Does mutual fund style category composition explain the benefits of family diversification?

by

Andrew Teodorescu

Jeong Ho (John) Kim
Adviser

Department of Economics

Jeong Ho (John) Kim

Adviser

Bree Ettinger

Committee Member

Caroline Fohlin

Committee Member

2020

Does mutual fund style category composition explain the benefits of family diversification?

By

Andrew Teodorescu

Jeong Ho (John) Kim

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Economics

2020

Abstract

Does mutual fund style category composition explain the benefits of family diversification?

By Andrew Teodorescu

This study examines the influence of family and investment style category factors on the cross-sectional variance and correlation structure of actively managed U.S. mutual fund returns. I find that actively managed fund family returns are largely driven by category factors, and that cross-family correlation rises when accounting for the category composition of the families. In particular, category structure explains roughly half of the cross-sectional differences in family return volatility. Diversification across categories within the same family seems to be a better bet than diversification across families within the same category. My findings suggest that category diversification explains a considerable portion of the rise in investor risk that previous studies have identified when investments are constrained to funds within just one family.

Does mutual fund style category composition explain the benefits of family diversification?

By

Andrew Teodorescu

Jeong Ho (John) Kim

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Arts with Honors

Department of Economics

2020

Acknowledgements

I am grateful for guidance from my thesis advisor, John Kim. I would also like to thank Stephen O'Connell and participants in Emory's 2019-2020 Economics Honors Program for their helpful comments at several seminars. All errors are my own.

Table of Contents

Figures

Tables

1. Introduction

The actively managed U.S. equity mutual fund industry has blossomed in size over the previous three decades, with over 4.3 trillion dollars under management as of March 2020. This industry's proliferation has been met with a surge in demand for the evaluation of fund performance on behalf of both individual and institutional investors. As a result, the mutual fund literature is today one of the largest in empirical finance. Academics and practitioners studying mutual fund performance and risk typically choose to examine funds at an aggregated level, with fund family being the most popular of these choices. While holding funds across fund families has been found to have positive diversification benefits, many investors are compelled to constrain their mutual fund investments within just one family. For example, Elton et al. (2006) study an array of 401k plans offered by employers and find that 45% of them restrict fund choices of participants to a single family. Load fees also deter investors from moving assets between families. Last, investors may exercise rational inattention by investing in only one family. Investigating the potential loss of diversification benefits by constraining funds to one family is an issue that has received little attention in the mutual fund literature.

Researchers have previously identified strong links between a fund's performance and its family membership. Brown and Wu (2016) examine cross-fund learning within fund families and find that flows to a member fund respond positively on average to family performance, suggesting that a fund's performance reflects not only its fund-specific characteristics but also the quality of the skill and resources shared across funds.[1] Others have found a similar impact of family performance on fund behavior and volatility in an array of settings (e.g. Kempf and Ruenzi, 2007; Massa, 2003; Nanda et al., 2004). Elton et al. (2007) find that, on average, portfolios of funds within families result in greater overall risk and greater risk clustering than similar portfolios created from funds across families. In particular, these

[1]Previous studies have found large cross-sectional differences in fund skill that are persistent (Berk and van Binsbergen, 2015) and grow over time (Pástor et al., 2015) in the actively managed mutual fund industry.

authors indicate that confining mutual fund investments to one family has a detrimental effect on investor risk.

Alongside fund family as a typical unit of analysis, the mutual fund literature has emphasized the importance of fund investment style in explaining differences in fund returns and volatility. Brown and Goetzmann (1997) propose an empirical method for categorizing fund investment styles and argue that style categories are partially responsible for cross-sectional differences in fund volatility as well as the correlation structure of family returns. Chan et al. (2015) find that differences in style are associated with differences in performance and compare several methods of style identification. Using Morningstar's style classification system for mutual funds, Teo and Woo (2004) find strong evidence that style-level value strategies based on annual style returns are profitable. Many papers have found a dominant role for style analysis driving the statistical moments of fund returns (e.g. Bogle, 1998; Brown et al., 2015; Kaniel and Parham, 2017).

If category composition is heterogeneous across families, family returns will contain category effects; similarly, category returns will be partially explained by family effects. My empirical estimation strategy for decomposing fund returns into category- and family-specific components runs parallel to Heston and Rouwenhorst (1994), who decompose international stock index returns into country and industry effects. Using monthly data for 3,768 funds that comprise the survivor-bias-free Center for Research in Security Prices (CRSP)-Morningstar universe of 317 families during 1991–2017, I present evidence that category composition and family-specific effects play a roughly equal role in explaining excess family returns. These category effects are not the result of family self-selection into funds by their risk profiles. My results are in contrast to studies that have found an overwhelming role for family effects in fund returns. In particular, these findings suggest within-family diversification as a means of risk reduction is not as dangerous as previously thought. I identify large category effects in family returns, but what remains is to determine the origination of these effects and those that drive family-specific idiosyncrasies.

The rest of the paper is organized as follows. Section 2 describes the CRSP-Morningstar data. Section 3 outlines the methodology, and my results appear in Section 4. Section 5 concludes.

2. Data

2.1. Sample construction

My data comes from a survivor-bias-free sample of mutual funds from the CRSP and Morningstar universes. I require that funds appear in both databases for the purpose of data validation. My process of merging the two databases and selecting only actively managed equity-only mutual funds in the United States follows exactly that of Berk and van Binsbergen (2015). Interested readers may refer to the data appendices of that paper for more detail. My mutual fund data set contains 3,768 funds with a partial or complete return history between 1991 and 2017.

I use Morningstar Category to assign funds to the nine types corresponding to Morningstar's 3×3 Style Box (Small Growth, Mid-Cap Value, etc.) or otherwise.[2] These nine style categories together compose 49% of the fund-months in the final sample and average 52% of the sample's total assets in each month.

The final sample includes 324 months of returns for 317 actively- anaged equity-only mutual fund families and 69 mutual fund Morningstar style categories from 1991 to 2017. In order to enter the sample each fund must, in a given sample month, belong to a family that is composed of at least two other categories and belong to a category that is composed of at least two other families. These two conditions are imposed because my decomposition method identifies family (category) effects within a category (family), which may either be absent or driven entirely by only one other family (category) without these conditions.

[2] Morningstar reviews category assignments semi-annually to reflect the primary investment focus of a fund over the past three years. Category assignments are intentionally quite stale, and category changes occur only when a fund has exhibited a strong, sustained shift into a new investment style. See Morningstar, Inc. (2018) for detail on Morningstar's category assignment methodology.

