total number of fund families and g_h be the number of funds in family *h*, we can test risk clustering using the following normally distributed test statistic:

[12] We thank Gary Simon for suggesting this approach.

$$
\begin{aligned}
Z &= \left(\sum_{h=2}^{H} T_{g,h} - \sum_{h=2}^{H} E\left[T_{g,h} \right] \right) \Bigg/ \sqrt{\sum_{h=1}^{H} \operatorname{var}\left(T_{g,h} \right)} \\
&= \left(\sum_{h=1}^{H} \left(Y_{g,h} - g_h/2 \right)^2 - \sum_{h=1}^{H} \left(g_h/4 \right) \right) \Bigg/ \sqrt{\sum_{h=1}^{H} \frac{g_h \left(g_h - 1 \right)}{8}}
\end{aligned}
\tag{10}
$$

Performing the risk concentration test for stock, combination, high-yield, and bond categories separately results in t-statistics equal to 3.02, 3.03, 0.63, and 2.8. The test across all fund objectives is 4.09. All are significant at the 1% level, with the exception of high-yield bonds. Another way of conceptualizing the extent of concentration of variance is to note that for 22 out of the 100 fund families in our sample every fund in the family was in the high or low variance group. High and low variances are more concentrated in families than would be expected by chance.

VI. Conclusion

Individual investors often restrict the mutual funds they select to the offerings of a single fund family. In addition, a common structure of 401K and 403b plans is that all the offerings are from one family. In this paper we show that this restriction causes investors to have higher risk portfolios than if they selected similar funds across different fund families. The principal reason for this higher risk is that funds within a family have higher correlation than if funds were selected from two families. This higher correlation holds for all ICDI categories involving stock and combination funds both when two funds are in the same ICDI category or when they are in two different ICDI categories.

Why does this increase in fund return correlation come about? When we split the increase in correlation between common response to market movements and increase due to residual correlation, we find that for most combinations more than 90% is due to residual correlation. Examining the effects of common holdings on the increase in correlation, we find about 60% of the increase in correlation is due to common holdings. Thus about 30% is due to a common response to factors other than the market, such as

industry and sector factors. The surprising result of this analysis is the size of common holdings. Depending on the group examined, between 4% and 34% are held in common with a median within family holding of roughly 16%.

There is another source of risk in addition to the increased fund return correlation. High and low risk funds are concentrated in different families. While this doesn't increase an investor's average risk, it does increase the distribution of risk across investors. The increased distribution raises the probability of having a bad outcome by investing in only one family. Overall, the results suggest that investors would be wise to build portfolios of funds from different families, and that retirement plan administrators would do well to include offerings for more than one mutual fund family.

References

Benartzi, Shlomo and Richard Thaler, 2001, Naïve diversification strategies in defined contribution saving plans, *American Economic Review* 91, 78-98.

Brown, Steven and William Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics* 43, 373-399.

Das, Sanjiv R., and Ranghuram K. Sundaram, 2002, Fee speech: Signalling, risk-sharing, and the impact of fee structures on investor welfare, *Review of Financial Studies* 15, 1465-1497.

Elton, Edwin J. and Martin J. Gruber, 1970, Homogeneous groups and the testing of economic hypotheses, *Journal of Financial and Quantitative Analysis 5, 581-602.*

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake, 2002, A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* 56(6).

Elton, Edwin J., Martin J. Gruber and Christopher R. Blake, 2003, The adequacy of investment choices offered by 401K plans, Unpublished manuscript, New York University.

Fama, Eugene and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Investment Company Institute and Securities Industry Association, 2002, *Equity Ownership in America*, *2002*.

Ivkovich, Zoran, 2001, Is blood thicker than water: Spillovers in mutual fund families, Unpublished manuscript, University of Illinois.

Khorana, Ajay, and Henri Servaes, 2003, An examination of competition in the mutual fund industry, Unpublished manuscript, Georgia Institute of Technology.

Liang, Nellie and Scott Weisbenner, 2002, Investor behavior and the purchase of company stock in 401K plan design. Unpublished manuscript, University of Illinois

Mayaysky, Harry, and Matthew Spiegel, 2001, A theory of mutual funds: optimal fund objectives and industry organization, Unpublished manuscript, Yale University.

Massa, Massimo, 2003, How do family strategies affect fund performance? When performance-maximization is not the only game in town, *Journal of Financial Economics*, 67, 249-304.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004, Family values and the star phenomenon, *Review of Financial Studies*, forthcoming.

