OXFORD

Trade Less and Exit Overcrowded Markets: Lessons from International Mutual Funds[*]

Teodor Dyakov[1], Hao Jiang[2], and Marno Verbeek[3]

[1]School of Business and Economics, Vrije Universiteit Amsterdam and Tinbergen Institute, [2]Broad College of Business, Michigan State University, and [3]Rotterdam School of Management, Erasmus University Rotterdam

Abstract

We study active investment skills in relation to returns to scale in the active mutual fund industry. Using a sample of 13,807 funds from sixteen domicile countries investing in forty-two equity markets from 2001 to 2014, we find that they achieve negative trading performance on average, driven mainly by particularly low returns to their trades in US equities. Exploring their investment environment, we find convincing evidence of decreasing returns to scale around the world, especially for the US market. Based on theory of optimal fund size, we estimate the optimal size of the active mutual fund industry. We find that the active mutual fund industry in USA has exceeded the optimal level, whereas in the international markets, there may still be room for further expansion. Consistent with this view, we find that mutual fund managers have been gradually reallocating their assets away from the USA and more into international equity markets.

1. Introduction

The asset management industry has been expanding tremendously around the globe. According to the Boston Consulting Group (2016), global assets under management in this industry grew from \$29 trillion in 2002 to \$71 trillion in 2015. Among global asset

managers, open-end mutual funds stand prominently in terms of industry size. The Investment Company Institute (2017) estimates that as of the second quarter of 2017, open-end mutual funds manage more than $36 trillion of assets worldwide, excluding funds of funds.[1] Since actively managed funds dominate the mutual fund industry, it is important to understand how the global rise of active fund managers influences their performance. Unfortunately, this question is not well understood for the global active fund industry.

In this paper, we fill the gap by studying how investment skills interact with the scale of the active fund industry to impact their performance. To infer investment skills, we exploit funds' holdings-based and trades-based performance, which the literature has considered to be more informative about active investment skills than performance measures based on overall fund returns (see, e.g., Grinblatt and Titman, 1989; Chen, Jegadeesh, and Wermers, 2000). In addition, the use of holdings information allows us to disaggregate the performance of international mutual funds across the countries they invest in. Our sample comprises 13,807 actively managed mutual funds from sixteen domicile countries investing in forty-two equity markets during the period 2001 to 2014. Through this global lens, we extend a growing literature on this important topic that focuses on the US active fund industry (e.g., Berk and Green, 2004; Pastor, Stambaugh, and Taylor, 2015; Berk and van Binsbergen, 2017).

We start by describing the average performance of trading by active funds around the world. We find that, in the aggregate, mutual funds tend to lose money on their trading, even before costs: the stocks they buy underperform those they sell by 18 basis points (bps) per month in the subsequent quarter (*t*-statistic $=-2.0$), after adjustments for passive benchmarks. Using the measure of dollar value added proposed by Berk and van Binsbergen (2015) (BvB), we estimate that global active mutual funds tend to destroy value by $1.19 billion per month (*t*-statistic $= -2.5$) in total through their trading activities. Although the negative trading performance comes from both US and internationally domiciled funds, it tends to concentrate in the US equities they trade. For instance, US domiciled funds achieve an average negative return of 34 bps per month (*t*-statistic $=-2.4$) to their trades in US equities, whereas their trades largely break even in the international equity markets. A similar pattern holds for internationally domiciled funds. This initial result suggests that the US equity market may be more crowded with active funds, which constrains their trading performance.

To formally examine the impact of the scale of active funds on their performance, we test for the presence of decreasing returns to scale in the USA and international equity markets. To this end, we extend the instrumental-variables approach developed by Pastor, Stambaugh, and Taylor (2015) with the modifications of Zhu (2018), and use both trading and holdings-based performance of mutual funds to test for diseconomies of scale. To measure benchmark-adjusted performance, we use both the traded funds approach proposed by BvB and the Daniel *et al.* (1997) (DGTW) adjustment procedure. At the industry level, we find strong evidence of decreasing returns to scale in active fund management when they invest in US equities. For instance, based on the BvB alpha, a 1% expansion of active funds relative to the US equity market value associates with a decline of 14 bps per month (*t*-statistic $=-3.1$) in returns to their equity trades, and a decline of 7 bps per month

---

1   The estimates in this paragraph are based on Boston Consulting Group's 2016 Global Asset Management report "Doubling Down on Data," and the Investment Company Institute's global research and statistics, available from https://www.iciglobal.org/iciglobal/research/stats.

(t-statistic $= -2.1$) in returns to their equity holdings; based on the DGTW alpha, we obtain a consistent pattern. These results clearly illustrate the adverse impact of crowded active investing at the industry level on individual funds' performance.

For international equities, we find that the DGTW alpha generates sharper results. This is primarily driven by the fact that the offering of region-specific index funds along the value and momentum dimension is a recent phenomenon, which does not allow us to use those style benchmarks in the traded funds approach for our international sample. It is for this practical reason that the DGTW procedure may be able to offer sharper inference on diseconomies of scale for international mutual funds. Grouping international mutual funds together, we find that an increase in active fund industry size has a strong negative impact on their trading performance. The magnitudes are comparable to those for the US funds. Looking at each region individually, we are able to find reliable evidence of decreasing returns to scale for the funds investing in Asia-Pacific, Europe, and Emerging Markets (EMs). Our data do not enable us to find statistically significant evidence of decreasing returns to scale for Canada and Japan, although the point estimates have the correct sign.

These findings naturally raise the question: What is the optimal size of the active mutual fund industry in these markets? To make initial progress in addressing this difficult but important question, we build on the optimal fund size model as in Berk and Green (2004) and Berk and van Binsbergen (2017). Assuming a linear relation between gross (before-fees) fund alpha and fund size, Berk and van Binsbergen (2017) postulate a simple closed-form solution for the optimal fund size. The optimal size is driven by two parameters, the gross alpha on the first cent a fund manager extracts from financial markets and the rate at which a fund's gross alpha decreases with fund size. Extending their theoretical results, we develop a simple statistical distribution theory for the BvB estimator of the optimal industry size. Our results indicate that the size of active fund industry has exceeded the optimal level at the 95% confidence level in the USA at the end of our sample period. For international markets, however, the actual size lies within the 95% confidence interval across the five regions. The point estimates for efficient industry size show that for Asia-Pacific and EMs, there is still substantial room for further expansion of the active fund industry.