These conditions restrict the number of fund-months in the sample by 27%, eliminating 69% of families and five of 74 categories.[3] Nevertheless, with a monthly average of 80% of total assets of the full sample, the restricted sample is fairly representative of the U.S. mutual fund landscape. Because fund value-weights within families (categories) are distorted by the removal of categories (families) that are not family-wise (category-wise) diverse, I present equally-weighted statistics alongside value-weighted statistics.

2.2. Category composition and return correlations

Table **1** reports the average monthly category composition and market capitalization value-weights of the top 10 mutual funds with a complete panel in my sample (ranked by monthly average inflation-adjusted assets under management). The category composition of families, as well as family distribution within categories, is not uniform. The top panel shows that several of the top families do not have funds in each of the Morningstar Style Box at any point in the sample period. The bottom panel gives the average weight measured as the percentage of the total capitalization of the funds in our sample, which I will refer to as the "U.S. market." Similar to Panel A, this panel shows that the value weights of categories within families varies widely.

Table **2** summarizes the performance of the top 10 families and the nine Morningstar Style Box categories over the sample period. The top panel shows that there are notable discrepancies across families in terms of average returns and the standard deviation of these returns. Blackrock and Morgan Stanley were among the poor-performing families, while State Street Global Advisors and Wells Fargo Funds were the highest performers. Measured by the standard deviation of returns, the value-weighted returns of Fidelity Investments were almost 20% more volatile than that of John Hancock. Although the time-series mean of the number of funds in each category (55.5) is three times that of the number of funds in

[3] The number of categories in my final sample grows from 18 in 1991 to 69 in 2017. The proliferation of category complexity, however, does not appear to bias Style Box category weights downwards in the latter sample years. Only 71 of 3,768 funds are reassigned from a Style Box category to a non-Style Box category in the final sample.

each family (16.5), the bottom panel shows that category performance was about as uniform as family performance. This is likely due to the portfolio holdings of families being more diversified than those of categories. Furthermore, the top panel also shows that most of the cross-family correlations are high, despite the fact that families often have well-diversified portfolios.

The value-weighted average correlation between families is 0.81. Equally-weighted correlations are on average slightly higher at 0.85. At the same time, the average of within-family correlations is not much lower than that of between-family correlations, at 0.80. That is, the average correlation between two funds within the same family is approximately the same as that between two funds from separate families. The value-weighted average of the category correlations is 0.71, and the equally-weighted average for is 0.73. Opposite to family correlations, the average within-category correlation is of roughly equal magnitude to between-category correlations, at 0.75.

3. Model and empirical design

3.1. Modeling decomposition into fixed effects

The crucial part of my analysis is the decomposition of mutual fund returns into family and category components. In a seminal paper in the international diversification literature, Heston and Rouwenhorst (1994) propose a reduced form model that decomposes excess country stock index returns into only country and firm industry effects.[4] I adapt this model to the mutual fund space to express return for fund i belonging to category j and family k in month t by

$$R_{it} = \alpha_t + \beta_{jt} + \gamma_{kt} + \epsilon_{it}, \tag{1}$$

where α_t is the average return of mutual funds in period t, β_{jt} is the category effect, γ_{kt} is the family effect, and ϵ_{it} is a fund-specific serially uncorrelated error term. Equation (1) does

[4] I refer readers interested in methodologically-related studies to Koren and Tenreyro (2007), Griffin and Karolyi (1998), and Vassalou and Xing (2004).

not allow for the possibility for interactions between the category-family effects. I assume that fund-specific disturbances have a zero mean and finite variance for returns in all families and categories, and are uncorrelated across funds.[5]

3.2. Choosing a baseline return

My data on 3,768 actively-managed funds spans 317 unique families and 69 categories in the time series. Let δ_{ij} be a category dummy that is 1 if fund i belongs to category j and 0 otherwise. Analogously, let θ_{ik} be a family dummy. In each month t, (1) can be expressed by

$$R_i = \alpha + \sum_j \beta_i \delta_{ij} + \sum_k \gamma_i \theta_{ik} + \epsilon_i. \tag{2}$$

Since the family dummies as well as the category dummies add up to the unit vector across firms, perfect multicollinearity between the δ_{ij} and the θ_{ik} across fund i limits us from estimating (2) cross-sectionally via ordinary least squares (OLS) regression. Every fund belongs to one family and one category, thus I can only measure cross-sectional differences between categories and cross-sectional differences between families. That is, family and category effects must be measured along a benchmark, say by selecting the "Mid-Cap Blend" category in the Fidelity Investments family as my baseline. I would then estimate (2) with the restriction that $\beta_{\text{Mid-Cap Blend}} = \gamma_{\text{Fidelity Investments}} = 0$. Then, β_j would measure the category effect of category j relative to the "Mid-Cap Blend" category, and γ_k would measure the family effect of family k relative to Fidelity Investments. However, this arbitrary choice of category and family is not very informative; instead, it makes more sense to ask how each category or family differs from the *average* fund in my sample. Equivalently, I am interested in measuring category and family effects relative to a fund-of-funds (FoF) representative of the U.S. equally-weighted mutual fund market. To avoid the interpretation problem of an arbitrary benchmark, we can impose the constraint that, for equally-weighted fund returns,

[5]The return decomposition (1) posits that the different components of a fund's return are orthogonal to one another. Hence it permits a simple variance decomposition in which all covariance terms are zero.

the sum of the family coefficients equals zero and the sum of the category coefficients equals zero (Suits, 1984; Kennedy, 1986). Making use of this definition requires the following restrictions in each month t:

$$\sum_j n_j \beta_j = 0, \tag{3a}$$

$$\sum_k m_k \gamma_k = 0, \tag{3b}$$

where n_j and m_k denote the number of funds in category j and family k, respectively.

Since estimated disturbances are orthogonal to all family and category dummies by construction, the average residual is zero in every family and in every category. The U.S. market FoF is simply the equally-weighted average over all families and categories, so the average disturbance for the U.S. FoF is also zero; the intercept in (2) can thus signify the equally-weighted market.

3.3. Components of family returns

The 'pure' category return, $\hat{\alpha} + \hat{\beta}_j$, is the OLS estimate of the return on a family-wise diversified FoF of funds in the j-th category. Here, a family-wise diversified FoF is one that has the same family composition as the U.S. equally-weighted mutual fund market, and is therefore free of family effects. Similarly, $\hat{\alpha} + \hat{\gamma}_k$ is an estimate of the 'pure' return on the family FoF k, which is category-wise diversified by having the same category composition as the U.S. market FoF, and therefore has no partial category effect.

