Daniel Fricke

Are specialist funds "special"?

Article (Accepted version)
(Refereed)

Original citation:
Fricke, Daniel (2018) Are specialist funds "special"? Financial Management. ISSN 1755-053X

DOI: https://doi.org/10.1111/fima.12257

Are Specialist Funds "Special"?

Daniel Fricke*

In this paper, I explore the relation between portfolio overlap and performance diversity. Using data on actively managed U.S. equity mutual funds, I find that the pairwise portfolio overlap between individual funds has increased over time and is significant compared to various randomized benchmarks. These findings motivate the main question of this paper, namely whether specialist funds (those with low levels of portfolio overlap with other funds) differ significantly from funds with high levels of overlap. Here, I find that these specialists differ with regard to certain portfolio- and fund-specific characteristics, but they do not appear to outperform other funds.

I am grateful for helpful comments from the editor (Uptal Bhattacharya), an anonymous reviewer, participants of the Portsmouth-Fordham Conference on Banking & Finance 2016, the Royal Economic Society Meeting 2017, and seminar participants at University College London, the University of Oxford, and the Deutsche Bundesbank. Special thanks go to Felix Reed-Tsochas and Hannes Wilke for numerous helpful discussions. The views expressed in this paper represent the author's personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank or its staff.

Daniel Fricke is a Research Economist at the Deutsche Bundesbank in the Directorate General Financial Stability. He is also affiliated with the Financial Computing & Analytics group at University College London, the LSE Systemic Risk Centre, and the CabDyN Complexity Centre at Saïd Business School at Oxford University.

A growing literature highlights the importance of overlapping portfolios on market dynamics and financial stability (e.g., Wagner, 2011; Greenwood, Landier, and Thesmar, 2015). In this paper, I take a different perspective on portfolio overlap by exploring its relation with performance diversity. The basic idea is simple: if two investors hold similar portfolios (i.e., high levels of portfolio overlap), their performances should be indistinguishable. Hence, depending on the levels of portfolio overlap in a given system, it may be difficult to detect investors with superior performances.

I investigate this relation for the set of open-ended, actively managed U.S. domestic equity mutual funds. The main motivation is illustrated in Figure 1, which shows the "typical" structure of the so-called *holdings matrix* at a certain point in time (here for March 2003). Put simply, the figure shows whether a given mutual fund (rows) holds a given stock (columns) in its portfolio: this is indicated as a black dot. Given the focus on actively managed funds, I have dropped index funds and funds with more than 300 stock holdings (closet indexers). The rows/columns in the figure are sorted according to their number of connections. This reordering shows a "triangular" matrix structure: some mutual funds hold many stocks in their portfolios (funds closer to the top of the figure), whereas others focus on only very few stocks (closer to the bottom of the figure).[1] The same is true from the other side: some stocks are held by practically all funds (closer to the left of the figure), and others are held by only a much smaller subset (closer to the right of the figure). The most interesting feature is that funds with very few connections tend to hold those stocks that are held by all other funds as well—otherwise there would be more connections in the bottom right part of the figure—which highlights the portfolio overlap among these mutual funds. (I will introduce concrete measures of portfolio overlap below.) This finding is remarkable given that my focus is on actively managed funds here, which supposedly aim to outperform both the market and other actively managed funds. The red line shows the cumulative portfolio share of the stocks shown on the x-axis relative to the total holdings of the funds in the sample. In line with the above reasoning, the economically meaningful investments are concentrated on a relatively small number of stocks: for the set of funds shown in figure 1, 80% (90%) of the total holdings are concentrated in only 358 (623) stocks.[2]

Taken together, several questions arise from Figure 1: Is the portfolio overlap between mutual funds significant? Does portfolio overlap increase or decrease over time? What is the relation between portfolio overlap and performance diversity? I tackle these questions in this paper.

Based on data for the period March 2003—December 2014, I quantify the pairwise

[1] Figure A.1 in the Appendix shows that a similar matrix structure can be found when aggregating stocks to the industry level. In other fields, such as ecology, such a triangular structure is called nested (e.g., Bascompte et al., 2003).

[2] This concentration is in line with the observation that the S&P 500 market index (comprising the largest listed companies in the United States) captures around 80% of the available market capitalization. See `https://us.spindices.com/indices/equity/sp-500`.

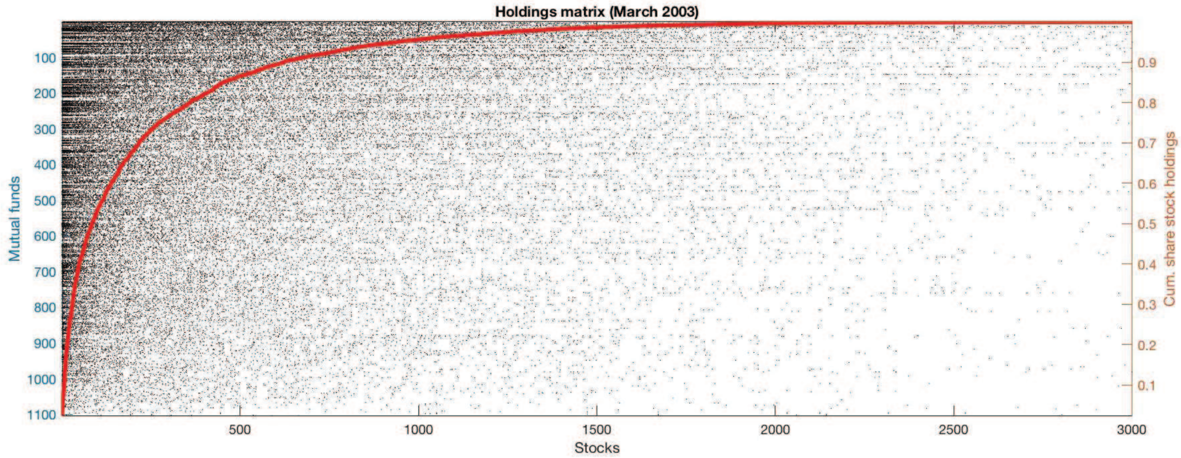

Figure 1. Binary Holdings Matrix in March 2003
Rows correspond to actively managed domestic equity mutual funds with at least 3 and at most 300 stock holdings (left y-axis), and columns to stocks. A link between a fund and a stock exists if the fund holds that particular stock: this is shown as a black dot. Rows and columns are sorted according to the number of connections. The red line shows the cumulative share of funds' holdings in these stocks relative to their total holdings (right y-axis).

portfolio overlap between actively managed mutual funds (mainly based on the cosine similarity measure using portfolio weights) for each month. My main findings are as follows: First, the typical value of the pairwise overlap measure increases over time, such that funds' portfolios have become more similar. This finding is remarkable given that actively managed funds have been reported to be particularly affected by increasing levels of competition (e.g., due to the rise of passive investment strategies (Malkiel, 2013), which should have a negative impact on funds' portfolio overlap because (increased) competition may provide incentives to construct innovative investment portfolios (e.g., Aghion et al., 2005; Sun, Wang, and Zheng, 2012). This, however, appears not to be the case for the set of mutual funds under study in this paper. Second, I assess the significance of the observed portfolio overlap by evaluating the hypothetical overlap that would arise if I disregarded funds' preferences for certain stocks (but fixed several basic portfolio characteristics, including the observed distribution of portfolio weights). In all cases I find that the observed portfolio overlap significantly exceeds these hypothetical benchmark values, indicating that the observed portfolio overlap is significant relative to these benchmarks. Third, I shed light on the relation between portfolio overlap and performance diversity. In line with the factor structure of portfolio returns, I find that even modest levels of portfolio overlap can imply substantial return correlations—that is, low levels of performance diversity. Last, I explore this finding in more detail by identifying *specialist* funds (those with low levels of portfolio overlap with all other funds) and testing whether these differ significantly from funds with high levels of overlap. Here, I find that specialists differ with regard to certain portfolio- and fund-specific characteristics (e.g., in terms of their total net assets [TNA]) but they do not appear to outperform other funds. Thus, specialists are not that special.

My paper is related to three streams of literature. First, I already mentioned the growing literature on portfolio overlap and market stability.[3] In essence, much of this literature predicts that higher levels of portfolio overlap can destabilize the system as a whole, but so far, little is known empirically about realistic levels of portfolio overlap. I quantify the portfolio overlap among a specific set of mutual funds, and the methodology can be generally applied to other kinds of financial institutions as well.

Second, the paper adds to the empirical literature on investor performance. The broad consensus is that detecting investors with superior performance is difficult and many actively managed mutual funds even generate returns (after fees and expenses) that are significantly lower than those of passive index-based strategies (e.g., Jensen, 1968; Grinblatt and Titman, 1989; Carhart, 1997; Berk and Green, 2004; Kacperczyk, Sialm, and Zheng, 2005, 2007; Cremers and Petajisto, 2009; Barras, Scaillet, and Wermers, 2010). These observations have been linked with the rise of passive investment strategies over the last decade, which puts additional pressure on active fund managers to come up with novel strategies to outperform the market. From this perspective, my findings are important in at least two ways: on the one hand, I find no evidence of an increase in the level of diversity of actively managed funds' asset portfolios. In other words, despite the increased competition within the mutual fund sector, there is no apparent tendency for innovation in funds' investment strategies. On the other hand, the fact that even those funds with the most "special" portfolios do not significantly outperform other funds implies that low levels of performance diversity can be in line with relatively low levels of portfolio overlap. By purely looking at the cross-section of investment portfolios, policy makers, regulators, and investors may get a false sense of (portfolio) diversity.

Last, my paper is related to the literature on herding among professional investors.[4] Conceptually, herding and portfolio overlap are strongly related (as illustrated, among others, by Coval and Stafford, 2007): on the one hand, herding is concerned with correlated trading in and out of the same assets (flow perspective). On the other hand, portfolio overlap results from the herding dynamics of investors (stock perspective). To the best of my knowledge, most of the literature has focused on the flow perspective, whereas I am explicitly interested in the stock perspective.

The remainder is structured as follows: Section I. reviews the relevant literature. Section II. defines measures of portfolio overlap. Section III. introduces the dataset and contains the empirical analyses. Section IV. concludes.

[3] A related literature explores the role of common ownership on the price/liquidity dynamics of individual assets (e.g., Antón and Polk, 2014).

[4] For example, Grinblatt, Titman, and Wermers (1995) find that roughly 77% of mutual funds are momentum investors—that is, buy stocks that were past winners. Wermers (1999) finds evidence of herding in mutual fund trading for small and growth-oriented stocks, which affects stock prices.

I. Literature Review

Let me briefly review the relevant literature that sheds light on the equilibrium set of portfolios shown in the cross-section of Figure 1. Which assets an investor wants to hold (security selection) remains the key aspect of investing. Given estimates of expected returns and covariances, modern portfolio theory produces the optimal portfolio weights for the relevant assets of interest. In the absence of heterogeneity and/or frictions, the capital asset pricing model (CAPM) suggests that in equilibrium the only portfolio that will be held is the market portfolio. In other words, in a CAPM-world, all investors will hold exactly the same diversified portfolio and thus portfolio overlap should be maximal. Existing empirical evidence (including Figure 1), however, suggests that many investors deviate from these theoretical predictions and tend to hold much more concentrated portfolios. Several theoretical contributions show why investors might indeed rationally choose different diversification levels in the presence of frictions/heterogeneity.[5]

Taking as given that investors might differ in their diversification levels does not tell us which securities they will actually select. In a similar fashion, the literature on (institutional) herding establishes different reasons for why investors might rationally prefer highly correlated performances with others (e.g., information cascades (Banerjee, 1992), correlated information (Shleifer and Summers, 1990), and incentive constraints. For example, it has been shown that when managers' remuneration depends on their performance relative to their peers, they will rationally choose highly correlated performances (e.g., Scharfstein and Stein, 1990; Chevalier and Ellison, 1999; Rajan, 2005). Hence, return correlations should be high if such career concerns are relevant for managers competing in the same job market—something that seems reasonable for the set of professional mutual fund managers.

In most of the above literature, the representative investor's optimization problem is typically absent of price effects due to asset liquidations. When a large fund faces unexpected redemptions, it will have to liquidate some of its assets on the market (Coval and Stafford, 2007). In an otherwise standard portfolio optimization framework, Caccioli et al. (2013) allow investors to take into account their *own* impact on asset prices (in case of liquidation). Thus, investors' trade-off asset's risk-adjusted return and its illiquidity against each other and, ceteris paribus, by taking market impact into account will reduce the exposure to illiquid assets.[6] Similarly, Lo, Petrov, and Wierzbicki (2003) construct a three-dimensional mean-variance-liquidity frontier, showing that portfolios close to each

[5] Examples include transaction costs (Constantinides, 1986), costs of information acquisition (Van Niewerburgh and Veldkamp, 2010), and borrowing constraints (Roche, Tompaidis, and Yang, 2013).

[6] Interestingly, Caccioli et al. (2013) find that, depending on the assumed market impact function, this approach can stabilize the optimization of large portfolios. A linear market impact function effectively adds an L_2-constraint on the portfolio weights in the optimization, and it is well known that such a constraint tends to yield more diversified portfolios. On the other hand, adding an L_1 constraint (in case of a bid-ask spread market impact function) yields sparser portfolios.

other on the mean-variance frontier can differ dramatically in terms of their liquidity.

Given the results of the literature on fire sales, it appears reasonable that mutual funds also take *others'* potential price impacts into account, in particular when those funds are subject to high levels of liquidation risk. For example, Greenwood and Thesmar (2011) argue that an asset's price can be fragile because of concentrated ownership or because its owners face correlated liquidity shocks. Based on this reasoning, they construct a fragility measure that predicts future volatility. Therefore, assuming that funds know about other funds' specific asset holdings (something that seems reasonable for mutual funds)[7], then they could take these systemic asset liquidations into account in their portfolio construction. This is exactly the framework of Wagner (2011), where it is shown that if assets are less than perfectly liquid and investors care about systemic liquidation risks, no two investors will hold exactly the same portfolio in equilibrium. The basic idea is that in order to reduce the risk of joint liquidation, investors want to distinguish themselves from one another and minimize their portfolio overlap with others. Another theoretical argument in favor of relatively low levels of portfolio overlap comes from the theory of industrial organization. A common prediction of this literature is that companies have incentives to avoid competition by creating products/services that cannot be easily replicated by their competitors (Aghion et al., 2005). The mutual fund sector has been shown to be highly competitive (Wahal and Wang, 2011) and in light of the rise of passive investment strategies over the last decade, competition among active managers is likely to have increased (see Sun et al. [2012] for similar arguments for the hedge fund industry). This should put pressure on active managers to come up with novel strategies to outperform the market.