Table 1
Mutual fund family summary statistics

Objectives	Number Of Funds	Number Of Families	Median Per Family	Maximum	Average Return	Average Risk
Stock	384	94	3	29	0.21	7.30
Aggressive Growth	166	75	2	10	0.26	8.08
Long-term Growth	218	82	2	19	0.16	6.70
Combination	275	83	2	18	0.23	4.49
Total Return	50	30	1	5	0.40	3.52
Growth and Income	128	65	1	9	0.19	5.48
Balanced	65	47	1	4	0.22	3.33
Income	65	47	1	4	0.17	4.38
High Yield Bond	39	31	1	3	-0.08	2.67
Bond	290	74	2	25	0.46	0.94
Ginnie Mae Bond	40	21	2	5	0.48	0.67
High Quality Bond	105	56	1	9	0.47	0.95
Municipal Bond	79	43	1	5	0.40	1.09
Government Securities	66	43	1	10	0.49	0.93
All Objectives	988	100	6	85	0.28	4.47

The table shows characteristics of the fund families considered in the study. The Stock, Combination, and Bond objectives are decomposed into subcategories. The Number Of Families refers to the number of families with at least one fund of that objective category. Median Per Family refers to the median number of funds for the subset of families that offer a fund of that objective. Maximum refers to the largest number of funds of that type for any family. Also reported is the average return and standard deviation for each objective classification. The sample period covers 1998 through 2002, and the return and objective data are taken from the CRSP mutual fund database.

Table 2
Return correlations by objective within and outside fund families

Objectives	Obs	Within Family Corr.	Outside Family Corr.	t-stat	Percent Larger	Binomial p-value
Stock – Stock	77	0.774	0.734	3.70	0.714	0.000
Aggressive Growth – Aggressive Growth	41	0.780	0.738	2.49	0.634	0.030
Aggressive Growth – Long-term Growth	62	0.757	0.718	3.66	0.661	0.004
Long-term Growth – Long-term Growth	42	0.805	0.774	2.33	0.690	0.004
Stock – Combination	78	0.757	0.709	4.71	0.667	0.001
Aggressive Growth – Total Return	28	0.710	0.706	0.26	0.536	0.286
Aggressive Growth – Growth and Income	48	0.695	0.660	2.50	0.667	0.007
Aggressive Growth – Balanced	39	0.719	0.681	2.57	0.718	0.002
Aggressive Growth – Income	22	0.644	0.631	0.68	0.636	0.067
Long-term Growth – Total Return	28	0.772	0.735	2.19	0.821	0.000
Long-term Growth – Growth and Income	52	0.793	0.757	2.26	0.673	0.004
Long-term Growth – Balanced	44	0.844	0.769	4.92	0.864	0.000
Long-term Growth – Income	26	0.740	0.716	1.17	0.577	0.163
Stock – High Yield Bond	31	0.498	0.495	0.34	0.645	0.035
Stock – Bond	70	-0.146	-0.142	-0.43	0.400	0.940
Combination – Combination	55	0.835	0.766	5.32	0.782	0.000
Total Return – Total Return	10	0.839	0.731	2.73	0.800	0.011
Total Return – Growth and Income	23	0.766	0.734	1.25	0.696	0.017
Total Return – Balanced	16	0.777	0.760	0.69	0.750	0.011
Total Return – Income	16	0.769	0.719	2.16	0.625	0.105
Growth and Income – Growth and Income	30	0.857	0.805	3.11	0.767	0.001
Growth and Income – Balanced	34	0.868	0.799	4.05	0.824	0.000
Growth and Income – Income	20	0.845	0.809	2.63	0.800	0.001
Balanced – Balanced	12	0.920	0.832	3.52	1.000	0.000
Balanced – Income	18	0.824	0.803	1.10	0.611	0.119
Income – Income	4	0.859	0.836	0.40	0.500	0.313
Combination – High Yield	31	0.476	0.472	0.48	0.484	0.500
Combination – Bond	67	-0.147	-0.132	-1.69	0.403	0.929
High-Yield Bond – High-Yield Bond	6	0.890	0.858	1.92	0.833	0.016
High Yield Bond – Bond	30	0.009	0.007	0.15	0.567	0.181
Bond – Bond	50	0.688	0.686	0.21	0.560	0.161
Ginnie Mae Bond – Ginnie Mae Bond	11	0.808	0.723	1.76	0.818	0.006
High Quality Bond – High Quality Bond	23	0.663	0.613	1.44	0.696	0.017
Municipal Bond – Municipal Bond	20	0.929	0.913	1.91	0.850	0.000
Government Securities – Government Securities	14	0.856	0.851	0.19	0.714	0.029

The table reports average return correlations of funds within and outside fund families. Correlations are averaged first within families and then across families. The number of observations is the number of families with at least one pair of funds matching the objectives being considered. "Stock" refers to funds with Aggressive Growth or Long-Term Growth objectives; "Combination" refers to Total Return, Growth and Income, Balanced, or Income; "Bond" refers to Ginnie Mae funds, High Quality Bond, High Quality Municipal Bond, or Government Securities. The sample period covers 1998 through 2002, and the return and objective data are taken from the CRSP mutual fund database.