Define R_k^{ew} as the equally-weighted "typical" fund of family k. Then, estimating (2) subject to restrictions (3a) and (3b) is convenient because it allows us to decompose R_k^{ew} into a component that is common to all families, $\hat{\alpha}$, the average of the category effects of the funds that it is composed of, and a family-specific component, $\hat{\gamma}_k$,

$$R_k^{ew} = \hat{\alpha} + \frac{1}{m_k} \sum_i \sum_j \hat{\beta}_j \delta_{ij} + \hat{\gamma}_k, \tag{4}$$

where the i-summation is taken over funds in family k.

3.4. Understanding variation in family returns

Equation (4) says that the return in Fidelity Investments may be different than that on the U.S. market FoF in period t for potentially two reasons:

1. The category composition of the Fidelity Investments "typical" fund is different from the category composition of the U.S. market. If, on average across the U.S., "Small-Cap Value" funds outperform and "Large Growth" funds underperform the U.S. market, the category effect for Fidelity Investments will be positive because Fidelity Investments has proportionally more "Small Value" funds and proportionally fewer "Large Growth" funds than the U.S. equally-weighted market.

2. The return on Fidelity Investments funds is different from funds which are in the same category but belonging to a different family. The family effect for Fidelity Investments is a measure of how well each category in Fidelity Investments performed relative to the average U.S. fund in that category (by total return for all fund assets belonging to the same category).

3.5. Components of category returns

Parallel to my decomposition in (4), each equally-weighted category FoF return R_j^{ew} can be decomposed into a component that is common to all categories, $\hat{\alpha}$, the weighted average of family components, and a category-specific component, $\hat{\beta}_j$:

$$R_j^{ew} = \hat{\alpha} + \hat{\beta}_j + \frac{1}{n_j} \sum_i \sum_k \hat{\gamma}_k \theta_{ik}, \tag{5}$$

where the i-summation is taken over funds in category j.

Since, by construction, the residuals for each family and category sum to zero in (4) and (5), these two equations do not include error terms.

The regressions above produce the category and family effects for month t. Collecting cross-sectional parameter estimates for each month, I obtain a time series of family-diversified category returns, $\hat{\alpha}_t + \hat{\beta}_{jt}$, and of category-diversified family returns, $\hat{\alpha}_t + \hat{\beta}_{kt}$. I suspect that variation in family and category "typical" fund returns can be partially explained by these returns.

3.6. Value-weighted decomposition

By using weighted least-squares to estimate (2), I can similarly deconstruct the value-weighted family and category "typical" funds. The weights are simply the total assets under management of each fund at the beginning of the month. The restrictions that imply that the value-weighted U.S. "typical" fund that has neither family effects nor category effects become, for each month t,

$$\sum_j w_j \beta_j = 0, \tag{6a}$$

$$\sum_k v_k \gamma_k = 0, \tag{6b}$$

where w_j and v_k are the value weights of category j and family k in the U.S. value-weighted market, and $\sum_k v_k = \sum_j w_j = 1$. After imposing these restrictions, the weighted least-squares estimate of the regression intercept now becomes the U.S. market value-weighted return. Similar to my earlier result, category and family returns are diversified portfolios in the sense that they have the same (value-weighted) family and category distribution as the U.S. market value-weighted return.

4. Decomposition results

4.1. Role of family and category effects in return volatility

I show in Equation (4) that a mutual fund family's return in excess of the U.S. mutual fund market can be decomposed into a pure family effect and a weighted average of cross-sectional category effects. Similarly, an excess category return equals a weighted average of cross-sectional family effects plus a pure category effect. In Table **3** I present the results of this decomposition. The ratio relative to the market gives the ratio of the variance of that component to the variance of that component's return in excess of the market. The top panel shows that the variance of excess equally-weighted family returns appears to be split roughly evenly between family and category effects. The bottom panel gives more insight into the decomposition: most of the variance of excess equally-weighted category returns can be attributed to category-specific effects, since the variance of the sum of family effects is on average only 3.6% of the variance of excess family returns. Not only is most of the variation in excess category returns is due to category effects, the average variance of the pure category effects is 8.27% squared, which is much larger than the average variance of the pure category effects (1.04% squared), due in part to the diversification of families across categories. Category effects in families are generally much larger than family effects in categories.[6] The right panel shows a similar pattern for the value-weighted FoFs.

The most important conclusion from Table **3** is that category composition of mutual fund families explains a considerable portion of the variance in family returns. Table **A.1** shows that the relative sizes of family and category effects are fairly consistent across time. Similarly, Tables **A.2** and **A.3** report consistency in the relationships between these effects when accounting for the number of funds in and the size of a family (category), respectively.

[6]Note that the family and category effects in FoFs are not uncorrelated. As a result the variance ratios of family effects and category effects do not add up to one, due to the relatively small covariances between them. In particular, the covariance between value-weighted pure family effects and the sum of category effects is -0.0126, and the covariance between value-weighted pure category effects and the sum of family effects is -0.0290.

4.2. Absolute family and category effects and model fit

Because families and categories differ in size, Panel A of Figure 2 reports the 36-month moving average absolute values of the family and category effects over time. On each date, the absolute values of the family and category effects were weighted by their respective market capitalizations.[7] The average category effect for the full sample is 1.40% per month (in absolute value), whereas the average absolute family effect is 0.62% per month. That is, a family-neutral tilt relative to the U.S. market has given rise to a tracking error that has been, on average, about 2.4 times as large as a category-neutral tilt of similar size. Panel B shows that, despite portraying a relatively restricted view of sources of mutual fund returns, the estimated regressions explain returns quite well. Category and family effects alone explain on average 62% of the variation in value-weighted returns and on average 69% of the variation in equally-weighted returns. Model fit (as measured by R^2) is relatively consistent over time and directly tracks market return. Comparing Panels A and B reveals that average absolute category and family effects inversely track the market, or that category and family effects are more dispersed during market downturns.

4.3. Decomposition corrections in family and category correlations

Table 4 reports family correlations corrected for category composition, similar to the raw correlations in Table 2. As shown in Table 3, around half of the variation of family returns in excess of the U.S. mutual fund industry can be attributed to category effects, which consequently has an economically-significant effect on family correlations. The average correlation between equally-weighted family correlation rises from 0.85 to 0.91 when corrected for category effects. The value-weighted correlations also rise on average from 0.81 to 0.88. This growth of category-corrected family correlations towards unity further confirms that category composition is important in explaining between-family differences in return

[7] The absolute family and category effects in month t were computed by $\sum_j w_j|\beta_j|$ and $\sum_k v_k|\gamma_k|$, respectively. Note that, without taking absolute values, these sums would be zero by construction.

volatility. Because family effects constitute a mcuh smaller proportion of the variance of category returns than category effects do of family returns, the family correction only has little effect on the category correlations: the average equally-weighted category correlation remains at 0.73, and the average value-weighted correlation falls slightly from 0.71 to 0.70. The absence of change in category correlations after family corrections instills confidence that category effects are not explained by family selection into categories.