Wagner (2011) also shows that portfolio diversity should increase with higher levels of (joint) liquidation risk. Hence, it seems reasonable that when there is heterogeneity in liquidation risk, individual funds may optimally choose different levels of portfolio overlap. This might then affect funds' performances. Both theoretical and empirical work has mainly focused on the question whether concentrated investors outperform more diversified ones, rather than on an investor's level of portfolio overlap. In this regard, various theories predict that uninformed investors should hold diversified portfolios, while those with superior information about certain assets should concentrate on these (Levy and Livingston, 1995; Van Niewerburgh and Veldkamp, 2010). Empirical evidence suggests that concentration can be beneficial in terms of performance (Kacperczyk et al., 2003, 2007).

To the best of my knowledge, this paper contains the first empirical analysis that explores whether funds with low levels of portfolio overlap (which may have superior information relative to other funds) tend to generate superior returns. It will become clear that my measure of portfolio overlap is conceptually related to, but distinct from, standard measures of portfolio diversification.

[7] For example, U.S. mutual funds have to report their holdings every quarter to the Securities and Exchange Commission (SEC), which are then published at the SEC's website.

II. Measuring Portfolio Overlap

Here, I briefly introduce the general setup: let $i \in \{1, \cdots, N_t\}$ be the set of investors (mutual funds) and $k \in \{1, \cdots, K_t\}$ the set of investments (stocks), where N_t and K_t are the total number of investors and possible investments at time t. At any given point in time, investors' portfolios can be represented as a weighted incidence matrix $\mathbf{W}_t$ of dimension $(N_t \times K_t)$, where each element $w_{i,k,t} \geq 0$ gives the portfolio weight of stock k in investor i's portfolio. By construction, $\sum_k w_{i,k,t} = 1$ and $w_{i,k,t} \geq 0 \; \forall i, k, t$ (short positions are disregarded).

I am also interested in the extensive margin and $\mathbf{B}_t$ denotes the binary version of $\mathbf{W}_t$, where each element $b_{i,k,t} = 1$ if $w_{i,k,t} > 0$ and zero otherwise. In other words, an investor and a stock are "connected" if the investor holds that stock in his/her portfolio. Hence, I refer to the binary incidence matrix $\mathbf{B}_t$ as the *holdings matrix*. In network jargon, I am dealing with bipartite networks ($\mathbf{W}_t$ is weighted, $\mathbf{B}_t$ is binary), where connections can only exist between mutual funds and stocks but not within the two sets.

In order to quantify the overlap of two funds' portfolios, I will use two different measures: one based on the binary version of the holdings matrix (BinOverlap) and one based on the matrix of portfolio weights (Overlap).[8] I define binary overlap between investors i and j $(i \neq j)$ as

$$\text{BinOverlap}_{i,j,t} = \frac{\sum_k b_{i,k,t} \times b_{j,k,t}}{\min(\sum_k b_{i,k,t}, \sum_k b_{j,k,t})} = \frac{K_{i,j,t}}{\min(K_{i.t}, K_{j,t})}, \tag{1}$$

which is the number of stocks jointly held by investors i and j $(K_{i,j,t})$, relative to the maximum possible number of joint holdings $(\min(K_{i,t}, K_{j,t}))$.[9] Note that BinOverlap is based on the extensive margin only (the presence of a given stock in a portfolio).

In order to also take the intensive margin into account, my second measure is based on the observed portfolio weights. We define

$$\text{Overlap}_{i,j,t} = \frac{\sum_k w_{i,k,t} w_{j,k,t}}{\sqrt{\sum_k w_{i,k,t}^2} \times \sqrt{\sum_k w_{j,k,t}^2}}. \tag{2}$$

This is the so-called cosine similarity measure, which is technically defined as the angle between the vectors of portfolio weights between fund i and fund j. Note that the denominator is simply the product of the square-root of the two investors' Hirschmann-Herfindahl Indices (HHIs):

$$\text{Overlap}_{i,j,t} = \frac{\mathbf{w}_{i,t} (\mathbf{w}_{j,t})^T}{\sqrt{\text{HHI}_{i,t}} \times \sqrt{\text{HHI}_{j,t}}}, \tag{3}$$

where $\mathbf{w}_{i,t}$ is the vector of portfolio weights of investor i. This shows that Overlap is related

[8] From a network perspective, the following matrices of portfolio overlap can be understood as a uni-partite projection of the original bipartite fund-stock network.

[9] The measure in Equation (1) is closely related to the so-called Hamming similarity measure that is widely used in information theory and computer science.

to, but distinct from, a fund's level of portfolio concentration (HHI). It also becomes clear that the numerator can be seen as a measure of cross-diversification in the joint investments of funds i and j.

Both overlap measures range between zero and one. For example, if two funds have no common assets, both measures are equal to zero; if they hold the same portfolios, these measures equal one. From a theoretical perspective, Overlap should be a more adequate measure (compared to BinOverlap) because it takes portfolio weights into account. I will therefore focus most of the attention on this measure.

Later in this paper, I will identify *specialist funds* based on their average portfolio overlap with all other funds

$$\text{MeanOverlap}_{i,t} = \frac{1}{N_t - 1} \sum_j \text{Overlap}_{i,j,t} \, . \tag{4}$$

From a network perspective, MeanOverlap corresponds to the (normalized) weighted degree of funds based on the projection of the original fund-stock holdings data.[10] MeanOverlap can be seen as an inverse measure of how *special* investor i's portfolio is: portfolios with low levels of MeanOverlap are special in the sense that they show little overlap with other mutual funds.[11] I therefore refer to these funds as *specialists* in the following. Of course, my specialists are not to be confused with the "Specialists" that acted as market makers on the New York Stock Exchange (NYSE); since these are nowadays referred to as "Designated Market Makers", I see no risk of confusion here (even more so because this paper is explicitly about mutual funds).

For the sake of completeness, I also take a look at the typical levels of portfolio overlap over time. For this purpose, I calculate the average values of portfolio overlap for each cross-section:

$$\langle \text{MeanOverlap} \rangle_t = \frac{1}{N_t} \sum_i \text{MeanOverlap}_{i,t}. \tag{5}$$

$\langle \text{MeanBinOverlap} \rangle_t$ is defined accordingly.

[10] Alternatively, instead of averaging over the pairwise portfolio overlap for each mutual fund i, I also calculated $\text{MeanOverlap}_i^{\text{mkt}} = \frac{\mathbf{w}_{i,t}(\mathbf{w}_{\text{mkt,t}})^T}{\sqrt{\text{HHI}_{i,t}} \times \sqrt{\text{HHI}_{\text{mkt,t}}}}$, where *mkt* refers to the market portfolio. This measure has the advantage of being even simpler in its computation. It turns out that it is almost perfectly correlated (Pearson-correlation of 0.981) with MeanOverlap in Equation (4) in the sample under study. The downside is that there is an implicit assumption that a typical mutual fund holds the market portfolio, thus ignoring the fact that the pairwise portfolio overlap is zero (or at least very small) between many fund-pairs. Not surprisingly, this alternative measure yields higher average values (0.314) compared with my main measure of interest (0.106; see Table I).

[11] Naturally, as one might expect from Equation (3), a given fund's portfolio overlap is negatively related to its concentration levels HHI. In my sample, I find an unconditional correlation of –0.251 between MeanOverlap and HHI.

III. Empirical Analysis

A. Data Set

My data come from different sources. I obtain mutual funds' portfolio holdings and additional fund-specific information from the Center for Research in Security Prices (CRSP) Survivor-Bias-Free Mutual Fund database. Following the literature, I aggregate different share classes to the fund level. The database provides accurate fund-specific information at the monthly level (such as fund returns and TNA), and detailed information on funds' portfolio holdings from March 2003 onwards.[12] My final sample contains 142 months during the period March 2003—December 2014. I update a given fund's portfolio whenever new holdings are reported.[13] In everything that follows, I disregard short positions.

I complement the mutual fund data with daily stock-specific information from the merged CRSP-Compustat data. In line with the literature (e.g., Antón and Polk, 2014), I focus on common stocks (share codes 10 and 11) traded on NASDAQ, NYSE, or AMEX. The merged data set gives detailed information on the domestic equity holdings of U.S. mutual funds. Therefore, I restrict myself to the set of *domestic equity* (DE) funds for which the diversification measure is likely to be most accurate.[14] I also drop index funds, funds with fewer than 24 monthly observations, fund observations with TNA below US$1 million, and/or fewer than 3 or more than 300 stock holdings (closet indexers). The final sample contains 5,887 unique funds, 6,664 unique stocks, and 295,973 fund-month observations.

B. System Size, Growth, and Overlap Dynamics

Table I illustrates the dynamics of the dataset. The first column shows that the size of the system (sum of total assets under management) has roughly doubled over this period, with the latest value being US$ 3.25 trillion. This growth has been relatively steady, except for the financial crisis period. The second column shows that the growth is largely driven by an increasing number of mutual funds, which has roughly doubled over the sample period (from 1,241 to 2,126). Hence, the level of competition appears to have increased among the set of actively managed funds under study here. On the other hand, the number of unique stocks that were held by these funds decreased slowly over the same period (from 3,474 to 2,635).[15] The number of stocks is generally larger than the number of mutual funds but

[12] Prior to 2003, mutual funds' equity holdings can be obtained from the Thomson Reuters Mutual Fund Ownership Database. In order to merge these holdings data with the CRSP Mutual Fund Database, a linking table is needed (the so-called MFLINKS table developed by Russ Wermers). Unfortunately, I do not have access to the MFLINKS table and therefore have to restrict my sample to post-2003 data.

[13] Note that there is a structural break in the fund identifiers in the CRSP Mutual Fund database: all fund IDs were replaced with new ones between the third and fourth quarter of 2010.

[14] More specifically, in line with the existing literature, I keep only funds with CRSP objective codes beginning with "ED".

[15] These patterns are in line with the results in Doidge et al. (2017). In comparison, the numbers reported here are generally lower, which is due to the fact that some stocks are not held by mutual funds.

Year	Total Assets (in US$ tn)	Mutual Funds	Stocks	⟨Mean-BinOverlap⟩	⟨Mean-Overlap⟩
2003	1.639	1,241	3,474	0.134	0.093
2004	1.805	1,254	3,471	0.155	0.101
2005	1.897	1,249	3,421	0.166	0.104
2006	2.046	1,393	3,380	0.176	0.110
2007	2.113	1,451	3,268	0.180	0.111
2008	1.394	1,768	3,192	0.177	0.110
2009	1.725	1,701	3,117	0.188	0.114
2010	2.867	2,625	2,898	0.131	0.090
2011	2.474	2,519	2,843	0.161	0.107
2012	2.565	2,414	2,760	0.176	0.114
2013	3.185	2,215	2,660	0.187	0.117
2014	3.250	2,126	2,635	0.193	0.116
Mean	2.175	1,794.261	3,119.599	0.166	0.106
Median	2.046	1,710	3,186	0.174	0.108
SD	0.586	493.979	293.396	0.022	0.010
Min	0.998	1,104	2,623	0.103	0.075
Max	3.284	2,625	3,537	0.194	0.117

Table I. Summary Statistics for the U.S. Domestic Equity Mutual Fund Sector The table shows the year-end values for total assets under management, the number of active mutual funds (domestic equity), the number of active stocks, and the average portfolio overlap (the average values of Equations (1) and (2), respectively), and the corresponding summary statistics over the entire sample period.

roughly comparable at the end of the sample period. To a certain extent, the decreasing number of stocks should lead to a somewhat mechanical increase of portfolio overlap in the sense that, ceteris paribus, fewer investment opportunities are available over time.

In this regard, Table I indeed shows an increasing trend in cross-sectional average values of portfolio overlap, ⟨MeanBinOverlap⟩ and ⟨MeanOverlap⟩, respectively. For the binary measure, the value is 0.193 at the end of the sample period (compared to 0.134 at the beginning), meaning that the average pair of mutual funds shares around 19% of stocks in their portfolios at the end of the sample period (average value across all years in 0.166). For the weighted measure, the average value is 0.106, with a value of 0.093 (0.116) at the beginning (end) of the sample.

To test whether these time trends are significant, I regress the cross-sectional average portfolio overlap measures on their own lag and/or a time trend.[16] Table II shows the results: there is indeed a small but significant positive time trend in both overlap measures. This finding is remarkable given that actively managed funds have been reported to be particularly affected by increasing levels of competition (e.g., due to the rise of passive investment strategies (Malkiel, 2013), which should have a negative impact on funds' portfolio overlap since competition may serve as an incentive to construct innovative investment portfolios (e.g., Aghion et al. , 2005; Sun et al. (2012)). The results indicate that this is not

[16] Note that these regressions are based on quarterly data, because most fund portfolios are not updated on a monthly basis.