Our findings on the optimal industry size need to be interpreted with caution, for three reasons. First, the international results are estimated using DGTW alpha and may be too harsh on active managers. Implementing factor-based strategies may be considered as a skill too, and the DGTW alpha cannot capture this skill component. Second, even though the holdings based approach is informative about managerial skill, it is based on quarter-end portfolio snapshots. Thus, we may miss the value active managers generate between the quarterly snapshots (e.g., Kacperczyk, Sialm, and Zheng, 2008). Finally, our sample period is relatively short and captures a very specific period of time when active management did not fair particularly well.

Although our statistical estimation is surely crude, it has a clear, directional implication: fund managers investing primarily in the US market would have incentives to diversify their investments into markets with a less crowded active fund industry. To examine this prediction, we compute changes in the amount of assets that US domiciled funds invest in the US and international equity markets. We find that, over our sample period from 2001 to 2014, US domiciled funds cumulatively withdrew $400 billion of assets out of US equity, while increasing their investments in international equity by a similar amount. As a result, the allocation to US equity by US domiciled funds decreased from 91% to 71% over our sample period (see, e.g., BvB for a related observation).

So far, our empirical analyses are at the level of individual mutual funds. To exploit the richness of our data sets, we perform multivariate regressions at the stock-level to test for the influence of diseconomies of scale on trading performance. Our panel regressions show that, in equity markets with more active mutual fund money chasing investment opportunities, fund trades tend to achieve lower performance. The negative association between stock returns and the interaction of mutual fund trades and fund industry size is strong, and robust to controlling for country-fixed, time-fixed, and stock-industry-fixed effects and many stock characteristics. The size of the active industry appears to be a statistically stronger predictor of future returns than stock-level herding. These results corroborate the close connection between poor trading performance and decreasing returns to scale in active fund management.

The remainder of this paper starts with a brief discussion of related literature evaluating the trading performance of active mutual funds. In Section 3, we provide more details on the data construction and descriptive statistics. After discussing the choice of benchmarks in Section 4, we continue analyzing the performance of aggregate mutual fund trades in Section 5 by relating changes in mutual fund holdings to subsequent stock returns. In Section 6, we relate the trading performance at the fund level to the size of the active fund industry in the country of investment and fund size to investigate the nature of the decreasing returns to scale. This section also includes our calculations of optimal industry size. We also relate performance at the stock-level to fund trading, the size of the active industry and herding. After a number of robustness checks in Section 7, Section 8 provides a more detailed analysis of the trading performance among US stocks by US mutual funds, for which a longer times series is available. The results confirm the poor trading performance since 2000, and support our general conclusion that the crowdedness of the US equity market has become detrimental to active funds' trading returns.

2. Related Literature

The literature on mutual fund performance is vast. To conserve space, we focus this review on the trading performance of actively managed mutual funds. This literature has offered a number of techniques to evaluate their trading skills.

First, the most commonly used approach is to proxy mutual fund trades using changes in their quarterly stock holdings. For instance, using this method, Chen, Jegadeesh, and Wermers (2000) show that stocks bought by domestic US equity mutual funds outperform stocks sold by 0.73% per quarter during the period 1975–95, after adjusting for common style exposures. Their evidence is in line with the estimates offered by Daniel et al. (1997). Baker *et al.* (2010) find that mutual funds' stock purchases outperform their sales around subsequent earnings announcements. These earlier studies point to the existence of trading skills among active mutual funds.

Studies using more recent data, however, paint a less optimistic picture. For instance, Duan, Hu, and McLean (2009) extend the sample of Chen, Jegadeesh, and Wermers (2000) by 8 years and find that during the period 1995–2003, the difference in abnormal returns between the stocks US mutual funds buy and sell is statistically indistinguishable from zero. In the cross-section of stocks they are able to find evidence of trading skills among stocks with higher idiosyncratic volatilities, consistent with the story of higher limits to arbitrage for these stocks. It is notable that the suggestive evidence reported in Duan, Hu, and McLean (2009) is in line with a general decline in mutual fund alpha observed by, for

example, Barras, Scaillet, and Wermers (2010) and Lewellen (2011). In this context, our study represents a leap in terms of the sample of mutual funds, equity markets, and time periods examined; it also brings us closer toward understanding the shifts in mutual fund trading performance in terms of increased competition among mutual funds in a deteriorating investment environment (see Berk and Green, 2004; Pastor and Stambaugh, 2012) and their increased tendency to trade in herds.

A number of studies, using the same quarterly stock holdings data, examine the performance of a specific form of mutual fund trading, namely, their herding activities. Using the LSV measure (Lakonishok, Shleifer, and Vishny, 1992), earlier studies such as Grinblatt, Titman, and Wermers (1995) and Wermers (1999) find a positive relation between mutual fund herding and subsequent returns. Our study using broader and more recent mutual fund data find an inverse relation between fund herding and subsequent stock returns. Our results are consistent with Dasgupta, Prat, and Verardo (2011) and Jiang and Verardo (2018), who show lower performance of herd-like trades.

Second, several recent studies have used institutional trading data from Abel Noser (ANcerno Ltd.) to assess trading performance. This data set covers the trades executed by the institutional clients of Abel Noser at daily frequency. With it, Puckett and Yan (2011) estimate that during the period between 1999 and 2005, interim (intraquarter) trades by these institutions generate abnormal returns between 0.20% and 0.26% per year after trading costs. Based on this evidence, they argue that studies using quarterly mutual fund trades are likely to underestimate the trading skills of mutual funds. In a subsequent study using the same data set, Chakrabarty, Moulton, and Trzcinka (2017) argue that the classification of interim trades by Puckett and Yan (2011) is overly narrow and represents only a small portion of short-term fund trades. With their broader definition of short-term fund trades, they find that short-term fund trading achieves negative returns on average. They argue that the high-frequency trading data support the conclusions reached by studies using quarterly fund holdings data.

Third, many studies have used the association between mutual fund turnover and fund performance to evaluate the trading skills of mutual funds. The literature has reached mixed conclusions. For instance, Elton et al. (1993) and Carhart (1997) find that turnover is negatively related to fund performance, Edelen, Evans, and Kadlec (2007) find an insignificant relation between turnover and fund returns, and Dahlquist, Engström, and Söderlind (2000) find a positive relation between turnover and fund returns. More recently, Pastor, Stambaugh, and Taylor (2017) argue that it is important to include fund-fixed effects in the turnover-performance regressions, which leads to a positive relation. There are at least two advantages of using fund turnover to capture fund trades: first, it is a catch-all measure of fund trading activities, reflecting both interim and interquarter fund trades; second, it can be directly connected to observed mutual fund alpha, which can be used by investors for mutual fund selection. The downside of this measure is that it combines mutual fund buys and sales at the fund portfolio level, which makes it less powerful to evaluate fund trading skills; on the other hand, stock-level trading measures could render the analysis of trading skills richer and statistically more powerful.