4.4. Implications for portfolio diversification

The relative size of family and category effects has important implications for FoF diversification. Figure **1** shows that randomly combining U.S. mutual funds in large FoFs reduces return variance to 69% of that of the average fund. This investment strategy diversifies both over families as well as over categories. Diversification across categories within a single category only reduces FoF return variance to 88% of the average fund variance. Diversification across categories within a single family reduces the FoF variance to 76%.[8] Correcting family returns for category effects results in a further seven percent decrease in return variance, corroborating my previous result that category effects explain a considerable portion of family returns. In line with the observation that family composition is insignificant to category returns, correcting for family effects in category returns only further reduces return variance by one percent. Diversification benefits for actively managed FoFs deteriorate somewhat quickly; across each of these diversification strategies, adding another fund to a FoF with 10 or more funds has a negligible effect on return variance. It is impossible in practice to

[8] An equally-weighted FoF of N such funds has a variance equal to

$$\overline{\text{Var}} \cdot \frac{1}{N} + \overline{\text{Cov}} \cdot \frac{N-1}{N},$$

where $\overline{\text{Var}}$ is the variance of the average monthly fund return and $\overline{\text{Cov}}$ is the average covariance among these funds. The average fund return has a variance of 0.0235 per month, and the average covariance in a large group of funds is just equal to the variance of an equally-weighted FoF. When diversifying across all U.S. mutual funds, the average covariance is $(0.0403)^2$, the variance of the equally-weighted U.S. FoF (see Table **2**). This is only 69% of the average variance of an individual fund. The weighted average variance of equally-weighted FoFs across families is 0.2110, and the weighted average variance of equally-weighted FoFs across categories is 0.1790. These numbers are 88% and 76% of the average fund variance, respectively.

perfectly diversify across categories within any one family, or to perfectly diversify across families within any one category. However, these results offer a heuristic accessible to all investors: risk reduction within a family relative to the U.S. market benchmark is improved as category composition tends towards that of the U.S. market.

5. Conclusions

This paper investigates the family and investment style category structure of U.S. mutual fund returns using data on 3,768 funds in 317 mutual fund families and 69 style categories. By separately measuring family and category effects, I am able to examine why family returns differ in volatility. Although I estimate a very reduced form model that identifies fund returns only in terms of a fund's family and category memberships, these effects together explain 62% of the variance in value-weighted fund returns. Category specialization explains about half of the variance in family returns. Controlling for category structure drives the correlations among U.S. family returns closer to unity and markedly reduces portfolio variance. Together, these findings suggest that constraining funds to one family is not as harmful as previous studies have found. These findings do not identify the origin of these strong independent family movements; I leave their consideration for future research.

References

Berk, Jonathan B., and Jules H. van Binsbergen, 2015, Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118, 1–20.

Bogle, John C., 1998, The implications of style analysis for mutual fund performance evaluation, *Journal of Portfolio Management* 24, 34.

Brown, David P., and Youchang Wu, 2016, Mutual fund flows and cross-fund learning within families, *The Journal of Finance* 71, 383–424.

Brown, Keith C., W. Harlow, and Hanjiang Zhang, 2015, Investment style volatility and mutual fund performance, *Working Paper, University of Texas McCombs School of Business* .

Brown, Stephen J., and William N. Goetzmann, 1997, Mutual fund styles, *Journal of Financial Economics* 43, 373–399.

Chan, Louis K. C., Hsiu-Lang Chen, and Josef Lakonishok, 2015, On mutual fund investment styles, *The Review of Financial Studies* 15, 1407–1437.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2006, The adequacy of investment choices offered by 401(k) plans, *Journal of Public Economics* 90, 1299 – 1314.

Elton, Edwin J., Martin J. Gruber, and T. Clifton Green, 2007, The impact of mutual fund family membership on investor risk, *Journal of Financial and Quantitative Analysis* 42, 257—-277.

Griffin, John M., and G. Andrew Karolyi, 1998, Another look at the role of the industrial structure of markets for international diversification strategies, *Journal of Financial Economics* 50, 351–373.

Heston, Steven L., and K. Geert Rouwenhorst, 1994, Does industrial structure explain the benefits of international diversification?, *Journal of Financial Economics* 36, 3–27.

Kaniel, Ron, and Robert Parham, 2017, WSJ Category Kings: The impact of media attention on consumer and mutual fund investment decisions, *Journal of Financial Economics* 123, 337–356.

Kempf, Alexander, and Stefan Ruenzi, 2007, Tournaments in mutual-fund families, *The Review of Financial Studies* 21, 1013–1036.

Kennedy, Peter, 1986, Interpreting dummy variables, *The Review of Economics and Statistics* 68, 174–175.

Koren, Miklós, and Silvana Tenreyro, 2007, Volatility and development, *The Quarterly Journal of Economics* 122, 243–287.

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.

Morningstar, Inc., 2018, Morningstar Style Box™ Methodology.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004, Family values and the star phenomenon: Strategies of mutual fund families, *The Review of Financial Studies* 17, 667–698.

Pástor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Suits, Daniel B., 1984, Dummy variables: Mechanics v. interpretation, *The Review of Economics and Statistics* 66, 177–180.

Teo, Melvyn, and Sung-Jun Woo, 2004, Style effects in the cross-section of stock returns, *Journal of Financial Economics* 74, 367–398.

Vassalou, Maria, and Yuhang Xing, 2004, Default risk in equity returns, *The Journal of Finance* 59, 831–868.

Figure 1
Estimated fund-of-funds variance across diversification benchmarks
This figure gives the equally-weighted FoF variance as the number of funds in the FoF increases, expressed as a percentage of the variance of a typical fund. The top black line is the variance of a fund that diversifies across families within a single category. Its corresponding red line reports this same variance corrected for family effects. The middle black line gives the variance of a FoF that diversifies across categories within a single family. Its corresponding red line reports this same variance corrected for category effects. The bottom FoF diversifies across both families and categories.