	Quarterly Data			
	$\langle \text{MeanBinOverlap} \rangle_t$		$\langle \text{MeanOverlap} \rangle_t$	
First lag	0.872**	0.836 **	0.809**	0.709**
	(0.017)	(0.019)	(0.036)	(0.044)
Time trend		8.8e-05 **		9.3e-05**
		(2.1e-05)		(2.8e-05)
Constant	0.024**	0.028 **	0.022**	0.030**
	(0.003)	(0.003)	(0.003)	(0.004)
Adj. R^2	0.983	0.986	0.921	0.936
Observations	43	43	43	43

** Significant at the 0.01 level.
* Significant at the 0.05 level.

Table II. Testing for Time Trends in Portfolio Overlap
Here, I regress the average overlap on its own first lag and a time trend. Results are based on ordinary least squares (OLS) with robust standard errors (in parentheses).

the case—or at least this component is not sufficient to undo a purely mechanical increase of portfolio overlap that arises due to a smaller number of stocks that mutual funds can invest in (cf. Table 1).[17]

Given that the time dynamics shown in Tables I and II are comparable for the two different overlap measures, I restrict ourselves to the weighted measure (Overlap) in everything that follows. So far the analysis focused on the cross-sectional average values of Overlap. Further evidence of a positive time trend on the fund-pair level is provided in Figure 2, which shows the empirical distributions of portfolio overlap. The left panel plots the cross-sectional cumulative distribution functions of the pairwise Overlap values (based on Equation (2)) for the end-of-year values in 2003, 2008, and 2014. Compared to the beginning of the sample, there is more probability mass on larger Overlap values in later years, in line with the overall time trends shown in Tables I and II. I also tested whether the Overlap values shown in Figure 2 are drawn from the same underlying distribution. Separate two-sample Kolmogorov-Smirnov tests between the different year-pairs all strongly reject this hypothesis (p-values <0.001). Lastly, the right panel of Figure 2 shows the probability distributions of fund-specific MeanOverlap values (as defined in Equation 4) for the same data. It becomes apparent that the distribution appears to have become bimodal over time, with most of the probability mass concentrated on values around 0.05, and another peak at values between 0.15 and 0.20.

C. On the Significance of Portfolio Overlap

Given the results from the previous subsection, I now wish to assess whether the typical portfolio overlap—that is, $\langle \text{MeanOverlap} \rangle$—is significant. While a broad literature explores

[17] I confirm that all of the randomization approaches discussed in the next section also yield a positive time trend in the MeanOverlap measure. As such, at least part of the positive time trend in the actual MeanOverlap is due to the mechanical effect mentioned in the main text.

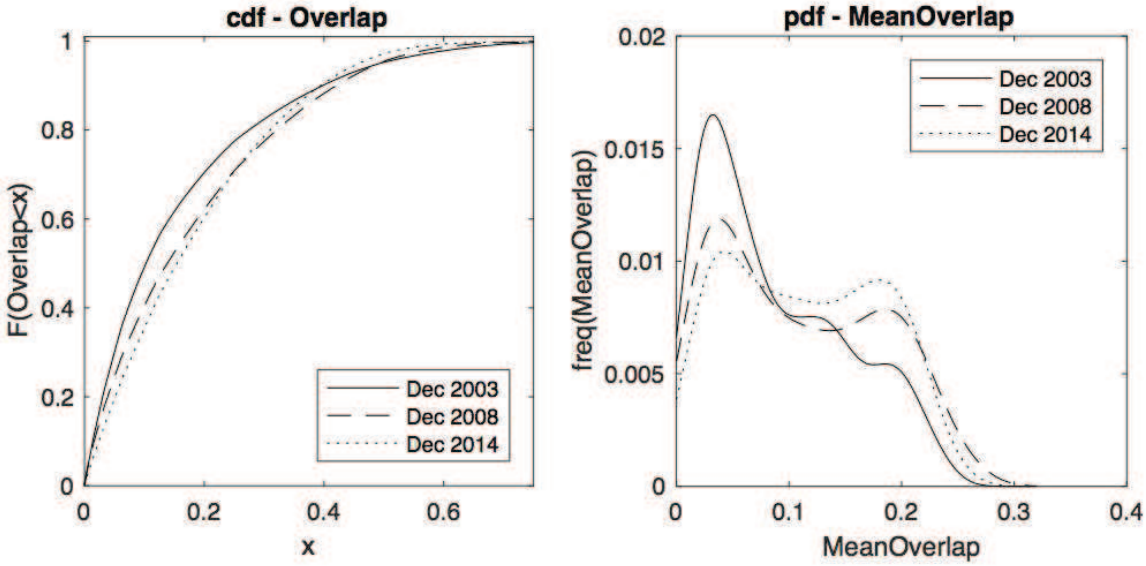

Figure 2. Empirical Distribution of Portfolio Overlap
Left panel: cumulative distributions of fund-pair-specific Overlap values based on end-of-year data in 2003, 2008, and 2014, respectively. For each date, I plot the full distribution of pairwise Overlap (as in Equation (2)). Right panel: probability distribution of fund-specific MeanOverlap based on the same data.

the importance of portfolio overlap, to the best of my knowledge there is little guidance on how such a significance analysis should be done. Before explaining my approach in detail, I therefore provide some background.

1. Background

As discussed in the literature review, numerous portfolio optimization models allow to compute an optimal portfolio, given an objective function (including certain constraints), the relevant model inputs (e.g., expected returns and covariances), and some free parameters (e.g., risk tolerance in a mean-variance framework). However, it is not clear how one should bring such models to the data and rigorously take investor heterogeneity into account. This is particularly relevant for large cross-sections (as in this paper), and I therefore refrain from using such approaches here.

Alternatively, theoretical (general) equilibrium models provide predictions on how investors' portfolios should look like in certain settings. For example, given a number of strong assumptions, the CAPM predicts that all investors should hold the market portfolio, thus implying perfect portfolio overlap across all investors. Clearly this is not a reasonable benchmark to compare my results with. Other models imply lower levels of portfolio overlap but make only qualitative predictions and thus cannot be used as quantitative theoretical benchmarks (e.g. Wagner, 2011).

Therefore, I use an alternative approach that is nonparametric, model-free, and does not rely on specific distributional assumptions. The idea is to fix several basic characteristics of observed fund portfolios and then perform a constrained randomization of these portfolios.[18] This allows me to quantify how the cross-section of mutual fund portfolios would look like on average under the null (matrix $\mathbf{W}_t^{\text{Null}}$), where funds' preferences for certain stocks are disregarded. In other words, comparing the actual values of portfolio overlap and those of the synthetic data gives an indication of whether reasonably simple probabilistic models can match the data. In the following, I define several of these benchmark models and explain their economic intuition.

2. Benchmarks

The benchmark (or null) models take mutual funds' observed portfolios as given and randomize these in different ways. The two basic models are as follows:

(1) "Shuffle‘': for each fund, I shuffle the portfolio weights among those stocks that are actually held by the fund. This approach keeps the observed heterogeneity in investor's portfolio weights but randomly reassigns these weights.

(2) "Randomize‘': for each fund, I shuffle all portfolio weights at random, disregarding the identities of stocks actually held by the fund. This approach removes any preferences of investors for certain stocks, such that any portfolio overlap arises purely by chance given the distribution of portfolio weights.

Note that both benchmarks take the observed distribution of portfolio weights for each fund as given, which makes most sense for the set of actively managed funds under study here. In other words, by fixing this distribution, I acknowledge the fact that funds choose to hold a certain number of stocks in their portfolios and desire a certain level of diversification. Such a procedure facilitates the comparison of the results.[19] Also note that, relative to the Randomize model, the Shuffle model is more constrained in the sense that it further incorporates a fund's preferences for holding specific stocks in its portfolio. As such, it is economically more relevant and should yield portfolios closer to the observed ones. An illustrative example of these benchmarks—with two mutual funds and two assets—is shown in Table III.

Note that the two benchmark models are relatively simplistic in that they ignore additional constraints that could drive the portfolio decisions of actively managed mutual funds. For example, it seems reasonable that these funds choose a certain industry exposure in their portfolio construction (e.g., Kacperczyk et al.,2005). Therefore, I also include

[18] The idea of randomizing investment portfolios is of course not novel. However, the literature has mainly used this approach for performance evaluation (e.g., Lisi, 2011).

[19] I also experimented with benchmark models that fix the number of stocks held by each fund but not their diversification levels. In this case, the interpretation of the results becomes much harder because differences in overlap can be driven by the numerator and the denominator of Equation (2).

Panel A. Original Portfolios

		Portfolio Weights					
		1	2	3	4	5	HHI
Fund	A	0.3	0.1	0.3	0.2	0.1	0.24
	B	0.6	0.4	0	0	0	0.52

$$\text{Overlap} = \frac{0.220}{\sqrt{0.240}\times\sqrt{0.520}} \approx 0.623$$

Panel B. Benchmarks

(1) "Shuffle"					(2) "Randomize"				
1	2	3	4	5	1	2	3	4	5
0.2	0.1	0.3	0.3	0.1	0.3	0.3	0.1	0.2	0.1
0.4	0.6	0	0	0	0	0	0.6	0	0.4
Overlap ≈ 0.396					Overlap ≈ 0.283				

Table III. Illustration of Portfolio Overlap in the Benchmark Models
Panel A shows an example of two funds' hypothetical portfolios, their HHIs and the corresponding portfolio overlap, as in Equation (2). Panel B shows examples of a hypothetical portfolio in the three benchmarks, and the corresponding portfolio overlap. Here, I show only a single realization.

more constrained versions of the above models that also fix the industry exposures of the funds:

(1b) "Shuffle—Industry10": same as "Shuffle" but including the additional constraint that a given fund's industry exposure matches what is observed in the data.

(2b) "Randomize—Industry10": same as "Randomized" but including additional constraint that a given fund's industry exposure matches what is observed in the data.

For example, in Model (2b), the observed portfolio weights for each industry will be shuffled randomly to stocks *within* that industry but ignoring the identity of stocks actually held by the fund. Hence, incorporating these additional constraints should yield portfolios closer to the observed ones. In both cases, I use Ken French's 10-industry classification. (I confirm qualitatively similar results for the 48-industry classification.)

3. Results

For each month and each of the four models, I generated 1,000 synthetic portfolio matrices ($\mathbf{W}_t^{\text{Null}}$), compute the corresponding Overlap-matrix, and then calculate the typical values of portfolio overlap (using Equation (5)) across replications for a given month. I then compare the levels of portfolio overlap with those observed in the data.

Figure 3 shows the results of this exercise. For each model, I show the distribution of relative deviations across all months between the portfolio overlap in the benchmarks and the real data $\left(\frac{\langle \text{MeanOverlap}^{\text{Null}}\rangle}{\langle \text{MeanOverlap}\rangle} - 1\right)$. For comparability, the solid black line indicates zero

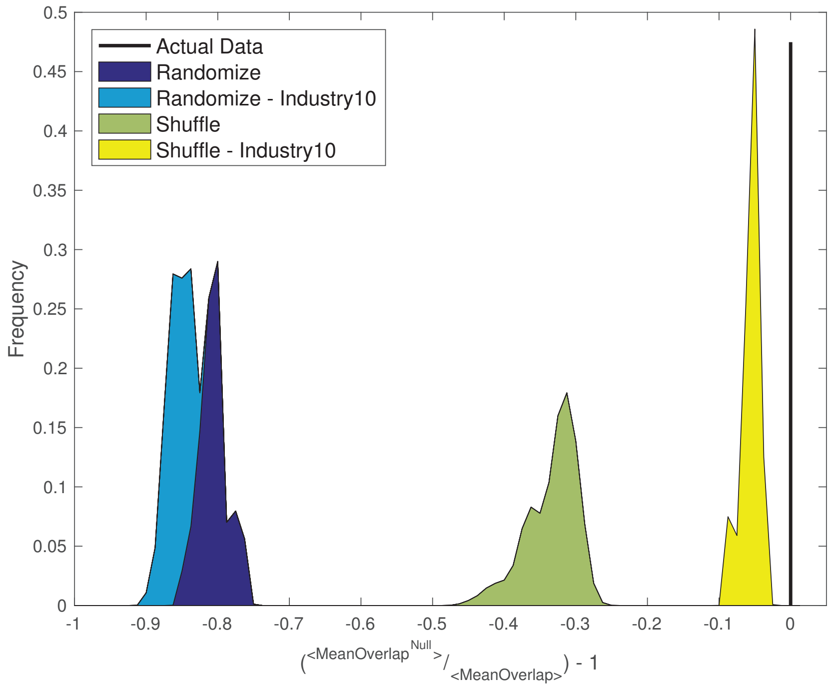

Figure 3. On the Significance of Portfolio Overlap
For each month and model, I generate 1,000 synthetic portfolio matrices ($\mathbf{W}^{\text{Null}}$), compute the corresponding Overlap-matrix, and then calculate the typical values of portfolio overlap (using Equation (5)) across replications. The plot shows the distribution of the relative deviations across all months between the benchmark models and the actual values, $\left(\frac{\langle \text{MeanOverlap}^{\text{Null}} \rangle}{\langle \text{MeanOverlap} \rangle} - 1\right)$. For the sake of readability, I show the actual values as a vertical black line with zero deviations. The benchmark models are defined in the main text.

deviations (actual data). In all cases, I observe negative values, such that all benchmark models predict significantly lower values of portfolio overlap.[20] For example, the first model ("Shuffle", shown in green) yields a portfolio overlap that tends to be around 35% smaller than what is observed in the data. In this case, taking the industry exposure into account ("Shuffle—Industry 10", yellow) yields values that are closer to the data but still significantly smaller (deviations of around 5%). Not surprisingly, the second model ("Randomize", dark blue) yields even lower values with deviations around 80%. Interestingly, in this case the constrained model ("Randomize—Industry 10", light blue) yields overlap values that are *lower* than those in the unconstrained model.[21]

Overall I find that relative to the benchmarks under study here, which disregard mutual funds' preferences for certain stocks, the observed portfolio overlap appears to be significant. While I acknowledge that these results are only indicative, it seems reasonable that the observed levels of portfolio overlap could not have arisen purely by chance.

[20] In fact, defining the corresponding p-values for each model as $p\text{-value}^{\text{Null}} = \#(\langle \text{MeanOverlap}_{\text{Null}} \rangle > \langle \text{MeanOverlap} \rangle)$, yields p-values of zero in all instances.

[21] This result is driven by fund pairs with no common industry exposure. In this case, the unconstrained reshuffling tends to lead to *higher* levels of portfolio overlap relative to the constrained case.