Our study is also related to a nascent literature on the performance of international mutual funds. BvB shows the growing importance of foreign equity for the performance of US mutual funds—the fraction of assets under management of funds that exclusively hold US equities has dropped from 45% in 1977 to <23% in 2011. Ferreira et al. (2013) provide the first systematic investigation of the net performance of mutual funds around the world.

They find that between 1995 and 2007, local mutual funds from twenty-seven countries, that is, those investing in their domestic markets only, underperform their benchmarks by 0.20% per quarter after fees. However, they do not study the performance of international funds, that is, those investing in both local and international markets. Moreover, Ferreira *et al.* (2017) compare the effect of local and foreign institutional ownership on subsequent stock returns. Using their broad sample of institutions, they find that the level of local institutional ownership forecasts future returns, but changes in local institutional ownership do not. They also find that trading by foreign institutions is negatively correlated with subsequent returns. However, it is difficult to infer what type of foreign institutions drives their results. Leippold and Rueegg (2018) find that internationally, most active funds have zero alphas when compared with investable benchmarks. Similarly to our work, their paper indicates that the Berk and Green equilibrium is unlikely to be rejected outside of the USA. There are, however, important differences between our studies. Their findings are based on estimated net alphas, whereas we employ gross alphas exploiting the underlying stock holdings. This allows us to estimate diseconomies of scale across different markets. Using the optimal size theory of BvB, we are able to show that the US industry has become larger than its optimal size. Our approach has the additional advantage that we are better able to identify where funds invest.

Several recent papers document the existence of decreasing returns to scale in the mutual fund industry. Building upon Berk and Green (2004) and Pastor and Stambaugh (2012), Pastor, Stambaugh, and Taylor (2015) find a negative relation between industry size and fund performance, controlling for the endogeneity of fund size using a recursive demeaning procedure. This analysis is extended by Zhu (2018). BvB stress that value added is a better measure of managerial skill than (gross or net) alpha; Berk and van Binsbergen (2017) expand upon this by stressing the implications of rational expectations equilibrium in money management. One implication is the existence of optimal sizes for mutual funds and the industry as a whole. Our paper is unique in fleshing out the link between trading performance and industry-level diseconomies of scale in international equity markets, and the first to empirically establish a rough estimate for the optimal size of the active mutual fund industry in the USA and other international markets.

3. Data Construction and Descriptive Statistics

For our analysis we construct a representative survivorship free data set of actively managed international mutual funds and their quarterly trades, with as little biases as possible. Our datasets combine portfolio holdings data from Factset and stock-level information from Datastream and Worldscope and cover quarterly snapshots of the equity holdings of active mutual funds around the world in the period 2001–14.[2] We complement our international trading dataset with the more traditional sample of trades by domestic US open-end mutual funds, starting in 1980, that combines the Thomson Financial/CDA S12 fund holdings database, the CRSP Mutual Fund Database, and the CRSP daily and monthly stock files. The complete sample construction is described in Appendices A–D.

---

2 Note that our sample selection procedures differ from earlier research utilizing the Factset holdings, such as Ferreira and Matos (2008), who focus on aggregate institutional ownership, including pension funds, insurances, etc., and do not restrict their sample to domiciles where reporting biases are least likely.

The summary statistics of the two samples are reported in Table I. In total, the 13,807 active funds in the international sample are domiciled in sixteen developed countries (Panel A), 4,569 of them in the USA. The US sample, starting in 1980, includes only 2,394 domestic equity funds. Thus, the international sample covers more US domiciled funds than the US sample. There are two reasons for this. First, the coverage of the international sample is broader—there are both domestic and international funds, as well as funds that may not be necessarily equity-only. In contrast, the US sample only covers actively managed domestic US equity mutual funds that cover specific investment objectives: growth, aggressive growth, or growth and income. Second, the data filters available in Factset used to identify actively managed open-ended funds may perform imperfectly and thus accidentally include funds that are not necessarily active or open-ended. Consistent with earlier research, we observe that the average size [total net assets (TNAs)] of mutual funds in the USA has been growing over time (e.g., BvB) and is much larger than for funds domiciled outside the USA (e.g., Khorana, Servaes, and Tufano, 2005; Ferreira et al., 2013). Means among both fund samples are higher than medians due to the presence of a few very large funds. Net fund returns among US funds are much smaller in the most recent decade, which is driven by the crisis period after 2007. Lastly, we note that reported turnover among the sample of US funds is generally higher than the turnover we infer from the reported holdings of funds in the international sample. Note that there is no information in Factset regarding net returns, flows, and expenses. Thus, the last three columns of Panel A are empty.

In Panels C and D, we report summary statistics of stock characteristics for the international and US samples, respectively. Note that the US stock sample data are based on CRSP, whereas the international stock sample comes from Datastream and Worldscope.[3] On average, stock ownership by active funds in the USA is twice as large as in the international sample (7.1% versus 3.7%). Trading, or changes in ownership, are at similar levels at 0.07% per stock per quarter. The mean stock size among international stocks is larger, because of the presence of many small stocks in the US sample. Notably, turnover among US stocks is larger, whereas most other stock characteristics are distributed similarly.

The average active fund ownership among international stocks, based on Factset holdings, is lower than the institutional ownership reported in previous research. For example, Ferreira and Matos (2008) report an average 7.4% institutional ownership among international stocks. In contrast, the average stock ownership among active funds in our sample is 3.7%. The difference arises due to two key data selection procedures. First, previous studies focus on total institutional ownership, while our focus is on ownership by active mutual funds only. Second, since we are interested in aggregate trading performance, we restrict our sample selection to fund domiciles where reporting biases are least likely. Appendix A outlines how we restrict our sample to funds from the sixteen domiciles listed in Table I and investing in forty-two equity markets.

4. Constructing Benchmarks

For the main part of our analyses, we use two different approaches to construct relevant benchmarks to evaluate the performance at the fund, stock, or aggregate level. Our primary methodology is based on comparing a fund's trading returns with a set of alternative

---

3 Further note that for consistency, US stock-specific information in the international sample is also based on data from Datastream and Worldscope.