Table 3
Influence of families on portfolio diversification

		Number of Funds Added	
Current Objective	Fund added	1	2
Stock	Stock	2	4
Stock	Combination	3	4
Combination	Combination	3	17
Combination	Stock	2	3

An investor is assumed to start with one mutual fund and adds one or two funds from outside the family. The table shows the number of funds from within the family that are necessary to arrive at the same level of risk as adding funds from outside the family. "Stock" refers to funds with Aggressive Growth or Long-Term Growth objectives; "Combination" refers to Total Return, Growth and Income, Balanced, or Income. Fund variances are assumed to be the average for that objective classification as reported in Table 1, and return correlations for each objective inside and outside the family are taken from Table 2. The sample period covers 1998 through 2002, and the return and objective data are taken from the CRSP mutual fund database.

Table 4
Determinants of the difference in fund return correlations inside and outside fund families

Panel A: Two-factor model	Obs	Return Correlation Difference	Systematic Comp.	Idiosyncratic Comp.	Ratio
Stock – Stock	77	0.040	-0.004	0.044	1.10
Aggressive Growth – Aggressive Growth	41	0.042	0.001	0.041	0.98
Aggressive Growth – Long-term Growth	62	0.039	0.005	0.034	0.87
Long-term Growth – Long-term Growth	42	0.032	0.000	0.032	1.00
Stock – Combination	78	0.048	0.000	0.048	1.00
Aggressive Growth – Growth and Income	48	0.035	0.012	0.023	0.66
Aggressive Growth – Balanced	39	0.038	-0.001	0.039	1.03
Long-term Growth – Total Return	28	0.037	0.007	0.029	0.78
Long-term Growth – Growth and Income	52	0.036	0.009	0.027	0.75
Long-term Growth – Balanced	44	0.075	0.000	0.075	1.00
Combination – Combination	55	0.068	-0.004	0.072	1.06
Total Return – Total Return	10	0.108	0.006	0.102	0.94
Total Return – Income	16	0.050	0.000	0.050	1.00
Growth and Income – Growth and Income	30	0.052	0.010	0.042	0.81
Growth and Income – Balanced	34	0.069	0.003	0.066	0.96
Growth and Income – Income	20	0.036	0.008	0.028	0.78
Balanced – Balanced	12	0.089	0.046	0.043	0.48
Panel B: Six-factor model					
Stock – Stock	73	0.036	0.008	0.028	0.78
Aggressive Growth – Aggressive Growth	36	0.037	0.007	0.030	0.81
Aggressive Growth – Long-term Growth	61	0.038	0.018	0.020	0.53
Long-term Growth – Long-term Growth	40	0.036	0.016	0.020	0.56
Stock – Combination	75	0.058	0.030	0.028	0.48
Aggressive Growth – Growth and Income	44	0.033	0.022	0.012	0.36
Aggressive Growth – Balanced	36	0.034	0.015	0.019	0.56
Long-term Growth – Total Return	26	0.043	0.023	0.020	0.47
Long-term Growth – Growth and Income	50	0.036	0.020	0.016	0.44
Long-term Growth – Balanced	42	0.079	0.043	0.036	0.46
Combination – Combination	52	0.075	0.031	0.044	0.59
Total Return – Total Return	8	0.127	0.053	0.075	0.59
Total Return – Income	12	0.040	0.017	0.023	0.58
Growth and Income – Growth and Income	28	0.050	0.025	0.025	0.50
Growth and Income – Balanced	33	0.078	0.044	0.034	0.44
Growth and Income – Income	17	0.047	0.024	0.023	0.49
Balanced – Balanced	10	0.103	0.070	0.034	0.33

The idiosyncratic component of return correlation is measured by the average covariance of fund return residuals, scaled by the standard deviation of each fund's returns; the systematic correlation is the correlation related to common exposure to return factors. Residual returns are obtained by regressing excess fund returns on the excess return of several index factors. The last column shows the ratio of the idiosyncratic component over the return correlation difference. Panel A shows the results for a two factor model, which includes the Value weighted CRSP Index (from Ken French) and the excess return on the Merrill Lynch aggregate U.S. Corp/Gov/Mortgage bond index. The six-factor model in Panel B adds equity size and value factors (SMB And HML), as well as mortgage and high yield indexes. The number of observations is the number of families with at least one pair of funds that matches the objectives being considered. Stock refers to Aggressive Growth and Long-Term Growth; Combination refers to Total Return, Growth and Income, Balanced, and Income; Bond refers to Ginnie Mae funds, High Quality Bond, High Quality Municipal Bond, and Government Securities. The sample period covers 1998 through 2002, and the return and objective data are taken from the CRSP mutual fund database.