D. Portfolio Overlap and Return Correlations

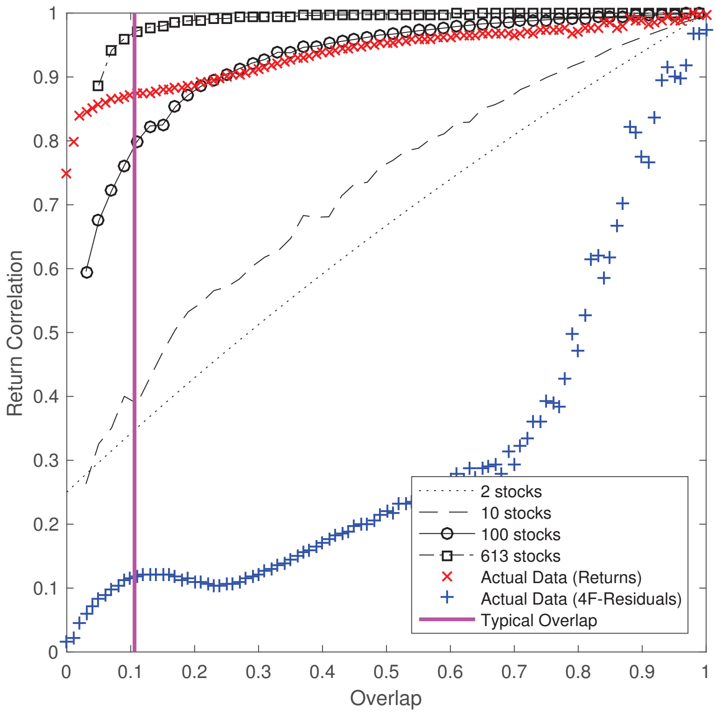

Figure 4. Portfolio Overlap and Return Correlations
The red crosses show the typical return correlations between mutual funds with varying levels of portfolio overlap (using only mutual fund pairs with at least 24 months of joint observations). In blue, I show the correlations between residuals based on the Carhart (1997) four-factor model. The vertical purple line shows the typical value of portfolio overlap over the full sample (from Table I). In order to construct the black lines, I take daily stock market returns from CRSP for the period under study (March 2003—December 2014). I then generate a large number of random portfolios with varying levels of portfolio overlap (as defined in Equation (2)) and calculate the return correlation of these synthetic portfolios. Results are shown for 2, 10, 100, and 613 stocks.

Before exploring the implications of portfolio overlap at the fund level, I wish to get a better understanding of the overall relation between portfolio overlap and performance diversity. For this purpose, for each fund-pair with at least 24 months of joint observations, I calculate the average portfolio overlap and the correlation between the two funds' portfolio returns. Figure 4 shows the typical relationship between the observed levels of portfolio overlap and the corresponding return correlations (in red). Note that I have averaged the data using different Overlap bins. The results are such that even for relatively low levels of portfolio overlap, the typical return correlation is very high (generally above 0.8). For the sake of reference, the purple vertical line shows the typical portfolio overlap of 0.106 observed during the sample period (cf. Table I). Hence, modest levels of portfolio overlap are in line with rather homogeneous performances.

To get a better understanding of this (empirical) relation, I take daily stock market returns from CRSP for the period under study (March 2003—December 2014) and construct a large number of random portfolios with varying levels of portfolio overlap (as defined in Equation (2)) and calculate the resulting return correlation of these random portfolios. I perform the exercise for a varying number of stocks, namely 2, 10, 100, and 613 stocks.[22] The results are shown as black lines in Figure 4. Clearly, there is a positive relation between portfolio overlap and return correlation in all cases. The relation becomes highly nonlinear when the number of stocks increases, with very strong correlations even for low levels of portfolio overlap: for portfolios consisting of 2 stocks, the implied return correlation would be around 0.3 for the observed values of portfolio overlap; for portfolios consisting of 100 stocks, the correlations are closer to 0.8; for portfolios with even more stocks the performances will be almost perfectly correlated.[23]

It is well known that a large part of the empirically documented return correlation can be captured by a small set of common factors (Carhart, 1997). Hence, the factor structure of asset returns translates modest levels of portfolio overlap into homogeneous performances. I therefore take a closer look at the factor structure of fund returns: the blue crosses show the correlation between four-factor residuals and portfolio overlap.[24] Clearly, the correlations are much smaller in this case suggesting that the relation is indeed driven by common factors of fund/stock returns. The hump for Overlap values between 0.1 and 0.2 is due to the bimodal distribution of MeanOverlap shown in the right panel of Figure 2.

E. On the Characteristics of Specialist Funds

I now turn my focus to the fund-level implications of portfolio overlap. In particular, as discussed in the literature review, most of the theoretical literature suggests that funds with superior information about certain assets should concentrate their investments on these. In my case, funds with low levels of portfolio overlap with other funds are special in the sense that their portfolios holdings are very different from those of all other funds. An obvious question is whether these funds, for example, tend to outperform other funds (hence the title of this paper).

I therefore proceed as follows: first, I identify specialist funds based on their observed levels of portfolio overlap with other funds. Second, I gather fund- and portfolio-specific characteristics over time. Last, I compare whether specialist funds differ from other funds across these characteristics. Let me briefly go through each of these steps.

[22] For simplicity, I restrict ourselves to 613 stocks that have complete data over the sample period.

[23] This result follows from the reduction of idiosyncratic risk due to diversification in portfolio theory: given a large enough number of stocks, any pair of portfolios will be strongly correlated because idiosyncratic risks get diversified away.

[24] The Carhart (1997) four-factor model includes (1) the excess return of the market portfolio over the risk-free rate, (2) the return difference between small and large market capitalization stocks, (3) the return difference between high and low book-to-market ratio (B/M) stocks, and (4) the return difference between stocks with high and low past returns. The factors were downloaded from Ken French's data library (`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library`).

Decile t \ t+1	MeanOverlap (Low) 1	2	3	4	5	6	7	8	9	(High) 10
1	0.957	0.041	0.002	.	.	.	.	.	.	.
2	0.036	0.910	0.051	0.002	.	.	.	.	.	.
3	.	0.045	0.894	0.058	0.002	.	.	.	.	.
4	.	.	0.053	0.884	0.061	0.002	.	.	.	.
5	.	.	.	0.055	0.879	0.063	0.001	.	.	.
6	.	.	.	.	0.058	0.878	0.062	0.002	.	.
7	.	.	.	.	.	0.057	0.879	0.060	0.003	.
8	.	.	.	.	.	.	0.057	0.876	0.065	0.001
9	.	.	.	.	.	.	.	0.059	0.885	0.055
10	.	.	.	.	.	.	.	.	0.049	0.950

Table IV. Quarterly Transition Matrix Between MeanOverlap Deciles
Each cell shows the probability of a mutual fund moving from decile *a* to decile *b*. Data are for the period March 2003—December 2014. Note: values below 0.1% are shown as ".".

1. Step 1: Sorting Funds into MeanOverlap Deciles

Here, I sort funds into deciles based on their observed levels of portfolio overlap, MeanOverlap, as defined in Equation (4). At any point in time, decile 1 (10) corresponds to funds with the lowest (highest) levels of portfolio overlap. Funds in decile 1 hold portfolios that are very different from those of all other funds.

Table IV shows the quarterly transition matrix based on the deciles for MeanOverlap over the entire sample period. The table shows substantial persistence in the different deciles, in particular for the two extreme categories: for example, a fund that is in decile 1 (10) today has a 95.7% (95%) probability of being in the same decile in the next quarter.[25]

2. Step 2: Gathering Fund-/Portfolio-Specific Characteristics

The next step is to gather fund-/portfolio characteristics that might differ between funds with different levels of portfolio overlap.

Portfolio-Specific Characteristics. The first set of characteristics is meant to capture how popular a given stock is among mutual funds in general, and the choice of characteristics is mainly driven by the literature on institutional preferences (e.g., Bennett, Sias, and Starks, 2003). In the following, I use a measure that is commonly used in the abovementioned literature, namely *institutional ownership* (InstOwn). It is defined as the sum of

[25] I show the quarterly transition matrix because fund portfolios tend to be updated quarterly. As a robustness check, Table B.I shows that the transition matrices are comparable for longer transition periods (one year and two years, respectively). Moreover, Table C.I shows the quarterly transition matrix based on funds' diversification levels (measured as the negative HHI) is very stable as well.

Panel A	*Portfolio-Specific Characteristics*				
	Mean	Median	Std. Dev.	Min.	Max
Age^P	15.718	12.500	13.837	1.083	97.667
Amihud^P	0.026	0.000	0.668	0.000	153.146
B/M^P	0.478	0.451	0.165	0.093	3.191
FirmRisk^P	0.175	0.143	0.116	0.026	4.295
InstOwn^P	0.392	0.388	0.117	0.070	0.796
Mcap^P	38.260	35.690	32.587	0.069	261.077
OwnConc^P	0.265	0.246	0.075	0.148	0.923
Price^P	55.609	49.116	27.617	2.097	520.399
Turnover^P	0.010	0.009	0.003	0.001	0.207
Panel B	*Fund-Specific Characteristics*				
	Mean	Median	Std. Dev.	Min.	Max
Abs(Flows)	0.030	0.013	0.076	0.000	1.996
Flows	0.002	–0.004	0.082	–0.997	1.996
HHI	0.029	0.017	0.052	0.003	0.978
MeanOverlap	0.106	0.101	0.065	0.001	0.285
NumberStocks	106.046	90	67.367	3	300
Alpha	0.000	0.000	0.007	–0.126	0.165
Beta	0.977	0.985	0.280	–3.634	3.836
Abs(Return)	0.038	0.029	0.034	0.000	0.829
Return	0.009	0.014	0.049	–0.462	0.829
Return-Style	0.000	0.000	0.017	- –0.439	0.913
Sharpe-Ratio	0.287	0.306	0.358	-1.602	2.565
TNA	1.216	0.224	5.066	0.001	202.30

Table V. Summary Statistics for Fund-/Portfolio-Specific Characteristics
Based on CRSP-Compustat data for the period March 2003—December 2014. I show the mean, median, minimum, maximum, and standard deviation for each variable. These are defined as follows: Age^P is the portfolio-weighted average age in months of the stocks held by a given fund. $Amihud^P$, B/M^P, $FirmRisk^P$, $InstOwn^P$, $Mcap^P$, $OwnConc^P$, $Price^P$, and $Turnover^P$ are the portfolio-weighted average of the corresponding stock-specific characteristics as defined in the main text. *Flows* are the raw monthly percentage inflows, and *Abs(Flows)* is the corresponding absolute value. *HHI* is the Hirschmann-Herfindahl Concentration Index based on investors' portfolio weights. *MeanOverlap* is the main measure of portfolio overlap as defined in Equation (4). *NumberStocks* is the number of stocks held by a fund. *Return* is the monthly fund-specific return, and *Return-Style* are fund-style-adjusted returns based on CRSP objective codes. *Alpha* and *Beta* are the intercept and the loading on the market factor from Carhart (1997) four-factor regressions, respectively, based on 12-month rolling windows. *Sharpe-Ratio* is the corresponding average return divided by the standard deviation. *TNA* are the total net assets. Note: Amihud is shown per US$ 1 million trading volumes; Mcap and TNA are shown in US$ billion.

shares held by mutual funds relative to the total number of shares outstanding,

$$\text{InstOwn}_{k,t} = \frac{\sum_i s_{i,k,t}}{\text{ShareOut}_{k,t}}, \tag{6}$$

which takes values between zero and one, where higher values correspond to more popular stocks.

I also define a measure of *ownership concentration* of stock k at date t ($\text{OwnConc}_{k,t}$) as the fraction of shares held by the five most important funds relative to the total fund holdings of that particular stock.

The remaining stock-specific characteristics are as follows: *Age* is the number of months since the stock first appeared in either CRSP or Compustat. *Amihud* is the monthly average value of the daily Amihud ratio (defined as the absolute return over the dollar trading volume). *Book-to-market value* (B/M) is the book value of equity relative to its market value at the end of any given month (taken from Compustat). *Firm-specific risk* (FirmRisk) is the idiosyncratic risk of a stock relative to its main industry. It is defined as the monthly sum of the squared differences between a stock's daily returns and its corresponding industry returns. *Mcap* is the stock's total market capitalization. *Price* is the stock price. *Turnover* is the average fraction of a stock's daily trading volume (in shares) relative to the total number of shares outstanding.

The next step is to aggregate the aforementioned stock-specific characteristics to the portfolio level (superscript P). I proceed as usual: observing a certain characteristic, say $z_{k,t}$, for each stock at a given point in time, a given fund i's portfolio-weighted average of that characteristic is

$$z^P_{i,t} = \sum_k^K w_{i,k,t} z_{k,t}. \tag{7}$$

The variables Age^P, Amihud^P, B/M^P, FirmRisk^P, InstOwn^P, Mcap^P, OwnConc^P, Price^P, and Turnover^P are constructed in this way.

Fund-Specific Characteristics. The CRSP Mutual Fund database contains characteristics such as fund size (*TNA*) and fund returns (*Return*).[26] I also use future returns, *Return(t+3)*, which are defined as funds' realized returns three months in the future. While raw fund returns are often used as a performance measure, a huge literature is devoted to capturing *excess returns*. In line with this literature, I use the following measures: first, I construct style-adjusted returns (*Return-Style*) that are in excess of a given fund's peers with the same CRSP objective codes. Second, I calculate fund alphas (*Alpha*) and betas (*Beta*) based on the Carhart (1997) four-factor model. Third, I include the *Sharpe-Ratio* as another measure of risk-adjusted returns. In the latter two cases, I perform rolling window estimations over the previous 12 months. Furthermore, I also include a fund's absolute

[26] CRSP mutual fund returns are net of fees, which are most relevant from an investor's perspective.

return, *Abs(Return)*, as a measure of return volatility.

Following the literature, I also calculate a fund's percentage (net) flows as

$$\text{Flows}_{i,t} = \frac{\text{TNA}_{i,t} - \text{TNA}_{i,t-1}(1 + \text{Return}_{i,t})}{\text{TNA}_{i,t-1}}. \tag{8}$$

Finally, in addition to the raw fund flows, I use Abs(Flows) as a measure of flow volatility. Table V shows summary statistics for each of the characteristics. Table E.I in the Appendix shows the correlation matrix between these characteristics.