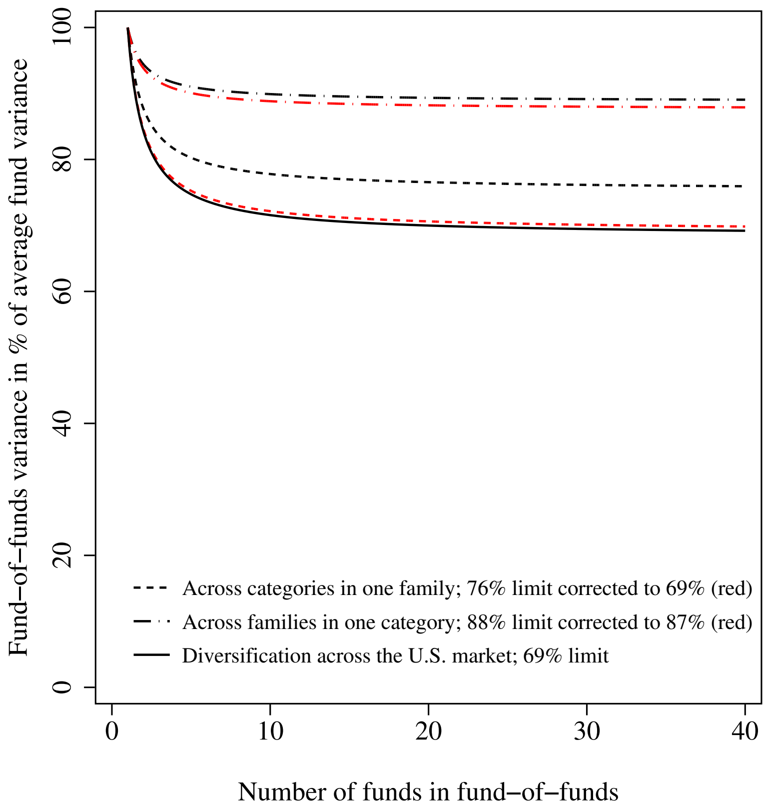
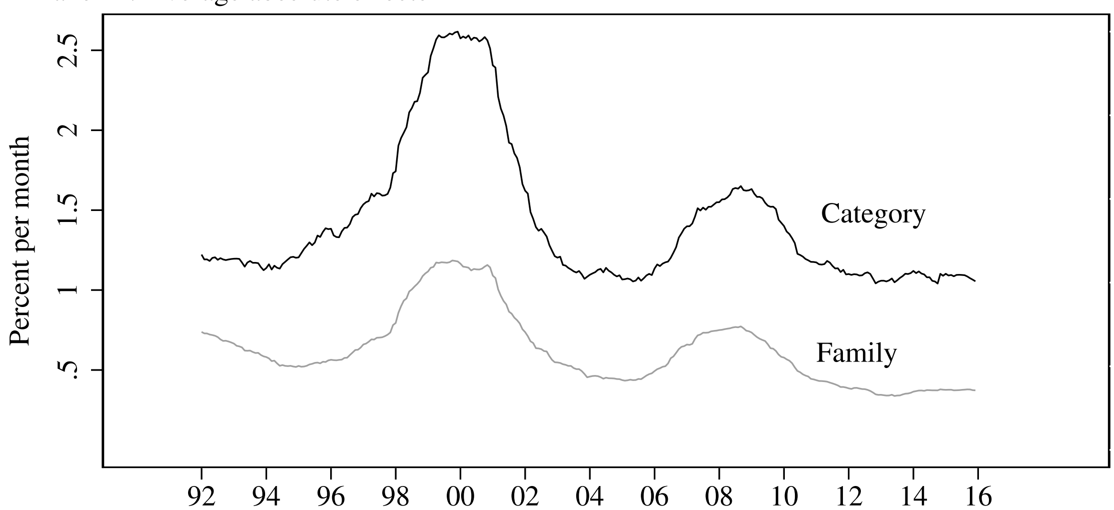


Figure 2
Absolute family and category effects and model fit, June 1992–2016 (moving average)
Panel A gives the 36-month moving average of absolute family and category effects weighted by their respective market capitalizations across the full period in percent per month. Panel B gives the 36-month moving average of the return on the value-weighted U.S. actively managed mutual fund market, measured in percent per month along the left vertical axis. Additionally, Panel B reports the 36-month moving average of the R-squared of the value-weighted cross-sectional regressions, measured in percent per month along the right vertical axis. All moving averages reported are centered with normalized month weights.

Panel A: Average absolute effects



Panel B: Model fit

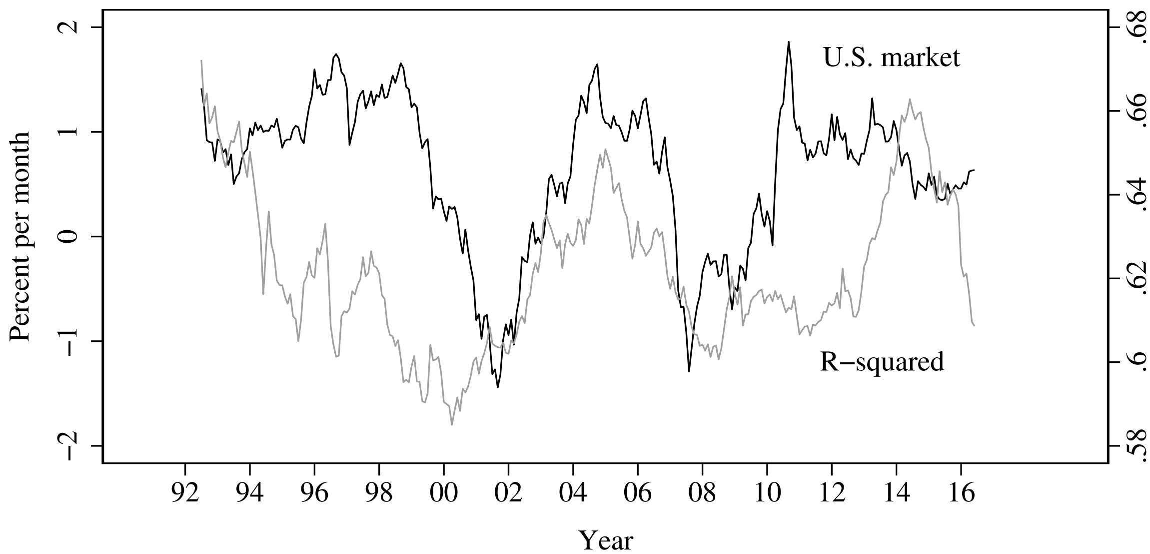


Table 1
Category composition in top mutual fund families
Morningstar Style Box category composition of the top 10 actively managed mutual funds (ranked by average real total assets under management in the time series) for funds 1991–2017. Panel A gives, for each of these 10 families, the number of funds included in the sample. Panel B gives the market share of these families across the time series by the nine Morningstar Style Box categories, expressed as a percentage of the actively managed mutual fund market. Style categories are abbreviated, e.g. Small Growth, Mid-Cap Blend, and Large Value correspond to SG, MB, and LV, respectively.