Table I Descriptive statistics

This table presents descriptive statistics on the number of unique funds and averages (mean and median) of the number of quarterly stock holdings, end of quarter net assets under the management (in $mill), net monthly return (in %, based on changes in NAV) and flows (in %), and yearly reported turnover and expense ratios (both in %), for both the international sample (Panel A) and the US sample (Panel B). In Panels C and D, we provide the mean, standard deviation, minimum, and maximum of stock-level variables separately for the sample of international stocks and US stocks, respectively. FracHold is the ownership by active funds, defined as the fractional holdings owned by all funds in our sample and expressed in percentages; ΔFracHold is the change in FracHold; BTM is the log of industry-adjusted book-to-market ratio; SIZE is the log of primary issue market capitalization in $mill; RET is the quarterly raw return; TURN is the stock turnover, defined as monthly trading volume scaled by the number of shares outstanding; VOL is the annualized stock volatility; PRICE is the stock price in $; MSCI is an indicator variable taking 1 if the stock is part of the MSCI World index and 0 otherwise (available only for the sample of international stocks); DY is the dividend yield in percent; ANALYSTS is the number of analysts following the stock in the IBES database; ILLIQ is the Amihud's illiquidity measure; and MOM is the 9-month return proceeding the calculation of RET. Data sources are provided in Section 3.

Domicile	Number of funds	Number of stockholdings		Net assets ($mill)		Yearly turnover (%)		Net monthly ret (%)		Monthly flow (%)		Yearly expense ratio (%)	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Panel A: International fund sample, 2001–14													
All domiciles	13,807	154	86	718	141	56	47						
Austria	162	108	87	95	48	48	39						
Belgium	272	147	95	129	67	38	32						
Canada	1,177	125	79	327	111	48	42						
Switzerland	245	196	106	197	101	43	29						
Germany	439	93	75	324	86	54	45						
Denmark	159	138	82	141	66	43	35						
Spain	326	80	66	79	37	46	36						
Finland	100	83	65	162	97	49	41						
France	690	111	71	295	107	53	41						
UK	1,358	122	77	466	173	51	44						
Ireland	750	188	90	312	120	52	45						
Luxembourg	3,099	151	85	328	97	51	44						
The Netherlands	133	120	91	627	204	49	41						

(continued)

Table I Continued

	Number of funds	Number of stockholdings		Net assets ($mill)		Yearly turnover (%)		Net monthly ret (%)		Monthly flow (%)		Yearly expense ratio (%)	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Panel A: International fund sample, 2001–14													
Norway	86	142	75	341	112	47	35					1.33	1.21
Sweden	242	167	115	386	167	55	44					1.35	1.22
USA	4,569	179	94	1,322	245	67	56					1.34	1.22
Panel B: US fund sample, 1980–2012													
Year													
1980–2000	2,372	91	60	754	127	89	66	1.26	1.39	2.32	0.20		
2001–12	1,752	118	71	1,932	359	86	65	0.29	0.79	0.66	−0.57		
1980–2012	2,394	104	65	1,291	205	88	66	0.82	1.11	1.35	−0.22		

Panel C: International stock sample, 2001–14

	N	Mean	Std	Min.	Max.
FracHold	793,981	3.770	6.934	0.000	86.204
ΔFracHold	793,981	0.068	1.597	−66.097	45.049
BTM	793,981	0.281	0.995	−1.666	2.070
SIZE	793,981	5.463	1.853	1.914	8.797
RET	793,981	0.035	0.190	−0.324	0.463
TURN	793,981	0.085	0.113	0.001	0.424
VOL	793,981	0.391	0.213	0.130	0.950
PRICE	793,981	17.343	22.058	0.175	81.643
DY	793,981	1.774	1.946	0.000	6.840
ANALYSTS	793,981	5.178	6.719	0.000	22.000
ILLIQ	793,981	0.254	0.734	0.000	4.425
MOM	793,981	0.117	0.387	−0.527	1.044
MSCI	793,981	0.096	0.294	0.000	1.000

Panel D: US stock sample, 1980–2012

	N	Mean	Std	Min.	Max.
FracHold	541,836	7.144	7.497	0.000	96.465
ΔFracHold	541,836	0.070	1.149	−72.593	52.856
BTM	541,836	0.350	0.864	−1.347	1.909
SIZE	541,836	5.039	1.894	1.965	8.656
RET	541,836	0.033	0.217	−0.352	0.513
TURN	541,836	0.283	0.291	0.022	1.092
VOL	541,836	0.455	0.242	0.154	1.045
PRICE	541,836	17.843	14.731	1.720	53.320
DY	541,836	0.814	2.188	0.000	8.062
ANALYSTS	541,836	5.758	6.516	0.000	22.000
ILLIQ	541,836	0.163	0.332	0.000	1.284
MOM	541,836	0.101	0.402	−0.533	1.049

investment opportunities as represented by low-cost passive funds (BvB). There are both theoretical and empirical reasons why this approach is more suitable than the traditionally used factor models, such as the Fama–French factor portfolios. First, factor portfolios are based on hypothetical stock portfolios and do not incorporate transaction costs, trade impact, and trading restrictions (Huij and Verbeek, 2009). Accordingly, they do not represent alternative investment opportunities. For example, investors do not have the opportunity to invest in momentum funds. From an empirical point of view, it is puzzling that index funds have positive alpha when their excess returns are regressed on the set of Fama–French factors. This could result in systematic biases in estimated fund alphas and thus lead to wrong inferences. Thus, we use a set of passive funds as the alternative investment opportunity set.

The benchmark-adjusted return of a fund's trades at any time is defined as the fund's trading return minus the closest return of the set of passive funds:

$$\alpha_{ft}^B = R_{ft} - \sum_{j=1}^{n(t)} \beta_f^j R_t^j, \tag{1}$$

where R_{ft} denotes the trading return of fund f in month t, R_t^j is the excess gross return earned by investors on the jth index fund at time t, and β_f^j is the sensitivity of fund f to the jth index fund. As reflected in the notation, the number of available benchmark funds may vary over time. To avoid a bias in selecting index funds, we follow BvB who select Vanguard index funds as benchmarks.[4] Vanguard funds are among the most popular passive investment opportunities and hence offer a reasonable representation of an investor's alternative investment opportunity set. We select passive funds offered by Vanguard in the following way. First, we select only equity funds and drop Morningstar Global Categories that span specific sectors of the stock market, such as technology and health care. Next, within each Global Category we select the oldest fund(s), offered in USD, that span all stocks in the category. We do not select funds from the Brazil Equity and Australia Equity Global Categories, as funds in those categories are not offered in USD and their coverage is already spanned by the EMs Equity category and the Asia-Pacific category, respectively. This selection procedure results in seven domestic US funds and six international funds. For US equity, we use the seven US funds. For international equity, we use the three Global Equity index funds. For European equity, we add the European Equity index fund. For Asia-Pacific equity, we add the Asia-Pacific Equity fund. Similarly, for EMs equity, we add the EMs equity fund. Due to geographical proximity, we further add the Asia-Pacific equity index fund to the alternative investment opportunity set for EM stocks from the Asia-Pacific region. For Canadian stocks, we add the S&P 500 index fund as a third passive alternative investment opportunity, due to geographical and economic proximity with the USA. The full list of benchmark funds is available in Panel B of Table II. Note that the resulting set of passive investment opportunities is very similar to that of BvB. Due to the international focus of our study, our alternative investment opportunity set includes more international funds. Importantly, there are no international benchmarks funds in our sample period with a distinctive regional value or momentum focus.[5]