3. Testing for Significant Differences

My sample covers a relatively long period (March 2003—December 2014). Following the literature (e.g., Antón and Polk, 2014), I therefore standardize each characteristic separately for each month, meaning that each cross-section has zero mean and unit standard deviation for each portfolio characteristic. The main advantages of this procedure are twofold: first, some of the characteristics under study here are likely to be driven by time trends (e.g., Amihud ratio and market capitalization), which can affect the comparison of different cross-sections along these dimensions. Second, it facilitates the readability of the output tables below in the sense that positive (negative) values indicate that a given decile differs relative to the typical values observed over time.[27]

In everything that follows, I use the MeanOverlap values from the previous month to sort funds into deciles and then calculate the average for a given portfolio characteristic in the next month for each decile.[28] This yields 141 monthly observations for each decile and characteristic. Here, I test whether there are significant differences between different deciles for each of the characteristics. In this regard, the main question is whether funds with low levels of portfolio overlap are significantly different from those with high levels of overlap. To answer this question, I directly compare deciles 1 and 10, and below-median (deciles 1 to 5) and above-median (deciles 6 to 10) diversification levels. Significance is assessed based on Wilcoxon rank-sum tests.[29]

I should stress that given that I perform multiple comparisons across the same groups, significance levels need to be adjusted in order to avoid rejecting the null hypothesis (no significant differences between deciles) too frequently.[30] For simplicity, I resort to Bonferroni-corrected p-values here. I compare fund deciles based on $n = 20$ different fund-/portfolio

[27] The characteristics in levels can be roughly retrieved by multiplying them by the standard deviation shown in Table V and adding the corresponding mean values. I checked that the results are robust to using all characteristics in levels; see Tables VI (standardized) and F.I (raw).

[28] This is akin to constructing equal-weighted portfolios of funds within each decile. I tested that value-weighted averages (based on fund TNAs) yield results comparable to those presented in the following.

[29] Standard t-tests yielded qualitatively similar results.

[30] Harvey, Liu, and Zhu (2016) provide details on the multiple testing problem. Barras et al. (2010) propose a simple multiple testing methodology that separates mutual funds into skilled, unskilled, and zero Alpha funds based on the estimated Alphas and their p-values.

characteristics. Hence, in order to achieve a nominal significance level of 5%, the p-value has to be below $0.05/n = 0.0025$.

Last, in order to mitigate the impact of extreme observations, I log-transform some of the portfolio characteristics before standardization: namely, I take the logarithm of Age^P, Amihud^P, Mcap^P, Price^P, Turnover^P, and TNA. If anything, this procedure will make it harder to detect significant differences between deciles in the significance tests.

Main Results. The main results are shown in Table VI. Each row in this table corresponds to a particular fund/portfolio-characteristic, and the first 10 columns show the standardized time average for each decile-characteristic combination. Note that due to the standardization procedure, the average value across the first 10 columns equals zero, such that a positive (negative) value indicates that a given decile tends to be above (below) the average mutual fund in terms of a given characteristic.[31] The last two columns show the average difference between deciles 1 and 10, and above-median (deciles 1 to 5) and below-median (deciles 6 to 10), respectively, including the Bonferroni-corrected significance levels.

The results can be summarized as follows: I find that specialist funds' portfolios (those with low levels of portfolio overlap) are significantly different from others with regard to most of the characteristics under study here. For example, panel A of Table VI shows that these funds tend to hold portfolios consisting of relatively small stocks (Mcap), which are relatively young (Age), illiquid (Amihud), and with higher firm-specific risk (FirmRisk). The stocks held by specialists also have a higher share of institutional ownership (InstOwn) and the ownership tends to be more concentrated (OwnConc). Panel B shows that funds with low levels of portfolio overlap tend to be relatively small (TNA), hold concentrated portfolios and receive larger inflows (Flows) with higher volatility.[32] In practically all of these cases, I find that the differences are statistically significant at the 1% level.

However, when it comes to the performance measures (Alpha, Return, Return-Style, and Sharpe-Ratio), the differences between deciles 1 and 10 are generally positive but mostly insignificant. This is despite the fact that specialists have significantly lower Betas.[33] When comparing below-/above-median funds (last column), I do find significant differences

[31] This is important to keep in mind, because some of the characteristics under study here—such as Abs(Return)—can take only positive values. As a robustness check, I also produced the same table based on the raw variables (nonstandardized); see Table F.I in the Appendix. In terms of the significance tests, the results are qualitatively comparable to the ones shown here.

[32] In Table H.I in the Appendix, I run standard flow-performance regressions, where I model fund flows as a function of past performance (Alpha(t–1)) and past flows (Flows(t–1)). In line with the results shown in Table VI, I find that specialist funds have highly persistent flows. Interestingly, however, these funds' flows react less strongly to past performance (parameter on Alpha is around 0.3, compared with values between 0.8 and 1.2 for the other deciles).

[33] In addition, I also tested for significant differences in the other factor loadings from the four-factor model (unreported results). In line with the results shown in panel A of Table VI, I find that funds with low levels of overlap have significantly higher loadings on both the *SMB* factor (as one would expect from the results on Age and Mcap) and the *HML* factor (as for B/M). The loadings on the *momentum* factor are not significantly different from those of other funds.

for fund Alphas but upon closer inspection it becomes clear that these results are mainly driven by the superior performance of deciles 2 to 5. The results for all other performance measures remain insignificant. Interestingly, I also find that specialists display more volatile returns. This is not surprising given that they tend to hold more concentrated portfolios (HHI significantly larger as well).

To get a better understanding of these findings, I also performed the nonparametric monotonicity-tests of Patton and Timmermann (2010). This methodology allows to test for a strictly decreasing pattern in the performances of MeanOverlap deciles versus the null hypothesis of a flat or weakly increasing pattern. Given that the results regarding fund Alphas in Table VI may be driven by the fact that MeanOverlap is not independent of fund size, I perform a double-sorting here. Specifically, I sort on MeanOverlap first and then on fund TNA. In line with the results presented in Table VI, I find that neither of the performance measures under study here exhibits a monotonic pattern at reasonable significance levels. As an illustration, Table VII shows the results for raw (nonstandardized) Alphas using both single- and double-sorted portfolios when sorting first in terms of MeanOverlap (rows) and/or then on TNA (columns). Note that in this case, the p-values are not Bonferroni-corrected but generally too large to reject the null hypothesis. This suggests that the results on the performance measures in Table VI are not overly driven by fund size.

Full Sample	Decile MeanOverlap(t–1)											(1:5) - (6:10)
	1 (Low)	2	3	4	5	6	7	8	9	10 (High)	1-10	
Panel A												
Age^P	–0.726	–0.711	–0.616	–0.384	–0.246	0.073	0.387	0.670	0.756	0.794	–1.520**	–1.073**
$Amihud^P$	1.523	0.937	0.545	0.266	–0.017	–0.255	–0.537	–0.764	–0.823	–0.869	2.392**	1.300**
B/M^P	0.756	0.353	0.187	0.063	–0.155	–0.246	–0.257	–0.205	–0.217	–0.276	1.032**	0.481**
$FirmRisk^P$	1.064	0.551	0.409	0.134	0.015	–0.163	–0.340	–0.526	–0.566	–0.574	1.638**	0.868**
$InstOwn^P$	0.281	0.874	0.777	0.485	0.351	–0.072	–0.436	–0.681	–0.758	–0.820	1.101**	1.107**
$Mcap^P$	–1.526	–1.227	–0.789	–0.269	0.056	0.426	0.695	0.826	0.879	0.921	–2.447**	–1.501**
$OwnConc^P$	1.889	0.830	0.388	0.036	–0.220	–0.358	–0.523	–0.623	–0.673	–0.739	2.628**	1.168**
$Price^P$	–1.287	–0.824	–0.439	–0.073	0.205	0.328	0.433	0.531	0.541	0.580	–1.868**	–0.966**
$Turnover^P$	0.113	0.362	0.481	0.391	0.370	0.002	–0.217	–0.483	–0.518	–0.501	0.614**	0.687**
Panel B												
Abs(Flows)	0.094	0.060	0.044	0.053	0.041	0.005	–0.026	–0.056	–0.090	–0.126	0.221**	0.117**
Flows	0.079	0.055	0.002	0.035	0.000	0.016	0.006	–0.031	–0.067	–0.094	0.173**	0.068**
HHI	1.053	0.045	–0.065	–0.040	–0.058	–0.040	–0.105	–0.188	–0.265	–0.333	1.386**	0.373**
Alpha	0.039	0.079	0.073	0.113	0.043	0.017	–0.045	–0.107	–0.115	–0.097	0.136	0.139**
Beta	–0.256	–0.071	–0.027	–0.026	0.022	0.004	0.064	0.072	0.105	0.112	–0.368**	–0.143**
Abs(Return)	0.430	0.221	0.145	0.080	0.048	–0.076	–0.146	–0.217	–0.223	–0.261	0.691**	0.369**
Return	0.052	0.076	0.056	0.037	0.028	–0.019	–0.054	–0.053	–0.052	–0.070	0.121	0.099
Return(t+3)	0.030	0.028	0.046	0.027	0.025	–0.006	–0.043	–0.031	–0.034	–0.043	0.073	0.063
Return-Style	–0.036	0.005	0.016	0.011	0.010	0.004	–0.013	0.006	0.003	–0.006	–0.029	0.003
Sharpe-Ratio	–0.119	–0.025	0.024	0.082	0.008	0.016	–0.034	0.025	0.037	–0.013	–0.106	–0.012
TNA	–0.275	–0.029	–0.071	–0.101	–0.074	–0.090	–0.049	0.086	0.273	0.328	–0.603**	–0.220**

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table VI. Are Specialist Funds Special?

This table shows the average standardized portfolio-/fund-characteristics for MeanOverlap deciles (based on first lag). Data are for March 2003—December 2014. For each month, I standardize the corresponding variables (rows) and calculate the mean value for each characteristic-decile combination. Columns (1) to (10) show the corresponding time-averages. The last two columns explicitly compare funds with low and high portfolio overlap (deciles 1 and 10), and those with below- and above-median values (deciles 1–5 versus 6–10), respectively. Significance tests are based on Wilcoxon rank-sum tests with Bonferroni-corrected p-values.

Alpha (4F)		Decile TNA (Small) 1	2	3	4	5	6	7	8	9	(Large) 10	MR p-value	Joint MR p-value
(Low)	1	–0.063	–0.042	0.052	–0.165	0.025	–0.050	–0.051	0.096	0.118	0.307	0.990	
	2	–0.045	–0.021	0.069	0.043	–0.065	0.101	0.105	0.157	0.162	0.226	0.794	
	3	–0.104	0.026	0.070	0.049	0.101	0.048	0.055	0.096	0.078	0.246	0.468	
Decile	4	–0.103	0.019	0.058	0.012	0.171	0.178	0.145	0.139	0.167	0.253	0.103	
MeanOverlap	5	–0.050	–0.053	0.050	–0.039	0.017	0.079	0.179	0.115	0.026	0.167	0.912	0.905
	6	–0.096	–0.051	–0.071	0.035	0.006	0.006	0.069	0.060	0.103	0.087	0.044	
	7	–0.136	–0.100	–0.064	–0.055	–0.025	–0.004	0.011	–0.033	–0.077	0.052	0.834	
	8	–0.126	–0.153	–0.124	–0.151	–0.056	–0.043	–0.090	–0.121	–0.076	–0.005	0.676	
	9	–0.105	–0.098	–0.160	–0.123	–0.123	–0.110	–0.068	–0.043	–0.083	–0.051	0.891	
(High)	10	–0.026	–0.099	–0.119	–0.144	–0.108	–0.108	–0.055	–0.102	–0.079	–0.088	0.823	
	MR p-value	0.392	0.707	0.527	0.998	0.963	0.934	0.791	0.927	0.784	0.005		
	Joint MR p-value					0.889							

Table VII. Monotonicity Tests Based on the Approach of Patton and Timmerman (2010).
The null hypothesis is that there is a flat or weakly increasing pattern (versus a monotonically decreasing one). The table reports the typical fund Alphas after sorting first on MeanOverlap and then on TNA. MR p-values show the p-values for each decile, while fixing the other characteristic. The joint MR p-values explore each sorting feature separately. (Note: I used the Matlab implementation as posted at Andrew Patton's website.)

Robustness Checks. I performed various additional robustness checks which generally yielded similar qualitative results. Let me briefly discuss the most important ones.

- **Excluding Financial Crisis Years.** It is well-known that stocks (and thus performances) comove more strongly during crisis periods (see Pollet and Wilson, 2010). I therefore checked whether the results are driven by the global financial crisis. For this purpose, Table VIII is similar to Table VI but excludes data from the crisis years (2008 and 2009). The results are generally very similar, and I find no significant differences in specialist funds' performances.

- **MeanBinOverlap Deciles (Binary Measure of Portfolio Overlap).** In earlier sections of this paper, I also analyzed the binary measure of portfolio overlap and found that the time dynamics are similar to my main measure of interest—namely the cosine similarity measure. The raw Pearson-correlation between MeanOverlap and MeanBinOverlap is 0.921, so the results should be very similar. I therefore also constructed deciles based on MeanBinOverlap and ran the same significance tests. Table IX shows that the results are very similar, both qualitatively and quantitatively. The only difference in this case is that specialists' Alphas are significantly larger, but the results remain insignificant for the other performance measures.

- **MeanOverlap Deciles Based on 48-Industry Portfolios.** My measure of portfolio overlap is based on highly granular portfolios. However, it might still be the case that funds with low levels of portfolio overlap cluster their investments on a certain set of industries as proposed by Kacperczyk et al. (2005); see also Figure A.1 in the Appendix. Hence, I also calculated funds' MeanOverlap based on 48-industry portfolios and then constructed deciles based on this alternative measure. Table X shows that most of the results are robust to this alternative classification, with signs and significance levels comparable to those reported in Table VI.

- **Deciles Based on Diversification Levels (–HHI).** Previous work has explored whether concentrated mutual funds tend to outperform more diversified ones (see e.g., Kacperczyk et al., 2005, 2007). Given that my measure of portfolio overlap (MeanOverlap) is related to fund diversification levels (negative HHI, or –HHI), an obvious question is whether I obtain similar results when sorting funds into diversification deciles instead.[34] Table XI shows the equivalent of Table VI, providing comparable results in terms of the statistical (in-)significance of the performance measures under study here.[35] Interestingly, concentrated funds tend to hold *larger*,

[34] Table C.I in the Appendix shows that this yields very persistent classifications as well.