A: Fund composition by family, category

	Style category										
Family	SG	SV	SB	MG	MV	MB	LG	LV	LB	O	US
Fidelity Investments	0.74	0.00	1.68	2.67	1.05	1.79	6.51	2.85	6.11	43.68	67.07
Columbia	3.70	0.50	0.55	2.63	0.81	1.49	3.47	3.13	4.91	25.45	46.62
Invesco	0.00	0.44	0.60	1.18	0.66	1.38	1.90	2.41	3.10	14.30	25.97
Morgan Stanley	2.37	0.97	0.08	1.22	0.05	0.10	2.55	0.87	1.42	13.67	23.29
John Hancock	2.06	0.23	1.02	1.44	0.79	0.39	1.63	1.29	3.29	7.40	19.56
T. Rowe Price	0.54	0.00	0.11	0.32	0.00	0.06	0.80	0.28	1.07	7.39	10.57
State Street	0.00	0.00	0.00	0.00	0.00	0.00	1.53	0.04	2.47	2.74	6.78
Wells Fargo Funds	1.16	0.07	1.93	0.47	0.39	0.38	2.89	1.56	1.54	7.42	17.81
BlackRock	0.15	0.51	0.65	0.52	0.85	0.15	0.77	0.79	1.27	6.68	12.34
Waddell & Reed	0.00	0.00	0.00	0.00	0.00	0.00	1.35	0.69	0.75	2.43	5.23
Other (O)	44.31	17.73	30.45	53.01	16.88	29.85	81.28	69.79	90.56	426.93	860.79
United States (US)	55.02	20.46	37.06	63.45	21.48	35.60	104.66	83.70	116.50	558.08	1,096.03

Table 1 (continued)

B: Average weights, U.S. value-weighted market

	Style category										
Family	SG	SV	SB	MG	MV	MB	LG	LV	LB	O	US
Fidelity Investments	0.36	0.00	0.42	0.82	1.70	0.91	4.04	1.60	1.85	12.59	24.29
Columbia	0.27	0.01	0.05	0.36	0.08	0.17	0.84	0.58	0.72	1.94	5.01
Invesco	0.00	0.05	0.08	0.14	0.02	0.10	0.55	0.74	1.11	1.32	4.10
Morgan Stanley	0.32	0.08	0.00	0.09	0.00	0.01	0.39	0.27	0.05	1.54	2.75
John Hancock	0.08	0.00	0.03	0.02	0.05	0.01	0.10	0.16	0.41	1.30	2.17
T. Rowe Price	0.11	0.00	0.01	0.00	0.00	0.00	0.05	0.07	0.31	1.08	1.63
State Street	0.00	0.00	0.00	0.00	0.00	0.00	0.42	0.04	0.91	0.28	1.65
Wells Fargo Funds	0.09	0.01	0.07	0.02	0.02	0.02	0.43	0.15	0.16	0.30	1.25
BlackRock	0.00	0.04	0.04	0.06	0.03	0.02	0.03	0.03	0.13	0.77	1.17
Waddell & Reed	0.00	0.00	0.00	0.00	0.00	0.00	0.39	0.07	0.12	0.33	0.91
Other (O)	1.69	0.66	1.47	3.32	0.80	1.77	6.26	6.03	8.53	24.52	55.06
United States (US)	2.93	0.86	2.17	4.82	2.71	3.00	13.51	9.73	14.28	45.99	100.00

Table 2
Summary statistics for monthly panel data, 1991–2017
Panel A summarizes the mean and standard deviation of the monthly equally-weighted and value-weighted family returns for the 10 families reported in Table 1. The bold correlations along the diagonal are the average correlations between fund returns within each family, where the U.S. market figure is equal-weighted. Panel B contains the summary statistics for the monthly returns on the nine Morningstar Style Box categories. The bold correlations along the diagonal are the average correlations between fund returns within each category, where the U.S. market figure is equal-weighted. All returns are expressed in percent per month. The correlations above the diagonal refer to value-weighted returns, and those below the diagonal are between the equally-weighted returns.

A: By family

	EW return		**VW return**		**Correlation matrix**										
	Mean	*SD*	*Mean*	*SD*	F	C	I	M	J	T	S	WF	B	WR	US
Fidelity Investments (F)	0.91	4.42	0.80	4.96	**0.697**	0.966	0.945	0.949	0.891	0.889	0.886	0.925	0.924	0.959	0.981
Columbia (C)	0.74	4.20	0.66	4.23	0.983	**0.809**	0.978	0.965	0.954	0.916	0.946	0.966	0.932	0.955	0.993
Invesco (I)	0.75	4.05	0.67	3.93	0.979	0.985	**0.768**	0.954	0.965	0.925	0.956	0.952	0.918	0.924	0.983
Morgan Stanley (M)	0.76	4.34	0.64	4.15	0.976	0.985	0.979	**0.690**	0.932	0.905	0.917	0.928	0.886	0.934	0.972
John Hancock (J)	0.77	4.42	0.67	4.06	0.968	0.970	0.975	0.966	**0.784**	0.904	0.973	0.940	0.882	0.892	0.952
T. Rowe Price (T)	0.70	4.71	0.73	4.74	0.950	0.951	0.935	0.935	0.910	**0.692**	0.884	0.870	0.858	0.868	0.922
State Street (S)	0.82	3.87	0.87	4.04	0.960	0.955	0.954	0.946	0.944	0.913	**0.929**	0.933	0.863	0.896	0.945
Wells Fargo Funds (WF)	0.87	4.32	0.87	4.49	0.963	0.964	0.970	0.966	0.963	0.900	0.937	**0.815**	0.913	0.930	0.963
BlackRock (B)	0.77	4.25	0.63	4.25	0.948	0.947	0.955	0.947	0.947	0.889	0.908	0.966	**0.579**	0.890	0.937
Waddell & Reed (WR)	0.74	3.88	0.75	4.02	0.956	0.947	0.942	0.943	0.934	0.906	0.937	0.937	0.911	**0.845**	0.962
United States (US)	0.74	4.03	0.69	4.15	0.991	0.993	0.991	0.986	0.983	0.944	0.962	0.978	0.963	0.959	**0.804**

Table 2 (continued)