---

4   See Section 5 and Table 1 in Berk and van Binsbergen (2015) for more details on their fund selection procedure.

5   Research suggests that global markets are not integrated (Fama and French, 2012) and risk premia are potentially driven by local factors (Griffin, 2002). Thus, the global Vanguard funds may not span

Table II. The aggregate performance of the stocks traded by active mutual funds—gross monthly alphas and monthly dollar value added

This table presents the performance of the aggregate trades of mutual funds in the international sample. We define buys (sales) as stocks with aggregate increases (decreases) in fractional holdings during quarter *t*. Next, we weigh stocks in the buys (sales) portfolio using aggregate volume bought (sold) during the quarter. Gross trading performance is calculated as the difference in performance between the buys and the sales. We track the excess return of the aggregate trading portfolio during the next 3 months and repeat the calculations. Next, we estimate the benchmark-adjusted trading performance using the Vanguard index funds as an alternative investment set. In Panel A, we report monthly alphas and aggregate dollar value added (in million USD) with standard errors in parentheses, separately for all, US, and non-US stocks as well as for all funds, US funds, and funds domiciled outside of the USA. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level. In Panel B, we report descriptive information about the set of Vanguard index funds used for estimating benchmark-adjusted returns in Panel A.

Panel A: Benchmark-adjusted performance of the aggregate trades

Domicile	Monthly gross alpha				Monthly dollar value added (in million USD)			
	All	USA	Non-USA	Difference	All	USA	Non-USA	Difference
All	−0.18**	−0.31***	−0.05	−0.26**	−1,193**	−917***	−276	−641**
	(0.09)	(0.11)	(0.09)	(0.11)	(466)	(309)	(254)	(321)
USA	−0.21*	−0.34**	0.03	−0.36**	−723**	−682**	−40	−642**
	(0.12)	(0.14)	(0.14)	(0.16)	(355)	(302)	(119)	(291)
Non-USA	−0.14*	−0.26***	−0.10	−0.16*	−354*	−179***	−174	−5
	(0.08)	(0.08)	(0.09)	(0.09)	(214)	(66)	(170)	(144)

Panel B: List of Vanguard benchmark funds

Fund name	ISIN	Asset class	Inception date
S&P 500 Index	US9229081081	US equity large cap blend	August 31, 1976
Value Index	US9229086783	US equity large cap value	November 13, 2000
Mid Cap Index	US9229088433	US equity mid cap blend	May 21, 1998
Extended Market Index	US9229082071	US equity mid cap growth	December 21, 1987
Small Cap Index	US9229087021	US equity small cap blend	October 3, 1960
Small Cap Growth Index	US9229088276	US equity small cap growth	May 21, 1998
Small Cap Value Index	US9229087930	US equity small cap value	May 21, 1998
Total Intl Stock Index	US9219096024	Global equity large cap blend	April 29, 1996
International value	US9219392035	Global equity large cap value	May 16, 1983
FTSE All-Wld ex-US SmCp Index	US9220427341	Global equity mid/Small cap blend	April 2, 2009
European Stock Index	US9220422052	Europe equity large cap blend	June 18, 1990
Pacific Stock Index	US9220421062	Asia equity large cap blend	June 18, 1990
Emerging Mkts Stock Index	US9220423043	Emerging markets equity large cap blend	May 4, 1994

all risk. Recently, index funds that track region-specific value and momentum risk premia have become available and thus should be considered as part of the set of benchmark traded funds available to current investors. For instance, Blackrock offers Europe-specific value and momentum ETFs since the beginning of 2015.

The benchmark loadings in (1) are estimated by regressing the fund's trading returns upon the relevant benchmark returns over the entire sample period that the fund is active. Here, we employ the benchmark funds' gross returns, defined as the reported net returns in Morningstar plus one-twelfth of the reported net annual expense ratio. Because one of the two global funds is not available throughout our sample period, we estimate betas by using an augmented basis of the factors where the factor returns are orthogonalized with respect to all other variables and missing returns are replaced with the mean of the orthogonalized factor. Alphas are then estimated by using the estimated betas and the augmented basis where we replace missing returns with zero.[6]

Our second approach is based on the comparison of every stock i with a set of stocks with similar size, book-to-market, and momentum characteristics [also known as DGTW-adjusted returns, following Daniel et al. (1997), Wermers (1999), and Wermers (2003), who introduced this methodology]. Specifically, the benchmark-adjusted return on a stock is given by

$$\alpha_{i,t}^{\mathrm{DGTW}} = R_{i,t} - R_{i,t}^{\mathrm{bench}}, \tag{2}$$

where $R_{i,t}^{\mathrm{bench}}$ denotes the return of a benchmark portfolio of stocks with similar size, book-to-market, and momentum characteristics. In Appendix E, we provide a detailed methodology for computing benchmark-adjusted returns for international stocks belonging to broad geographical regions, where we tackle a number of problems related to the size of equity markets and differences in accounting standards.[7] Where relevant, the stock level alphas from Equation (2) are aggregated to fund or industry level using the appropriate weights. The DGTW methodology offers several advantages. First, it identifies the closest benchmark for each individual asset traded and thus offers a relatively precise risk-adjustment. Second, calculated alphas are not affected by estimation error, which can be substantial during our relatively short sample period. Third, as they compare the local return of assets with the local return of a benchmark portfolio, DGTW returns are not affected by currency returns. On the negative side, the DGTW benchmark portfolio may not represent the actual investment opportunity set faced by fund managers, as they might be constrained in their trading, due to regulation, prohibitive trading costs, or other frictions.