[35] I should note that Kacperczyk et al. (2005, 2007) found that concentrated mutual funds outperform more diversified ones, while I find no significant differences here. These opposing findings can be driven by various factors. First, I should stress that the data sets are constructed in a similar fashion but cover different sample periods. Kacperczyk et al. use data only up until 2003, whereas I have access

older, and *more liquid* stocks in their portfolios, while the results were exactly the opposite based on the MeanOverlap classification. This suggests that portfolio overlap and portfolio concentration indeed measure something different and tend to yield different classifications.[36]

- **Fund-Level Regressions.** Constructing deciles and comparing these in terms of their performances is a standard approach in finance. An alternative approach is to explore whether funds from different deciles differ in their performances by running fund-level regressions and controlling for (unobserved) fund-level heterogeneity. This is done in Table G.I in the Appendix, where I wish to explain fund Alpha's as a function of their corresponding MeanOverlap decile. Here, I include dummy variables for each decile, with decile 1 serving as the baseline category. Given that Alphas are estimated using 12-month rolling windows, I use MeanOverlap(t–12)—that is, the 12-month lag. I also control for lagged flows and (log-)size and include fund-fixed effects in all specifications. Standard errors are clustered by month, and all variables (except for the dummy variables) are standardized. I run regressions for current Alphas but also for values 6- and 12-months in the future. In line with decreasing returns to scale, I find that larger funds perform significantly worse. When it comes to MeanOverlap deciles, however, most of the parameters are insignificant. If anything, specialist funds tend to marginally outperform funds from other categories. This outperformance, however, does not survive longer time windows, because the corresponding coefficients are generally insignificant for Columns (2) and (3).

to data only from 2003 onwards. Therefore, my findings suggest that there have been changes in the relative performances of mutual funds with different diversification levels. This seems reasonable given the increasing levels of competition in the mutual fund industry and the dramatic growth of the system over the last two decades, which may drive some of my findings. Second, the increased portfolio overlap over my sample period suggests that career concerns of mutual fund managers, and their relative performances, might also have become more important. Last, average stock return correlations have increased over time, which makes it more difficult for active managers to generate superior performances.

[36] While the unconditional correlation between MeanOverlap and HHI is –0.251, the resulting deciles are generally very different. Table D.I provides further evidence along those lines, where I show the probability of a fund being in MeanOverlap decile *a* and negative HHI decile *b*. If the classification were similar, I would expect very large diagonal elements, which is not the case here.

No Crisis (excl. 2008–09)	Decile MeanOverlap(t–1) 1 (Low)	2	3	4	5	6	7	8	9	10 (High)	1-10	(1:5) - (6:10)
Panel A												
Age^P	–0.756	–0.737	–0.624	–0.383	–0.249	0.076	0.406	0.690	0.779	0.793	–1.549**	–1.099**
$Amihud^P$	1.488	0.925	0.554	0.267	–0.010	–0.245	–0.527	–0.760	–0.826	–0.860	2.348**	1.288**
B/M^P	0.751	0.384	0.218	0.078	–0.154	–0.253	–0.254	–0.227	–0.241	–0.299	1.051**	0.510**
$FirmRisk^P$	1.089	0.597	0.411	0.134	0.015	–0.170	–0.362	–0.541	–0.579	–0.588	1.677**	0.897**
$InstOwn^P$	0.313	0.896	0.772	0.481	0.342	–0.071	–0.448	–0.691	–0.776	–0.815	1.128**	1.121**
$Mcap^P$	–1.517	–1.231	–0.786	–0.272	0.063	0.422	0.697	0.826	0.880	0.912	–2.429**	–1.496**
$OwnConc^P$	1.875	0.826	0.397	0.036	–0.221	–0.360	–0.518	–0.627	–0.671	–0.730	2.605**	1.164**
$Price^P$	–1.291	–0.832	–0.445	–0.085	0.197	0.322	0.432	0.538	0.559	0.600	–1.891**	–0.981**
$Turnover^P$	0.109	0.345	0.448	0.380	0.363	0.013	–0.216	–0.461	–0.505	–0.476	0.585**	0.658**
Panel B												
Abs(Flows)	0.089	0.059	0.046	0.050	0.045	0.006	–0.022	–0.056	–0.093	–0.123	0.212**	0.115**
Flows	0.071	0.047	0.009	0.031	0.006	0.016	0.004	–0.031	–0.061	–0.091	0.161**	0.065**
HHI	1.092	0.033	–0.058	–0.042	–0.065	–0.054	–0.113	–0.193	–0.264	–0.331	1.423**	0.383**
Alpha	0.022	0.073	0.067	0.103	0.048	0.015	–0.043	–0.094	–0.096	–0.093	0.115	0.125**
Beta	–0.177	–0.045	–0.026	–0.010	0.035	–0.005	0.043	0.037	0.070	0.078	–0.256**	–0.090**
Abs(Return)	0.458	0.228	0.136	0.077	0.042	–0.076	–0.145	–0.222	–0.232	–0.264	0.721**	0.376**
Return	0.070	0.070	0.048	0.032	0.028	–0.019	–0.047	–0.054	–0.057	–0.070	0.139	0.099 *
Return(t+3)	0.052	0.040	0.046	0.027	0.029	–0.011	–0.046	–0.040	–0.046	–0.051	0.104	0.078
Return-Style	–0.014	0.009	0.013	0.011	0.012	–0.002	–0.010	–0.001	–0.006	–0.013	–0.001	0.013
Sharpe-Ratio	–0.053	0.006	0.042	0.090	0.033	0.025	–0.045	–0.021	–0.015	–0.061	0.008	0.047
TNA	–0.276	–0.047	–0.071	–0.101	–0.066	–0.081	–0.033	0.075	0.275	0.323	–0.599**	–0.224**

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table VIII. Are Specialist Funds Special? Robustness Check Excluding the Financial Crisis
This table shows the average standardized portfolio-/fund-characteristics for MeanOverlap deciles (based on first lag). Data are for March 2003—December 2014, excluding the financial crisis years 2008 and 2009. For each month, I standardize the corresponding variables (rows) and calculate the mean value for each characteristic-decile combination. Columns (1) to (10) show the corresponding time-averages. The last two columns explicitly compare funds with low and high portfolio overlap (deciles 1 and 10), and those with below- and above-median values (deciles 1–5 versus 6–10), respectively. Significance tests are based on Wilcoxon rank-sum tests with Bonferroni-corrected p-values.

Full Sample	Decile MeanBinOverlap(t–1)											(1:5)
	1 (Low)	2	3	4	5	6	7	8	9	10 (High)	1-10	- (6:10)
Panel A												
Age^P	–0.903	–0.754	–0.501	–0.318	–0.224	0.111	0.553	0.648	0.725	0.612	–1.515**	–1.070**
Amihud^P	1.692	0.977	0.547	0.269	0.001	–0.287	–0.648	–0.836	–0.902	–0.740	2.432**	1.380**
B/M^P	0.710	0.294	0.153	0.125	–0.103	–0.154	–0.162	–0.255	–0.308	–0.278	0.988**	0.467**
FirmRisk^P	1.216	0.628	0.363	0.086	–0.008	–0.203	–0.468	–0.555	–0.534	–0.469	1.685**	0.903**
InstOwn^P	0.326	0.942	0.626	0.433	0.344	–0.042	–0.509	–0.697	–0.757	–0.626	0.952**	1.061**
Mcap^P	–1.727	–1.216	–0.593	–0.164	0.070	0.394	0.682	0.801	0.861	0.821	-2.548**	–1.438**
OwnConc^P	1.968	0.812	0.379	0.038	–0.200	–0.367	–0.537	–0.659	–0.705	–0.655	2.623**	1.184**
Price^P	–1.421	–0.820	–0.377	–0.024	0.185	0.266	0.423	0.530	0.575	0.611	-2.032**	–0.972**
Turnover^P	0.065	0.457	0.457	0.359	0.354	0.035	–0.408	–0.514	–0.489	–0.301	0.366**	0.674**
Panel B												
Abs(Flows)	0.104	0.056	0.133	0.052	–0.027	–0.024	–0.033	–0.079	–0.065	–0.114	0.218**	0.127**
Flows	0.075	–0.007	0.047	0.051	0.005	0.030	–0.009	–0.035	–0.091	–0.061	0.136**	0.067**
HHI	0.240	0.226	0.266	0.113	–0.006	–0.054	–0.082	–0.133	–0.217	–0.344	0.583**	0.334**
Alpha	0.054	0.046	0.151	0.100	0.042	–0.017	–0.092	–0.089	–0.107	–0.082	0.136**	0.156**
Beta	–0.132	–0.073	–0.122	–0.037	–0.012	0.059	0.047	0.066	0.082	0.111	–0.243**	–0.148**
Abs(Return)	0.311	0.307	0.206	0.109	0.023	–0.081	–0.189	–0.226	–0.228	–0.219	0.530**	0.380**
Return	0.064	0.072	0.066	0.030	0.005	–0.019	–0.046	–0.058	–0.071	–0.042	0.107	0.095
Return(t+3)	0.022	0.036	0.058	0.030	0.005	–0.013	–0.034	–0.037	–0.040	–0.028	0.050	0.060
Return-Style	–0.026	–0.009	0.012	0.017	0.011	0.006	–0.009	–0.009	–0.017	0.021	–0.047	0.003
Sharpe-Ratio	–0.064	–0.076	0.042	0.076	0.009	0.028	0.028	–0.001	–0.064	0.020	–0.085	–0.005
TNA	–0.291	–0.004	–0.142	–0.092	–0.075	–0.047	–0.079	0.096	0.208	0.407	–0.698**	–0.238**

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table IX. Are Specialist Funds Special? Robustness Check Using MeanBinOverlap

This table shows the average standardized portfolio-/fund-characteristics for MeanBinOverlap deciles (based on first lag). Data are for March 2003—December 2014. For each month, I standardize the corresponding variables (rows) and calculate the mean value for each characteristic-decile combination. Columns (1) to (10) show the corresponding time-averages. The last two columns explicitly compare funds with low and high portfolio overlap (deciles 1 and 10), and those with below- and above-median values (deciles 1–5 versus 6–10), respectively. Significance tests are based on Wilcoxon rank-sum tests with Bonferroni-corrected p-values.

Full sample	Decile MeanOverlap(t–1) **Using 48-Industry Portfolios**											(1:5)
	1 (Low)	2	3	4	5	6	7	8	9	10 (High)	1-10	- (6:10)
Panel A												
Age^P	0.166	–0.276	–0.177	–0.041	0.054	0.024	–0.013	–0.015	0.024	0.250	–0.084**	–0.109**
$Amihud^P$	0.432	0.433	0.329	0.167	–0.011	–0.116	–0.170	–0.258	–0.357	–0.454	0.885**	0.541**
B/M^P	0.442	0.198	0.107	0.189	0.105	–0.021	–0.172	–0.259	–0.332	–0.259	0.702**	0.417**
$FirmRisk^P$	0.088	0.275	0.189	0.018	–0.041	–0.032	–0.024	–0.070	–0.135	–0.266	0.354**	0.211**
$InstOwn^P$	–0.099	0.074	0.143	0.048	–0.027	–0.011	0.027	0.043	0.007	–0.204	0.105**	0.055**
$Mcap^P$	–0.013	–0.103	–0.262	–0.203	–0.070	–0.020	0.009	0.092	0.185	0.385	–0.398**	–0.260**
$OwnConc^P$	0.523	0.336	0.319	0.212	0.037	–0.082	–0.172	–0.294	–0.394	–0.490	1.013**	0.572**
$Price^P$	–0.236	–0.097	–0.273	–0.184	–0.096	–0.018	0.052	0.184	0.289	0.380	–0.616**	–0.355**
$Turnover^P$	0.043	0.049	0.006	–0.155	–0.136	–0.050	0.039	0.100	0.109	–0.006	0.048	–0.077**
Panel B												
Abs(Flows)	0.254	0.056	0.049	–0.001	–0.021	–0.039	–0.054	–0.064	–0.079	–0.104	0.358**	0.135**
Flows	0.069	0.042	0.072	0.052	0.014	–0.010	–0.022	–0.040	–0.080	–0.098	0.167**	0.100**
HHI	1.527	0.212	–0.009	–0.097	–0.168	–0.218	–0.261	–0.299	–0.333	–0.375	1.901**	0.590**
Alpha	0.146	0.124	0.028	–0.007	–0.031	–0.027	–0.045	–0.075	–0.039	–0.073	0.219 *	0.104**
Beta	–0.303	–0.067	–0.021	–0.054	0.007	0.041	0.077	0.110	0.106	0.103	–0.406**	–0.175**
Abs(Return)	0.453	0.068	0.024	–0.061	–0.072	–0.053	–0.058	–0.067	–0.095	–0.146	0.600**	0.166**
Return	0.044	0.004	0.013	–0.005	–0.002	–0.003	–0.007	–0.022	–0.008	–0.015	0.058	0.022
Return(t+3)	0.074	–0.003	–0.008	–0.012	–0.002	–0.009	–0.020	–0.019	–0.004	0.000	0.073	0.020
Return-Style	–0.014	0.008	–0.004	–0.006	0.002	–0.003	0.000	–0.012	0.007	0.022	–0.036	–0.005
Sharpe-Ratio	–0.031	0.031	0.036	0.048	0.040	–0.001	–0.032	–0.047	–0.045	–0.003	–0.028	0.050 *
TNA	–0.049	–0.110	–0.139	–0.097	–0.025	0.017	0.046	0.084	0.136	0.136	–0.186**	–0.168**

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table X. Are Specialist Funds Special? Robustness Check Using 48-Industry Portfolios
This table shows the average standardized portfolio-/fund-characteristics for MeanOverlap deciles (based on first lag) using 48-industry portfolios. Data are for March 2003—December 2014. For each month month I standardize the corresponding variables (rows) and calculate the mean value for each characteristic-decile combination. Columns (1) to (10) show the corresponding time-averages. The last two columns explicitly compare funds with low and high portfolio overlap (deciles 1 and 10), and those with below- and above-median values (deciles 1–5 versus 6–10), respectively. Significance tests are based on Wilcoxon rank-sum tests with Bonferroni-corrected p-values.