B: By category

	EW return		VW return		Correlation matrix									
	Mean	*SD*	*Mean*	*SD*	SG	SV	SB	MG	MV	MB	LG	LV	LB	US
Small Growth (SG)	1.06	6.03	0.98	6.20	**0.897**	0.831	0.925	0.969	0.731	0.899	0.896	0.714	0.814	0.893
Small Value (SV)	1.04	4.65	0.99	4.57	0.863	**0.889**	0.926	0.812	0.911	0.900	0.758	0.869	0.840	0.848
Small Blend (SB)	0.96	4.86	0.90	4.91	0.923	0.960	**0.883**	0.909	0.845	0.939	0.850	0.814	0.855	0.902
Mid-Cap Growth (MG)	0.96	5.40	0.89	5.33	0.976	0.833	0.898	**0.871**	0.769	0.936	0.957	0.776	0.883	0.943
Mid-Cap Value (MV)	0.89	4.06	0.88	4.49	0.796	0.939	0.911	0.814	**0.855**	0.904	0.777	0.948	0.904	0.878
Mid-Cap Blend (MB)	0.90	4.25	0.89	4.60	0.905	0.934	0.956	0.929	0.958	**0.839**	0.926	0.901	0.948	0.972
Large Growth (LG)	0.80	4.53	0.77	4.87	0.888	0.794	0.846	0.950	0.845	0.933	**0.888**	0.823	0.940	0.965
Large Value (LV)	0.80	3.88	0.76	3.88	0.742	0.874	0.850	0.784	0.969	0.934	0.865	**0.881**	0.957	0.902
Large Blend (LB)	0.79	3.95	0.80	3.97	0.832	0.850	0.868	0.889	0.930	0.960	0.964	0.961	**0.891**	0.970
United States (US)	0.74	4.03	0.69	4.15	0.905	0.881	0.921	0.940	0.920	0.975	0.958	0.917	0.968	**0.746**

Table 3

Decomposition of excess returns into family and category effects

This table gives the variance of the funds composing the equally-weighted (EW) and value-weighted (VW) excess family and category returns over the U.S. market. Each excess family return is decomposed in a pure family effect and a sum of J category effects,[1] where J is the number of categories that funds in family k belong to in month t.[2] Analogously, each excess category return is decomposed in a pure category effect and the sum of K family effects,[3] where K is the number of families that funds in category j belong to in month t.[4] Returns are measured in percent per month. The ratio relative to the market gives the ratio of the variance of that component to the variance of that component's return in excess of the market.

A: By family	**EW fund-of-funds**				**VW fund-of-funds**			
	Pure family effect		Sum of category effects		Pure family effect		Sum of category effects	
Family	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
Fidelity Investments	0.25	0.615	0.13	0.309	0.64	0.464	0.25	0.180
Columbia	0.10	0.406	0.15	0.581	0.19	0.771	0.07	0.268
Invesco	0.15	0.524	0.18	0.642	0.38	0.644	0.16	0.262
Morgan Stanley	0.18	0.413	0.25	0.576	0.23	0.340	0.55	0.803
John Hancock	0.25	0.437	0.36	0.620	0.58	0.556	0.25	0.245
T. Rowe Price	0.49	0.236	1.52	0.735	0.78	0.222	3.13	0.892
State Street	0.44	0.362	0.52	0.428	0.68	0.373	0.65	0.357
Wells Fargo Funds	0.19	0.415	0.24	0.518	0.36	0.376	0.50	0.529
BlackRock	0.49	0.643	0.42	0.544	0.77	0.452	1.44	0.844
Waddell & Reed	0.70	0.527	0.59	0.443	0.90	0.739	0.25	0.206
Cross-family average	1.04	0.550	1.11	0.568	1.61	0.580	2.04	0.643

[1] The pure family effect measures the average return of funds in a family relative to funds which belong to the same category but a different family.

[2] The sum of the category effects represents the component of a family's return that can be attributed to the difference between its own category composition and that of the U.S. market.

[3] The pure category effect measures the average return of funds in a category relative to funds which belong to the same family but a different category.

[4] The sum of the family effects represents the component of a category's return that can be attributed to the difference between its own family composition and that of the U.S. market.

Table 3 (continued)

B: By category	EW fund-of-funds				VW fund-of-funds			
	Pure category effect		Sum of family effects		Pure category effect		Sum of family effects	
Category	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
Small Growth	7.47	0.958	0.02	0.003	8.91	1.022	0.06	0.007
Small Value	4.65	0.953	0.06	0.013	5.95	0.991	0.19	0.032
Small Blend	3.65	0.963	0.04	0.009	4.11	0.959	0.09	0.021
Mid-Cap Growth	4.09	0.965	0.01	0.003	4.11	1.124	0.17	0.047
Mid-Cap Value	2.38	0.819	0.08	0.028	6.52	1.227	0.35	0.067
Mid-Cap Blend	1.00	1.029	0.04	0.043	1.79	1.225	0.13	0.088
Large Growth	1.61	0.972	0.01	0.004	1.85	0.944	0.03	0.014
Large Value	2.55	0.945	0.01	0.004	2.96	0.871	0.07	0.019
Large Blend	0.97	1.008	0.01	0.006	0.90	0.904	0.03	0.035
Cross-category average	8.27	0.960	0.09	0.036	9.41	1.019	0.28	0.071

Table 4
Summary statistics for estimated return decompositions
Analogous to Table 2, this table gives summary statistics for the estimated return decompositions. Panel A summarizes the mean and the standard deviation of the monthly equally-weighted (EW) and value-weighted (VW) estimated family returns, corrected for category effects. Panel B contains the summary statistics for the monthly returns on equally- and value-weighted category FoFs, corrected for family effects. All returns are expressed in percent per month. The correlations above the diagonal refer to value-weighted FoF returns, and those below the diagonal are between the equally-weighted FoF returns.

A: By family	**EW return**		**VW return**		**Correlation matrix**										
	Mean	*SD*	*Mean*	*SD*	F	C	I	M	J	T	S	WF	B	WR	US
Fidelity Investments (F)	0.87	4.28	0.82	4.65		0.976	0.969	0.978	0.937	0.963	0.958	0.970	0.958	0.965	0.989
Columbia (C)	0.73	4.18	0.67	4.22	0.989		0.985	0.985	0.977	0.974	0.979	0.985	0.961	0.965	0.994
Invesco (I)	0.72	4.00	0.63	3.96	0.987	0.993		0.976	0.975	0.973	0.977	0.978	0.950	0.949	0.989
Morgan Stanley (M)	0.76	4.19	0.67	4.26	0.987	0.990	0.987		0.961	0.963	0.971	0.980	0.957	0.967	0.990
John Hancock (J)	0.68	4.05	0.59	3.94	0.978	0.988	0.984	0.980		0.959	0.976	0.965	0.933	0.931	0.971
T. Rowe Price (T)	0.83	4.13	0.72	4.22	0.976	0.983	0.982	0.977	0.983		0.963	0.961	0.952	0.938	0.977
State Street (S)	0.85	4.04	0.78	4.09	0.979	0.985	0.985	0.980	0.981	0.977		0.970	0.937	0.944	0.980
Wells Fargo Funds (WF)	0.81	4.21	0.75	4.23	0.985	0.989	0.986	0.986	0.977	0.974	0.977		0.959	0.961	0.988
BlackRock (B)	0.78	4.27	0.73	4.58	0.971	0.972	0.965	0.971	0.965	0.964	0.956	0.975		0.936	0.966
Waddell & Reed (WR)	0.80	3.97	0.76	4.00	0.972	0.974	0.965	0.971	0.956	0.957	0.955	0.974	0.956		0.972
United States (US)	0.74	4.03	0.69	4.15	0.993	0.998	0.994	0.993	0.988	0.985	0.987	0.993	0.977	0.977	