Quantifying the impact of every possible investment constraint is a daunting task. To obtain some idea about the relevance of constraints due to frictions in international equity markets, we zoom into the holdings of the largest passively managed international fund in the Morningstar database—Vanguard Global Stock Index Fund. Because the fund is passively managed, it should ideally be able to closely mimic its benchmark, the MSCI World Index. However, potential frictions in financial markets should result in deviations from its benchmark portfolio. We collect index constituents from Morningstar and hand-match

---

6   The Appendix in Berk and van Binsbergen (2015) shows that alphas can be consistently estimated using this approach for dealing with missing passive index returns. Because the set of passive funds differs across equity markets, the augmented basis is calculated separately for European, Asia-Pacific, Canadian, EMs from Asia-Pacific, and other EMs equity.

7   The DGTW benchmark returns are available from the first author upon request.

them to Datastream and Worldscope.[8] We then construct the fund's Active Share in the spirit of Cremers and Petajisto (2009) which quantifies funds' deviations from the benchmark. According to Petajisto (2013), index funds keep their Active Share below 20%. The Active Share of Vanguard's fund stands at 17% at the beginning of our sample period drops to 10% in 2004 and remains at levels under 5% after 2005. Thus, any potential investment constraints in the first couple of years of our sample have quickly disappeared.

Mutual funds, however, may also constrain their investment universe based on geographical preferences or perceived information advantages. A large literature documents the tendency of investors to overweight geographically close assets, potentially because of the difficulty of acquiring information for distant stocks (e.g., Coval and Moskowitz, 1999) or because of cognitive biases (e.g., Graham, Harvey, and Huang, 2009). This "home-bias" is also the driver behind Vanguard's benchmark deviations in the early years of our sample.[9] Therefore, equities that are not within close geographical proximity may offer superior returns but will not be part of the investment opportunity set. For these reasons, the DGTW risk-adjustment methodology is a second choice to the alternative set of index funds.

As a robustness check, we also estimate alphas using traditional factor regressions. This standard approach computes alphas by subtracting the realized factor portfolio returns times the estimated fund factor sensitivities of a fund's excess returns. We consider the CAPM, the Fama–French three factor model, the Fama–French three factor plus momentum (Carhart, 1997), and the Fama–French five factor models, using, where relevant, international versions of the factor returns.

5. The Performance of Aggregate Mutual Fund Trades

5.1 Gross Alpha

Consistent with previous studies (e.g., Chen, Jegadeesh, and Wermers, 2000), we use changes in fractional holdings for classifying the aggregate buys and sales of mutual funds. For each stock at each point in time, fractional holdings are defined as the number of shares owned by funds in our sample relative to the total number of shares outstanding. We define stock i in quarter t as a buy (sale) if funds increased (decreased) their fractional holdings in that stock between quarters t and $t–1$. Consequently, the portfolio of aggregate buys (sales) of the actively managed equity funds consists of all stocks that experience an increase (decrease) in fractional holdings across two consecutive quarters. We weigh the stocks in the buys and sales portfolios using dollar volume traded. This way we give higher weight to stocks for which there is a stronger trading consensus among mutual funds, represented by the difference among the buying and selling volume in those stocks (the aggregate change in holdings times the price per share at the end of quarter $t–1$). We define trades as the difference between the buys and sales portfolios.

We track the subsequent returns of the trades portfolio and report its benchmark-adjusted performance in Table II. Overall, mutual fund trades worldwide have a poor trading record—the stocks they purchase underperform the stocks they sell by 0.18% per

---

8　We contacted MSCI to double-check the quality of Morningstar Data. MSCI sent us four monthly snapshots of the MSCI World Index constituents which we verified are identical to the constituents data provided by Morningstar.

9　The home bias is 13% in the beginning of the sample and decreases to below 1% after 2005.

month, after comparing their returns with the returns of the Vanguard index funds. Among US stocks, the poor trading record is even more pronounced and amounts to −0.31% per month. In the aggregate, trades among US stocks significantly underperform trades among non-US stocks. Among US domiciled funds, trades in the domestic stocks underperform trades among international stocks by 0.36% per month. Non-US funds also perform poorly among US stocks, but the difference in performance with respect to internationals stocks is weaker. In Section 7, we show that these findings are robust to using DGTW-adjusted returns and conventional factor regressions as well as alternative definitions of aggregate trades.

5.2 Dollar Value Added

The economic size of the aggregate trading performance can be further assessed using a dollar measure of value added. The dollar measure of performance is particularly useful in distinguishing skilled from unskilled fund managers. BvB show that in competitive markets, a fund with a small gross alpha but relatively large amount of dollar value added is more skilled than a fund with a relatively large gross alpha but small amount of dollar value added. We therefore follow BvB and quantify the amount of money added or destroyed by the trades of fund managers. In our study, the quarterly aggregate dollar value added is defined as the alpha on the funds' trading portfolio scaled by the dollar amount traded.

Time-series averages are reported in Panel B of Table II. Among US stocks, funds in the international sample destroy combined $1,193 million per month via their trades. This number corresponds to an average of $85,700 destroyed per fund per month. In contrast, BvB report that the average US fund adds $270,000 per month. There are, however, important differences between our studies. The focus of BvB is on total fund performance, whereas we study trading performance only. Thus, a likely explanation for our findings is that long-term fund holdings may capture fund value-adding decisions, whereas funds may destroy value using impatient trades. This view is consistent with Cremers and Pareek (2016) and Lan, Moneta, and Wermers (2018), who find that only fund managers with longer investment horizons are able to outperform the market. In addition, the industry may be beyond its optimal size and new dollars flowing into funds may end up in value-destroying trades. We examine this conjecture more thoroughly in the subsequent sections.

Similarly to the gross alpha findings in Panel A, US funds destroy significantly more value via trades in domestic stocks—an average of $682 million per month. Non-US funds, in contrast, destroy a combined $179 million per month.

5.3 Trading Costs

Data from Investment Technology Group[10] indicate that average round-trip commission and brokerage costs among international stocks range between 47 bps in the UK and 90 bps in Asia-Pacific emerging markets during the 2009 to 2014 period. Edelen, Evans, and Kadlec (2013) investigated the transaction costs among active US equity funds and find bid-ask spreads of similar order of magnitude to commission costs. Assuming a comparable relation among international stocks, a conservative estimate of the total round-trip transaction costs of active funds trading outside of the USA is at least 100 bps. Although an

---

10 See the company's Global Cost Review on https://www.itg.com/assets/ITG_Global-Cost-Review-2017Q2-Prelim-BrokerCostUpdated.pdf.

investigation of the net returns to investors in international markets is beyond the scope of our study, these returns are likely to be more similar to the net returns to investing in US stocks.