Full sample	Decile –HHI(t–1) 1 (Low)	2	3	4	5	6	7	8	9	10 (High)	1-10	(1:5) - (6:10)
Panel A												
Age^P	0.090	0.095	0.138	0.187	0.151	0.139	0.104	–0.044	–0.246	–0.613	0.703**	0.264**
Amihud^P	0.229	–0.035	–0.064	–0.175	–0.157	–0.193	–0.135	–0.036	0.090	0.474	–0.246**	–0.080**
B/M^P	0.217	0.091	0.063	–0.006	–0.080	–0.080	–0.047	–0.068	–0.070	–0.021	0.238**	0.114**
FirmRisk^P	0.131	–0.110	–0.129	–0.174	–0.147	–0.155	–0.094	0.028	0.157	0.492	–0.361**	–0.171**
InstOwn^P	–0.144	–0.262	–0.230	–0.224	–0.164	–0.155	–0.022	0.135	0.362	0.701	–0.844**	–0.409**
Mcap^P	0.101	0.260	0.230	0.230	0.186	0.157	0.055	–0.102	–0.313	–0.803	0.904**	0.402**
OwnConc^P	0.437	–0.010	–0.047	–0.113	–0.140	–0.182	–0.140	–0.077	–0.006	0.276	0.161**	0.051**
Price^P	0.013	0.075	0.052	0.095	0.110	0.137	0.039	–0.036	–0.106	–0.377	0.390**	0.138**
Turnover^P	–0.025	–0.237	–0.217	–0.226	–0.191	–0.120	–0.024	0.111	0.327	0.601	–0.626**	–0.358**
Panel B												
Abs(Flows)	0.111	0.089	0.047	–0.001	0.009	–0.028	–0.040	–0.058	–0.052	–0.077	0.188**	0.102**
Flows	0.067	0.074	0.051	0.039	0.020	–0.023	–0.030	–0.054	–0.057	–0.087	0.154**	0.100**
HHI	1.861	0.194	0.000	–0.114	–0.192	–0.253	–0.303	–0.348	–0.392	–0.447	2.307**	0.698**
Alpha	0.104	0.020	0.016	0.002	–0.002	–0.019	–0.027	–0.053	–0.014	–0.027	0.131	0.056**
Beta	–0.261	–0.065	–0.028	–0.033	–0.009	0.042	0.047	0.101	0.108	0.096	–0.358**	–0.158**
Abs(Return)	0.323	–0.009	–0.026	–0.096	–0.091	–0.099	–0.069	–0.043	0.005	0.106	0.217**	0.040
Return	–0.001	–0.025	–0.026	–0.020	–0.017	–0.014	0.006	–0.003	0.039	0.061	–0.062	–0.036
Return(t+3)	0.024	–0.015	–0.013	–0.020	–0.016	–0.010	–0.001	–0.008	0.018	0.041	–0.017	–0.016
Return-Style	–0.033	–0.009	–0.014	–0.008	0.003	0.001	0.013	0.000	0.029	0.018	–0.051	–0.024
Sharpe-Ratio	–0.046	0.004	0.031	0.031	0.022	0.020	0.026	–0.039	0.012	–0.062	0.016	0.017
TNA	–0.069	–0.145	–0.141	–0.047	–0.001	0.075	0.122	0.124	0.074	0.007	–0.076**	–0.161**

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table XI. Are Specialist Funds Special? Robustness Check Using Negative HHI
This table shows the average standardized portfolio-/fund-characteristics for –HHI deciles (based on first lag). Data are for March 2003—December 2014. For each month, I standardize the corresponding variables (rows) and calculate the mean value for each characteristic-decile combination. Columns (1) to (10) show the corresponding time-averages. The last two columns explicitly compare funds with low and high portfolio overlap (deciles 1 and 10), and those with below- and above-median values (deciles 1–5 versus 6–10), respectively. Significance tests are based on Wilcoxon rank-sum tests with Bonferroni-corrected p-values.

IV. Conclusions

In this paper, I connected two different streams of literature—namely, the literature on common asset holdings (or overlapping portfolios) and the literature on investor performances. I explored whether specialist mutual funds (those with low levels of portfolio overlap with other funds) are different from others. The main finding is that, while specialists indeed tend to hold portfolios that are significantly different from those of other funds, they do not generate significant performances. This result is driven by the factor-structure of stock returns, under which relatively modest levels of portfolio overlap can explain highly homogeneous performances.

Overall, these findings illustrate that common asset holdings have important implications for the diversity of investor performances. By purely looking at the cross-section of investment portfolios, policy makers, regulators, and investors may get a false sense of (portfolio) diversity, because even those mutual funds with very special portfolios do not necessarily generate performances that are significantly different from those of other funds. These findings are particularly important because I have focused on actively managed funds. In theory, these funds should have incentives to differentiate themselves from other funds in terms of their investment strategy, even more so in an increasingly competitive environment. I find no evidence of an increased tendency for innovation among these funds. An important avenue for future research is therefore why this is the case in order to understand in more detail the drivers of the increased portfolio overlap among mutual funds.

References

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt, 2005, "Competition and Innovation: An Inverted-U Relationship," *Quarterly Journal of Economics*, 120, 701–728.

Antón, M., and C. Polk, 2014, "Connected Stocks", *Journal of Finance*, 69, 1099-1127.

Banerjee, A., 1992, "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797–817.

Barras, L., O. Scaillet, and R. Wermers, 2010, "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas," *Journal of Finance*, 115, 811–841.

Bascompte, J., P. Jordano, C.J. Melián, and J.M. Olesen, 2003, "The Nested Assembly of Plant-Animal Mutualistic Networks," *PNAS*, 100, 9383–9387.

Bennett, J.A., R.W. Sias, and L.T. Starks, 2003, "Greener Pastures and the Impact of Dynamic Institutional Preferences," *Review of Financial Studies*, 16, 1203–1238.

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

Caccioli, F., S. Still, M. Marsili, and I. Kondor, 2013, "Optimal Liquidation Strategies Regularize Portfolio Selection," *European Journal of Finance*, 19, 554–571.

Carhart, M., 1997, "On Persistence in Mutual Fund Performance," *Journal of Finance*, 52, 56–82.

Chevalier, J., and G. Ellison, 1999, "Career Concerns of Mutual Fund Managers," *Quarterly Journal of Economics*, 104, 389–432.

Constantinides, G.M., 1986, "Capital Market Equilibrium with Transaction Costs," *Journal of Political Economy*, 94, 842–862.

Coval, J., and E. Stafford, 2007, "Asset Fire Sales (and Purchases) in Equity Markets," *Journal of Financial Economics*, 86, 479–512.

Cremers, M., and A. Petajisto, 2009, "How Active is Your Fund Manager? A New Measure that Predicts Performance," *Review of Financial Studies*, 22, 3329–3365.

Doidge, C., G.A. Karolyi, and R.M. Stulz, 2017, "The U.S. Listing Gap," *Journal of Financial Economics*, 123, 464–487.

Goldstein, I., H. Jiang, and D.T. Ng, 2017, "Investor Flows and Fragility in Corporate Bond Funds," *Journal of Financial Economics*, 126, 592–613.

Greenwood, R., A. Landier, and D. Thesmar, 2015, "Vulnerable Banks," *Journal of Financial Economics*, 115, 471–485.

Greenwood, R., and D. Thesmar, 2011, "Stock Price Fragility," *Journal of Financial Economics*, 102, 471–490.

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

Grinblatt, M., S. Titman, and R. Wermers, 1985, “Momentum Investment Strategies: Portfolio Performance, and Herding: A Study of Mutual Fund Behavior,” *American Economic Review*, 85, 1088–1105.

Harvey, C.R., H. Liu, and H. Zhu, 2016, “... and the Cross-Section of Expected Returns,” *Review of Financial Studies*, 29, 5–68.

Jensen, M., 1968, “The Performance of Mutual Funds in the Period 1945—1964,” *Journal of Finance*, 23, 389–416.

Kacperczyk, M., C. Sialm, and L. Zheng, 2005, “On the Industry Concentration of Actively Managed Equity Mutual Funds,” *Journal of Finance*, 60, 1983– 2011.

Kacperczyk, M., C. Sialm, and L. Zheng, 2007, “Industry Concentration and Mutual Fund Performance,” *Journal of Investment Management*, 5, 50–64.

Levy, A., and M. Livingston, 1995, “The Gains from Diversification Reconsidered: Transaction Costs and Superior Performance,” *Financial Markets, Institutions, and Instruments*, 4, 1–60.

Lisi, F., 2011, “Dicing with the Market: Randomized Procedures for Mutual Funds Evaluation,” *Quantitative Finance*, 11, 163–172.

Lo, A.W., C. Petrov, and M. Wierzbicki, 2003, “It’s 11 PM—Do You Know Where Your Liquidity Is? The Mean-Variance Liquidity Frontier,” *Journal of Investment Management*, 1, 55–93.

Malkiel, B.G., 2013, “Asset Management Fees and the Growth of Finance,” *Journal of Economic Perspectives*, 27, 97–108.

Patton, A.J., and A. Timmermann, 2010, “Monotonicity in Asset Returns: New Tests with Applications to the Term Structure, the CAPM, and Portfolio Sorts,” *Journal of Financial Economics*, 98, 605–625.

Pollet, J., and M. Wilson, 2010, “Average Correlation and Stock Market Returns,” *Journal of Financial Economics*, 96, 364–380

Rajan, R., 2005, “Has Financial Development Made the World Riskier?,” *Proceedings—Economic Policy Symposium—Jackson Hole*, Federal Reserve Bank of Kansas City, August, 313–369.

Roche, H., S. Tompaidis, and C. Yang, 2013, “Why Does Junior Put All His Eggs in One Basket? A Potential Rational Explanation for Holding Concentrated Portfolios,” *Journal of Financial Economics*, 109, 775–796.

Scharfstein, D.S., and J.C. Stein, 1990, “Herd Behavior and Investment,” *American Economic Review*, 80, 465–479.

Shleifer, A., and L.H. Summers, 1990, “The Noise Trader Approach to Finance,” *Journal of Economic Perspectives*, 4, 19–33.

Sun, Z., A. Wang, and L. Zheng, 2012, “The Road Less Traveled: Strategy Distinctiveness and Hedge Fund Performance,” *Review of Financial Studies*, 25, 96–143.

Van Nieuwerburgh, S., and L. Veldkamp, 2010, "Information Acquisition and Under-Diversification," *Review of Economic Studies*, 77, 779–805.

Wagner, W., 2011, "Systemic Liquidation Risk and the Diversity-Diversification Trade-Off," *Journal of Finance*, 66, 1141–1175.

Wahal, S., and A. Wang, 2011, "Competition Among Mutual Funds," *Journal of Financial Economics*, 99, 40–59.

Wermers, R., 1999, "Mutual Fund Herding and the Impact on Stock Prices," *Journal of Finance*, 54, 581–622.

Appendix

A. Industry Holdings Matrix

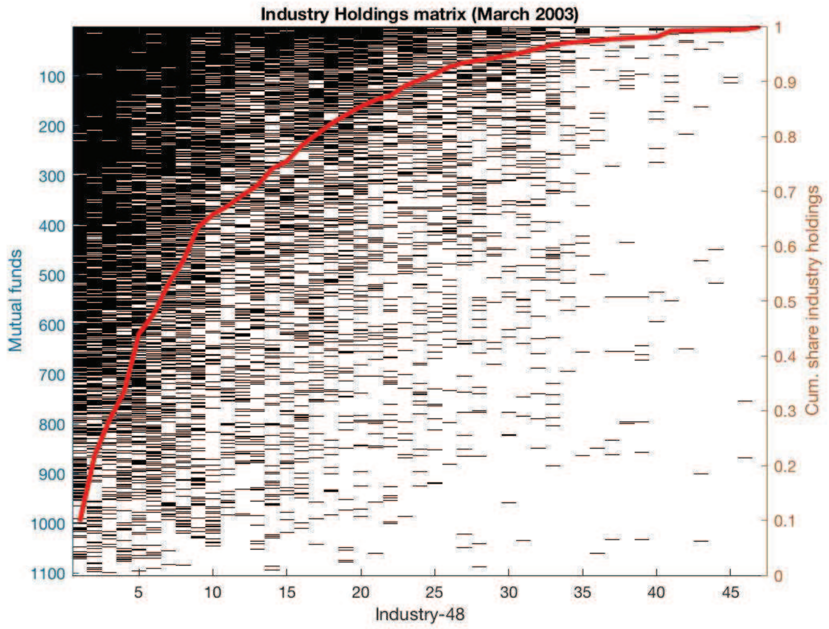

Figure A.1. Binary Industry Holdings Matrix in March 2003
Rows correspond to actively managed domestic equity mutual funds with more than 3 and less than 300 stock holdings (left y-axis), and columns to industries using the 48-industry classification of Ken French (`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library`).
A link between a fund and an industry exists if the fund holds more than one stock from that particular industry - this is shown as a black dot. Rows and columns are sorted according to the number of connections. The red line shows the cumulative share of the industry holdings relative to total fund holdings (right y-axis).