Table 4 (continued)

B: By category	EW return		VW return		Correlation matrix									
	Mean	*SD*	*Mean*	*SD*	SG	SV	SB	MG	MV	MB	LG	LV	LB	US
Small Growth (SG)	1.05	5.95	1.01	6.21		0.839	0.923	0.969	0.696	0.888	0.899	0.721	0.824	0.893
Small Value (SV)	1.02	4.72	1.01	4.69	0.874		0.922	0.812	0.912	0.906	0.768	0.861	0.838	0.852
Small Blend (SB)	0.91	4.82	0.85	4.87	0.918	0.957		0.903	0.827	0.933	0.845	0.824	0.856	0.900
Mid-Cap Growth (MG)	0.95	5.36	0.92	5.47	0.974	0.844	0.894		0.725	0.925	0.963	0.783	0.891	0.945
Mid-Cap Value (MV)	0.90	4.13	0.84	4.43	0.820	0.941	0.925	0.840		0.888	0.736	0.937	0.872	0.848
Mid-Cap Blend (MB)	0.88	4.32	0.88	4.69	0.901	0.935	0.952	0.928	0.971		0.919	0.914	0.948	0.968
Large Growth (LG)	0.82	4.55	0.80	4.86	0.890	0.805	0.843	0.952	0.865	0.936		0.834	0.947	0.965
Large Value (LV)	0.79	3.89	0.76	3.91	0.751	0.872	0.858	0.795	0.967	0.942	0.870		0.956	0.912
Large Blend (LB)	0.80	3.96	0.79	3.99	0.831	0.851	0.866	0.890	0.939	0.963	0.964	0.963		0.973
United States (US)	0.74	4.03	0.69	4.15	0.905	0.886	0.920	0.942	0.935	0.976	0.959	0.922	0.967	

Appendix

A. Tables

Table A.1
Decomposition of excess returns into family and category effects by period
Analogous to Table 3, this table gives the variance of the funds composing the equally-weighted (EW) and value-weighted (VW) excess family and category returns over the U.S. market by period.

A: By family	**EW fund-of-funds**				**VW fund-of-funds**			
	Pure family effect		Sum of category effects		Pure family effect		Sum of category effects	
Period	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
1991–1999	1.85	0.535	1.62	0.566	2.35	0.624	2.82	0.648
2000–2007	1.25	0.597	1.15	0.559	1.89	0.618	2.21	0.680
2008–2009	1.57	0.597	1.52	0.559	2.73	0.683	2.45	0.608
2010–2017	0.67	0.521	0.84	0.600	1.02	0.549	1.52	0.666

B: By category	**EW fund-of-funds**				**VW fund-of-funds**			
	Pure category effect		Sum of family effects		Pure category effect		Sum of family effects	
Period	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
1991–1999	17.63	0.993	0.25	0.026	22.24	1.071	0.56	0.078
2000–2007	8.74	1.063	0.13	0.060	10.47	1.072	0.45	0.086
2008–2009	14.19	0.971	0.12	0.035	15.76	0.983	0.59	0.168
2010–2017	5.85	0.965	0.05	0.032	6.27	1.010	0.16	0.066

Table A.2
Decomposition of excess returns into family and category effects by number of funds
Analogous to Table 3, this table gives the variance of the funds composing the equally-weighted (EW) and value-weighted (VW) excess family and category returns over the U.S. market by number of funds.

A: By family	**EW fund-of-funds**				**VW fund-of-funds**			
	Pure family effect		Sum of category effects		Pure family effect		Sum of category effects	
Number of funds	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
3–9	1.05	0.527	1.42	0.714	1.48	0.723	2.68	0.963
10–19	0.91	0.512	1.22	0.603	1.54	1.066	2.09	1.296
20–29	1.04	0.931	1.33	0.954	1.74	1.635	2.25	1.807
30+	1.05	0.548	1.09	0.550	1.65	0.575	2.02	0.639

B: By category	**EW fund-of-funds**				**VW fund-of-funds**			
	Pure category effect		Sum of family effects		Pure category effect		Sum of family effects	
Number of funds	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
3–9	8.48	0.971	0.09	0.030	9.46	1.025	0.28	0.063
10–19	7.86	0.957	0.08	0.037	8.96	1.016	0.26	0.073
20–29	10.00	0.959	0.13	0.031	12.17	1.026	0.36	0.055
30+	6.79	0.967	0.07	0.039	7.62	1.005	0.27	0.076

Table A.3

Decomposition of excess returns into family and category effects by size

Analogous to Table **3**, this table gives the variance of the funds composing the equally-weighted (EW) and value-weighted (VW) excess family and category returns over the U.S. market by family (category) size. Size is measured by total assets under management in each month and reported in quintiles.

A: By family	**EW fund-of-funds**				**VW fund-of-funds**			
	Pure family effect		Sum of category effects		Pure family effect		Sum of category effects	
Size quintile	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
1	1.15	0.640	1.28	0.662	1.66	0.627	2.17	0.669
2	1.02	0.994	1.08	1.139	1.59	0.796	2.15	0.957
3	1.28	0.516	1.00	0.545	2.20	0.554	2.71	0.613
4	1.61	0.457	1.03	0.527	3.25	0.513	2.41	0.609
5	1.93	0.581	0.66	0.679	3.68	0.564	1.08	0.716

B: By category	**EW fund-of-funds**				**VW fund-of-funds**			
	Pure category effect		Sum of family effects		Pure category effect		Sum of family effects	
Size quintile	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market	*Var.*	Ratio to market
1	8.90	0.965	0.12	0.042	10.28	1.032	0.37	0.092
2	5.58	0.989	0.05	0.043	6.73	1.031	0.29	0.099
3	7.00	0.964	0.02	0.023	7.03	1.014	0.11	0.086
4	10.15	0.978	0.01	0.005	8.03	1.056	0.40	0.047
5	1.43	0.997	0.01	0.010	1.59	0.987	0.03	0.050