6. Has the Active Industry in the USA Become Too Large?

6.1 Active Industry Size

The US domestic market has witnessed a dramatic increase in the size of the fund industry. At the same time, the direct holdings by retail investors have shrunk by >50% in the past three decades (French, 2008). Such crowding of the investment management industry in the USA might have pronounced effects on the potential of fund managers to identify profitable opportunities for stock picking. For instance, Stein (2009) demonstrates that when too much capital from sophisticated investors is chasing the same opportunities, prices might deviate from fundamentals due to correlated trading. Related, Berk and Green (2004) and Pastor, Stambaugh, and Taylor (2015) show that increases in the fund industry can have a perverse impact of fund performance. Across different countries, Khorana, Servaes, and Tufano (2005) report an overall fraction of the market owned by funds that is much larger in the USA than the rest of the world, which is consistent with our data. As a result, the pessimistic picture of the crowded US equity market may not necessarily translate to international markets. Consistent with this conjecture, our results in the previous section document that the trading performance of active mutual funds is statistically lower among US relative to non-US stocks.

To further analyze this, we define active industry size in country (market) m as the total ownership of stocks in that market by all funds in our sample scaled by the total size of the market, that is,

$$\text{AIS}_{m,t} = \frac{\sum_i \text{Hold}_{i,t} \times \text{Price}_{i,t}}{\sum_i \text{SO}_{i,t} \times \text{Price}_{i,t}}, \tag{3}$$

where $\text{Hold}_{i,t}$ refers to active fund ownership (holdings) in stock i at time t, defined as the number of shares owned by all funds, $\text{SO}_{i,t}$ refers to total shares outstanding in stock i at time t, and where summations are taken over all stocks i in country m. Note that the size of the active fund industry is defined in terms of the country where investments take place (i.e., the market), not the country where the funds are domiciled.[11]

The average Active Industry Size (AIS) between 2001 and 2014 for the forty-two stock markets represented in our sample is provided in Table III. The fund industry is largest in the USA, where active funds from the international sample hold on average 13.2% of the market capitalization of all stocks. In the other countries, the size of the active industry amounts to on average 0.9–7.9%. The ownership of active funds is typically higher among developed markets and lower in emerging markets, with some exceptions. We also report Active Industry Size at the end of our sample period (2014). Most notably, the US fund industry has decreased from an average of 13.2–11.4%. The 2014 active industry size is higher than its mean in most emerging markets countries. Among developed markets, the fund industry in the UK has the highest growth of >2%. Growth in other countries is more moderate while some developed markets have even experienced a decrease. Further note

---

11 This is different from Ferreira et al. (2013), who explain fund performance from, among others, country characteristics related to a fund's domicile.

Table III. Average size of the active mutual fund industry across countries

This table presents average size of the active fund industry across investment countries between 2001 and 2014, based on the ownership of active funds in the international sample. Active industry size in a given market is defined as the total equity ownership by all active funds in that market, scaled by the combined market capitalization of all equities in that market.

Region	Country name	Industry size		Region	Country name	Industry size	
		Average	End of sample			Average	End of sample
North America				Asia Pacific			
	Canada	0.071	0.069		Australia	0.023	0.018
	USA	0.132	0.114		Hong Kong	0.030	0.036
Japan					New Zealand	0.019	0.015
	Japan	0.031	0.027		Singapore	0.040	0.038
					South Korea	0.062	0.058
Europe				Emerging markets			
	Austria	0.038	0.041		Brazil	0.010	0.022
	Belgium	0.030	0.042		Chile	0.011	0.020
	Denmark	0.048	0.071		China	0.011	0.015
	Finland	0.079	0.060		Colombia	0.004	0.010
	France	0.045	0.047		Czech Republic	0.028	0.028
	Germany	0.011	0.010		Hungary	0.052	0.070
	Greece	0.026	0.016		India	0.023	0.055
	Ireland	0.037	0.035		Indonesia	0.053	0.035
	Italy	0.029	0.038		Malaysia	0.025	0.027
	Luxembourg	0.019	0.027		Mexico	0.049	0.031
	The Netherlands	0.055	0.056		Peru	0.009	0.023
	Norway	0.055	0.034		Philippines	0.043	0.038
	Portugal	0.031	0.042		Poland	0.032	0.027
	Spain	0.026	0.027		Russia	0.016	0.038
	Sweden	0.054	0.054		Taiwan	0.057	0.066
	Switzerland	0.045	0.057		Thailand	0.033	0.041
	UK	0.062	0.088		Turkey	0.042	0.058

that the descriptive statistics reported in Table III are based on aggregation across the holdings of funds from the sixteen domiciles covered by our database and thus understate the amount of actively managed capital.

6.2 Theoretical Framework

In order to analyze whether the active industry in the USA has become too large, we need a theoretical model that relates performance to scale. Berk and Green (2004) and BvB propose a rational equilibrium framework that helps explain some well-known stylized facts of the active industry, such as the lack of return persistence and the predictability of fund flows. In the context of our study, the rational equilibrium has predictions for the effect of

the size of the industry on performance. Below we restate a basic version of the model of Berk and Green (2004) and BvB under neoclassical assumptions.

First, note that managers cannot infinitely scale positive NPV projects. In other words, as investors allocate money to successful funds, managers eventually run out of ideas and cannot generate extra alpha. In addition, as funds grow larger, their trades have growing impact on prices. Empirical evidence by Pastor, Stambaugh, and Taylor (2015) and Zhu (2018) provide ample support that funds do not operate under constant returns to scale. The literature establishes two related arguments why fund performance may suffer in a largely developed market, reflecting diseconomies of scale at either the fund or industry level. For instance, larger funds may run out of ideas or suffer from large price impact of their trades (Berk and Green, 2004). Alternatively, all funds in a relatively large fund industry may suffer from the fierce competition among them (Pastor and Stambaugh, 2012). Of course, the two arguments are closely related as a large fund industry can only arise if individual funds grow to be sufficiently large. To set the stage, assume that a fund's gross alpha a^g is decreasing in industry size:

$$\alpha^g = a - b\text{AIS}. \tag{4}$$

In this equation, $b > 0$ stands for diseconomies of scale and a corresponds to the gross alpha on the first dollar invested. In the original work of BvB, α^g is decreasing in fund size. However, because we are interested in the optimal industry size, we treat the aggregate industry as one fund. Thus, we assume returns are decreasing in the aggregate industry size.