B. One Year and Two Year MeanOverlap Transition Matrices

Decile	MeanOverlap									
	(Low)									(High)
t \ t+1yr	1	2	3	4	5	6	7	8	9	10
1	0.963	0.035	0.001	0.001	.	.	.	.	.	.
2	0.031	0.922	0.046	.	.	.	.	.	.	.
3	.	0.043	0.897	0.057	0.003	.	.	.	.	.
4	.	.	0.052	0.888	0.058	0.002	.	.	.	.
5	.	.	.	0.053	0.891	0.055	.	.	.	.
6	.	.	.	.	0.045	0.895	0.059	0.002	.	.
7	.	.	.	.	.	0.050	0.890	0.057	0.002	.
8	.	.	.	.	.	.	0.052	0.883	0.063	0.002
9	.	.	.	.	.	.	0.001	0.056	0.884	0.059
10	.	.	.	.	.	.	.	.	0.050	0.950
t \ t+2yr	1	2	3	4	5	6	7	8	9	10
1	0.964	0.033	0.001	0.002	.	.	.	.	.	.
2	0.034	0.916	0.048	0.001	.	.	.	.	.	.
3	.	0.051	0.884	0.063	0.001	.	0.001	.	.	.
4	.	.	0.061	0.876	0.061	0.002	.	.	.	.
5	.	.	0.001	0.057	0.887	0.054	.	.	.	.
6	.	.	.	.	0.047	0.894	0.056	0.002	.	.
7	.	.	.	.	.	0.054	0.891	0.052	0.002	.
8	.	.	.	.	.	.	0.051	0.888	0.058	0.003
9	.	.	.	.	.	.	0.001	0.059	0.885	0.056
10	.	.	.	.	.	.	.	.	0.053	0.947

Table B.I. Transition Matrices at Longer Horizons
One year (top panel) and two year (bottom panel) transition matrices between deciles based on MeanOverlap. Each cell shows the probability of a mutual fund moving from decile *a* to decile *b*. Data are for the period March 2003—December 2014. Note: values below 0.1% are shown as "–".

C. Quarterly Diversification Level (–HHI) Transition Matrix

Decile	Diversification (–HHI)									
	(Low)									(High)
t+1 / t	1	2	3	4	5	6	7	8	9	10
1	0.935	0.056	0.004	0.002	0.001	.	.	.	.	.
2	0.053	0.856	0.078	0.007	0.003	0.002	.	.	.	.
3	0.002	0.081	0.810	0.095	0.007	0.002	0.002	.	.	.
4	.	0.002	0.100	0.787	0.097	0.009	0.004	.	.	.
5	.	.	0.003	0.100	0.772	0.109	0.012	0.003	.	.
6	.	.	.	0.002	0.114	0.764	0.109	0.008	0.002	.
7	.	.	.	0.001	0.002	0.111	0.774	0.104	0.007	.
8	.	.	.	.	.	0.002	0.099	0.796	0.098	0.004
9	.	.	.	.	.	.	0.002	0.088	0.838	0.072
10	.	.	.	.	.	.	.	0.001	0.061	0.937

Table C.I. Quarterly Transition Matrix Between Deciles Based on Fund Diversification Levels (Negative HHI)
Each cell shows the probability of a mutual fund moving from decile a to decile b. Data are for the period March 2003—December 2014. Note: values below 0.1% are shown as “.”.

D. Classification - MeanOverlap versus –HHI

Decile	t \ t+1	Diversification (–HHI) (Low) 1	2	3	4	5	6	7	8	9	(High) 10
(Low)	1	0.315	0.103	0.116	0.094	0.098	0.080	0.069	0.054	0.038	0.034
	2	0.126	0.090	0.065	0.074	0.082	0.094	0.098	0.136	0.122	0.112
	3	0.131	0.114	0.085	0.083	0.078	0.061	0.059	0.075	0.121	0.193
	4	0.148	0.145	0.085	0.066	0.068	0.076	0.112	0.101	0.085	0.115
Mean-	5	0.134	0.182	0.116	0.076	0.051	0.045	0.049	0.067	0.140	0.141
Overlap	6	0.101	0.197	0.188	0.117	0.084	0.057	0.040	0.034	0.047	0.136
	7	0.034	0.114	0.217	0.195	0.133	0.091	0.072	0.053	0.044	0.049
	8	0.009	0.041	0.099	0.195	0.180	0.176	0.133	0.080	0.055	0.032
	9	.	0.013	0.025	0.080	0.162	0.204	0.209	0.161	0.101	0.044
(High)	10	.	.	.	0.022	0.066	0.114	0.159	0.238	0.249	0.148

Table D.I. Comparison of Classifications Using MeanOverlap and –HHI
Each cell shows the probability of a mutual fund being in MeanOverlap decile a and HHI decile b. Data are for the period March 2003—December 2014. Note: values below 0.1% are shown as “.”.

E. Correlation Matrix

	Correlation matrix																		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)
(1) Age^P	1.000	–0.016	–0.044	–0.113	0.006	0.158	–0.137	0.116	–0.063	–0.094	–0.062	0.017	–0.013	0.006	–0.014	–0.001	0.002	0.013	0.288
(2) Amihud^P	–	1.000	0.132	0.066	–0.018	–0.036	0.093	–0.046	–0.014	0.005	0.002	0.009	–0.015	–0.020	0.023	–0.007	0.001	–0.033	–0.008
(3) B/M^P	–	–	1.000	0.145	0.091	–0.175	0.083	–0.403	0.096	0.018	–0.018	0.075	–0.098	0.007	0.247	–0.123	–0.019	–0.334	–0.039
(4) FirmRisk^P	–	–	–	1.000	0.071	–0.435	0.410	–0.358	0.426	0.056	0.029	0.065	–0.040	0.070	0.209	0.109	0.019	–0.084	–0.069
(5) InstOwn^P	–	–	–	–	1.000	–0.430	–0.495	0.136	0.353	–0.010	–0.038	–0.068	–0.009	0.042	0.125	–0.018	0.002	–0.034	–0.044
(6) Mcap^P	–	–	–	–	–	1.000	–0.414	0.384	–0.425	–0.041	–0.018	–0.042	–0.025	–0.004	–0.164	–0.004	–0.002	0.087	0.118
(7) OwnConc^P	–	–	–	–	–	–	1.000	–0.467	–0.054	0.041	0.058	0.146	0.050	–0.070	0.015	0.050	0.009	0.039	–0.019
(8) Price^P	–	–	–	–	–	–	–	1.000	–0.142	–0.035	–0.011	–0.032	0.029	0.008	–0.151	0.053	0.023	0.257	0.065
(9) Turnover^P	–	–	–	–	–	–	–	–	1.000	0.043	–0.014	0.094	–0.066	0.205	0.426	–0.170	–0.007	–0.366	–0.072
(10) Abs(Flows)	–	–	–	–	–	–	–	–	–	1.000	0.455	0.017	0.028	–0.012	0.038	0.027	0.022	0.001	–0.046
(11) Flows	–	–	–	–	–	–	–	–	–	–	1.000	0.015	0.085	–0.034	0.005	0.052	0.037	0.070	–0.006
(12) HHI	–	–	–	–	–	–	–	–	–	–	–	1.000	0.008	–0.110	0.121	0.004	–0.009	–0.042	–0.017
(13) Alpha	–	–	–	–	–	–	–	–	–	–	–	–	1.000	–0.296	–0.035	0.061	0.151	0.253	0.019
(14) Beta	–	–	–	–	–	–	–	–	–	–	–	–	–	1.000	0.147	–0.045	–0.006	–0.158	–0.014
(15) Abs(Return)	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1.000	–0.058	0.062	–0.328	–0.034
(16) Return	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1.000	0.362	0.236	0.003
(17) Return-Style	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1.000	0.096	0.004
(18) Sharpe-Ratio	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1.000	0.043
(19) TNA	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1.000

Table E.I. Correlation Matrix for Fund Characteristics
This table shows the (unconditional) correlations between the portfolio-/fund-characteristics shown in Table V in the main text.

F. Robustness: Significance Tests with Raw Variables

Full Sample	Decile MeanOverlap(t–1)											(1:5)
	1 (Low)	2	3	4	5	6	7	8	9	10 (High)	1-10	- (6:10)
Panel A												
Log(Age^P)	22.662	22.847	23.895	26.431	27.961	31.499	34.945	38.060	39.025	39.461	–16.799**	–11.839**
Log(Amihud^P)	–18.619	–19.703	–20.425	–20.937	–21.459	–21.898	–22.419	–22.836	–22.945	–23.030	4.412**	2.397**
B/M^P	0.542	0.490	0.467	0.453	0.426	0.416	0.414	0.420	0.419	0.413	0.129**	0.059**
FirmRisk^P	0.266	0.219	0.206	0.186	0.172	0.159	0.143	0.126	0.123	0.122	0.144**	0.075**
InstOwn^P	0.383	0.418	0.412	0.391	0.381	0.353	0.330	0.316	0.312	0.309	0.073**	0.073**
Log(Mcap^P)	21.467	21.903	22.533	23.284	23.757	24.295	24.691	24.883	24.962	25.026	–3.559**	–2.183**
OwnConc^P	0.387	0.331	0.307	0.289	0.275	0.266	0.256	0.250	0.247	0.243	0.144**	0.065**
Price^P	3.477	3.638	3.769	3.893	3.992	4.038	4.072	4.099	4.099	4.112	–0.635**	–0.330**
Turnover^P	–4.660	–4.581	–4.554	–4.580	–4.587	–4.692	–4.761	–4.849	–4.860	–4.857	0.197**	0.211**
Panel B												
Abs(Flows)	0.037	0.035	0.034	0.034	0.033	0.031	0.029	0.027	0.024	0.021	0.016**	0.008**
Flows	0.010	0.008	0.004	0.006	0.004	0.005	0.004	0.001	–0.002	–0.004	0.014**	0.005**
HHI	0.085	0.033	0.026	0.027	0.025	0.026	0.024	0.020	0.016	0.012	0.072**	0.019**
Alpha	–0.001	0.000	0.000	0.000	–0.001	–0.001	–0.001	–0.001	–0.001	–0.001	0.001	0.001
Beta	0.899	0.948	0.960	0.961	0.974	0.968	0.985	0.983	0.993	0.994	–0.095**	–0.036**
Abs(Return)	0.040	0.037	0.035	0.034	0.033	0.031	0.029	0.028	0.028	0.027	0.013**	0.007
Return	0.014	0.014	0.013	0.013	0.013	0.012	0.011	0.011	0.011	0.011	0.003	0.002
Return(t+3)	0.009	0.009	0.009	0.009	0.009	0.008	0.007	0.008	0.007	0.007	0.002	0.001
Return-Style	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sharpe-Ratio	0.366	0.378	0.389	0.399	0.386	0.390	0.382	0.396	0.399	0.390	–0.023	–0.008
Log(TNA)	18.750	19.209	19.131	19.072	19.126	19.096	19.171	19.421	19.773	19.877	–1.127**	–0.410**

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table F.I. Are Specialist Funds Special? Robustness Check Using Nonstandardized Characteristics
This table shows the average **nonstandardized** portfolio-/fund-characteristics for MeanOverlap deciles (based on first lag). Data are for March 2003—December 2014. For each month, I calculate the mean value for each characteristic-decile combination. Columns (1) to (10) show the corresponding time-averages. The last two columns explicitly compare funds with low and high portfolio overlap (deciles 1 and 10), and those with below- and above-median values (deciles 1–5 versus 6–10), respectively. Significance tests are based on Wilcoxon rank-sum tests with Bonferroni-corrected p-values.

G. Fund-Level Performance Regressions (4F-Alpha)

	Alpha		
	(t)	(t+6)	(t+12)
Flows(t–1)	0.057**	0.037**	–0.019**
	(0.005)	(0.004)	(0.004)
Log(TNA(t–1))	–0.055	–0.316**	–0.406**
	(0.039)	(0.047)	(0.050)
MeanOverlap(t–12)			
Decile 2	–0.109	–0.161**	–0.127
	(0.057)	(0.061)	(0.079)
Decile 3	–0.126 *	–0.084	–0.012
	(0.063)	(0.064)	(0.088)
Decile 4	–0.153 *	–0.147 *	0.012
	(0.066)	(0.068)	(0.088)
Decile 5	–0.158 *	–0.127	0.056
	(0.072)	(0.078)	(0.095)
Decile 6	–0.163 *	–0.136	0.045
	(0.080)	(0.088)	(0.100)
Decile 7	–0.151	–0.071	0.109
	(0.085)	(0.094)	(0.104)
Decile 8	–0.185 *	–0.073	0.147
	(0.090)	(0.096)	(0.109)
Decile 9	–0.145	–0.043	0.153
	(0.094)	(0.101)	(0.110)
Decile 10	–0.147	–0.033	0.153
(High)	(0.096)	(0.106)	(0.111)
Fund FEs	Yes	Yes	Yes
Obs.	211,367	185,128	155,130
adj.-R^2	0.220	0.229	0.218

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table G.I. Fund-Level Performance Regressions
All variables (except for discrete MeanOverlap deciles) are standardized (zero mean, unit standard deviation) for each cross section. Time FEs are therefore insignificant. Standard errors clustered by month.

H. Flow-Performance Relation by MeanOverlap Decile

Dep. Var.:	Decile MeanOverlap(t–1)									
Flows(t)	1 (Low)	2	3	4	5	6	7	8	9	10 (High)
Alpha(t–1)	0.297**	0.952**	0.831**	0.825**	0.803**	1.100**	1.080**	1.173**	0.964**	1.097**
	(0.044)	(0.126)	(0.133)	(0.094)	(0.098)	(0.112)	(0.122)	(0.139)	(0.137)	(0.147)
Flows(t–1)	0.215**	0.073 *	0.029	0.095**	0.095**	0.115**	0.111**	0.081**	0.152**	0.048
	(0.021)	(0.034)	(0.045)	(0.034)	(0.028)	(0.032)	(0.029)	(0.028)	(0.020)	(0.027)
Log(TNA)	0.002**	0.001**	0.002**	0.001**	0.001 *	0.001**	0.002**	0.002**	0.002**	0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(Age)	−0.011**	−0.011**	−0.012**	−0.009**	−0.008**	−0.008**	−0.007**	−0.006**	−0.005**	−0.006**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Time FE's	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs	26,364	25,355	24,922	24,849	24,853	24,853	24,615	24,543	24,837	26,334
adj.-R^2	0.077	0.029	0.021	0.025	0.024	0.038	0.032	0.021	0.032	0.013

** Significant at the 0.01 level.

* Significant at the 0.05 level.

Table H.I. Flow-Performance Relation by MeanOverlap Decile

This table shows flow-performance relations for the sample of domestic equity funds from March 2003—December 2014. I use the same methodology as Goldstein, Jiang, and Ng (2017) and cluster standard errors by mutual fund.