Table 5
Common stock holdings for funds within and outside fund families

Objectives	Obs.	Common Holdings Within Family	Common Holdings Outside Family	t-stat
Panel A: Percentage of total net assets				
Stock – Stock	47	0.133	0.056	3.51
Aggressive Growth – Aggressive Growth	25	0.139	0.029	2.72
Aggressive Growth – Long-term Growth	41	0.095	0.038	2.93
Long-term Growth – Long-term Growth	25	0.169	0.104	2.13
Stock – Combination	49	0.149	0.071	4.28
Aggressive Growth – Growth and Income	25	0.040	0.029	1.46
Aggressive Growth – Balanced	18	0.059	0.024	1.76
Long-term Growth – Total Return	15	0.157	0.079	1.79
Long-term Growth – Growth and Income	29	0.222	0.129	3.66
Long-term Growth – Balanced	18	0.214	0.084	3.77
Combination – Combination	30	0.274	0.128	4.77
Total Return – Income	5	0.176	0.081	1.96
Growth and Income – Growth and Income	14	0.236	0.174	2.69
Growth and Income – Balanced	14	0.340	0.119	4.55
Growth and Income – Income	6	0.272	0.150	2.01
Panel A: Percentage of stock holdings				
Stock – Stock	47	0.144	0.060	3.50
Aggressive Growth – Aggressive Growth	25	0.145	0.030	2.79
Aggressive Growth – Long-term Growth	41	0.101	0.040	2.93
Long-term Growth – Long-term Growth	25	0.184	0.110	2.14
Stock – Combination	49	0.184	0.083	4.29
Aggressive Growth – Growth and Income	25	0.042	0.030	1.47
Aggressive Growth – Balanced	18	0.092	0.034	1.87
Long-term Growth – Total Return	15	0.253	0.121	1.95
Long-term Growth – Growth and Income	29	0.236	0.136	3.58
Long-term Growth – Balanced	18	0.326	0.119	3.56
Combination – Combination	30	0.404	0.163	4.50
Total Return – Income	5	0.251	0.121	1.74
Growth and Income – Growth and Income	14	0.250	0.184	2.71
Growth and Income – Balanced	14	0.546	0.167	4.31
Growth and Income – Income	6	0.328	0.169	2.02

The table reports the average percentage of holdings in common for funds within and outside fund families. For each fund pair, the common percentage holdings are calculated as $\Sigma_s \min(X_{si}, X_{sj})$ where X_{si} represents the percentage of fund j's holdings in stock s. Panel A calculates holdings as a percentage of total net assets, and Panel B reports holdings as a percentage of total stock holdings. The number of observations is the number of families with at least one pair of funds that matches the objectives being considered. Stock refers to Aggressive Growth and Long-Term Growth; Combination refers to Total Return, Growth and Income, Balanced, and Income. Fund holdings are taken from Thomson Financial and are measured in December of 2000.

Table 6
Percentage of the difference in return correlation inside and outside fund families due to common holdings

		Difference in Correlation		Portion due to Common Holdings	
Panel A: Percentage of stock holdings	Obs	Return	Residual	Return	Residual
Stock – Stock	47	0.039	0.036	0.44	0.48
Aggressive Growth – Aggressive Growth	25	0.041	0.042	0.27	0.27
Aggressive Growth – Long-term Growth	41	0.035	0.032	0.35	0.38
Long-term Growth – Long-term Growth	25	0.014	0.014	1.37	1.44
Stock – Combination	49	0.051	0.053	0.43	0.42
Aggressive Growth – Growth and Income	25	0.010	0.014	0.58	0.44
Aggressive Growth – Balanced	18	0.058	0.049	0.33	0.39
Long-term Growth – Total Return	15	0.040	0.033	0.62	0.75
Long-term Growth – Growth and Income	29	0.040	0.030	0.47	0.62
Long-term Growth – Balanced	18	0.092	0.100	0.67	0.61
Combination – Combination	30	0.074	0.066	0.89	0.99
Total Return – Income	5	0.056	0.040	2.11	2.96
Growth and Income – Growth and Income	14	0.018	0.015	1.15	1.38
Growth and Income – Balanced	14	0.126	0.135	0.94	0.88
Growth and Income – Income	6	0.038	0.036	0.66	0.69

The table shows the influence of common stock holdings on the difference between fund return correlations within and across fund families. The influence of common holdings is measured by increased correlation over a benchmark portfolio with different stocks. The number of observations is the number of families with at least one pair of funds that matches the objectives being considered. Stock refers to Aggressive Growth and Long-Term Growth; Combination refers to Total Return, Growth and Income, Balanced, and Income; Bond refers to Ginnie Mae funds, High Quality Bond, High Quality Municipal Bond, and Government Securities. The sample period covers 1998 through 2002, and the return and objective data are taken from the CRSP mutual fund database. Fund holdings are taken from Thomson Financial and are measured in December of 2000.