Similarly to BvB and Berk and van Binsbergen (2017), we assume that managers maximize value-added V (AIS). In other words, their combined objective function maximizes the total dollar value extracted by the aggregate fund industry

$$V(\text{AIS}) = \text{AIS}\alpha^g = \text{AIS}(a - b\text{AIS}). \tag{5}$$

Taking first-order conditions with respect to the size of the active industry and setting it to zero produces

$$\text{AIS}^* = \frac{a}{2b}. \tag{6}$$

This implies the following maximum aggregate value added by the active industry (provided $a > 0$ and $b > 0$):

$$\alpha^g(\text{AIS}^*) = \frac{a^2}{4b}. \tag{7}$$

We can interpret the skill measure (7) as the upper bound of the dollar amount that the active industry can generate, relative to the total size of the market [see Equation (3)]. When markets are competitive and agents rational, investors allocate capital to funds with good past performance, as measured by net alpha. However, because projects are not infinitely scalable, managers cannot extract the same percentage return from financial markets. An equilibrium is reached when the industry has grown up to levels where net alpha going forward is zero.

Our focus is on the prediction of the optimal active industry size as given in Equation (6). Because managers' objective function is quadratic in the size of the industry, there is an optimal industry size that maximizes the total value added of the industry (provided $a > 0$ and $b > 0$). Beyond this optimal size, extra dollars cannot be put into productive use, which

could explain why in the aggregate funds destroy value via their trades. Consider an analogy with equity investments. Rational investors would bid the prices of undervalued stocks up until their returns going forward are zero on a risk-adjusted basis. However, if they bid the prices too high, then future returns would be negative. Similarly, rational investors would allocate capital to active funds as long as managers can generate value. Beyond the optimal point, investors would earn negative returns. In the next two subsections we give empirical content to these predictions.

6.3 Fund-Level Regressions: Estimating Diseconomies of Scale

In this subsection, we test empirically for the impact of scale on performance. We build on Pastor, Stambaugh, and Taylor (2015) and Zhu (2018) and estimate diseconomies of scale separately for USA and international markets. Consider a group of mutual funds, indexed $f = 1,\ldots,N$, which can invest in multiple markets $m = 1,2,\ldots,M$.[12] Denote the benchmark-adjusted return in month t of fund f in market m as r_{ft}^m. Denote the total market value of the fund at the end of the previous month as $q_{f,t-1}$ We then regress the benchmark-adjusted performance of mutual funds in a particular market on the size of the active industry in this market and the natural logarithm of the total size of the fund. That is,

$$r_{ft}^m = \alpha_f^m + \beta_1^m \, \text{AIS}_{m,t-1} + \beta_2^m \log q_{f,t-1} + \epsilon_{ft}^m. \tag{8}$$

In this equation, α_f^m captures unobserved market-specific managerial skill (which is assumed to be time-invariant). The coefficient $\beta_1^m < 0$ identifies decreasing returns at the industry level. Similarly, the coefficient $\beta_2^m < 0$ identifies decreasing returns to scale at the fund level. We include the natural logarithm of the total dollar value of the fund due to its robustness to outliers. The α_f^m are treated as fund-market-fixed effects, absorbing the cross-sectional variation in fund skill within a given market, and their inclusion is crucial for identifying the effect of $\log q_{f,t-1}$ on trading performance. We consider specifications where the dependent variable tracks either the total holdings or trading performance of a fund. The effect of diseconomies of scale is likely to be reflected in both.

A standard fixed effects estimator requires the regressors in Equation (8) to be strictly exogenous. That is, regressors should be uncorrelated with ϵ_{ft}^m across all time periods. As stressed by Pastor, Stambaugh, and Taylor (2015) this is not the case here, because (a) fund size mechanically relates to past performance (even without flows), and (b) investor flows respond to past performance. In addition, in our case, (c) funds may reallocate across markets depending upon past performance. To address this problem, we follow Pastor, Stambaugh, and Taylor (2015) and Zhu (2018) and first eliminate the fixed effects α_f^m by forward-demeaning Equation (8). The forward-demeaned version of a variable x is defined as

$$\bar{x}_{ft} = x_{ft} - \frac{1}{T_f - t + 1} \sum_{s=t}^{T_f} x_{fs}, \tag{9}$$

where T_f denotes the number of time periods for which fund f is observed. The coefficients in Equation (8) are then estimated by two-stage least squares (2SLS), employing instruments that are plausibly uncorrelated with the forward-demeaned error term. Pastor, Stambaugh, and Taylor (2015) propose to use backward-demeaned fund size as an

---

12   Note that not every fund needs to invest in every market.

instrument for forward-demeaned fund size, where the backward-demeaned version of a variable x is defined as

$$\underset{-}{x}_{f,t-1} = x_{f,t-1} - \frac{1}{t-1}\sum_{s=1}^{t-1} x_{f,s-1}. \tag{10}$$

We implement this by means of a 2SLS approach, where in a first stage a reduced form is estimated for the endogenous regressor, the fitted values of which are substituted into the forward-demeaned version of Equation (8) (without an intercept).

Zhu (2018) argues that, unlike Pastor, Stambaugh, and Taylor (2015), an intercept term should be included in the reduced forms, and we follow this recommendation. In addition, she advocates the use of lagged fund size $q_{f,t-1}$ as an instrument, because it is obviously correlated with the forward-demeaned lagged fund size and it is plausibly uncorrelated with the forward-demeaned error term. This instrument could be stronger if the fit of the first-stage regressions is improved.

We implement three versions of the recursive-demeaning 2SLS estimator. The first version follows Pastor, Stambaugh, and Taylor (2015) while allowing for an intercept term in the reduced form. We refer to this estimator as RD1. The second one extends Zhu (2018) and employs lagged fund size as an instrument for the forward-demeaned version. We refer to this estimator as RD2. Both estimators are expected to be (asymptotically) unbiased, their precision depending upon the relevance of the employed instruments. Simulation results in Zhu (2018) suggest that RD2 is more accurate than RD1. Given the availability of multiple instruments, it is natural to combine them into one estimator, which should be even more precise. We therefore also consider a third estimator that includes both the backward-demeaned and the lagged values of fund size as instruments. The resulting estimator, which is our preferred one, is referred to as RD3.[13] In order to minimize the impact of estimation error on our findings, we drop fund-market observations with <4 years of data. The specific steps to construct the three estimators are described in more detail in Appendix F.

The results from the diseconomies of scale regressions are summarized in Table IV. As results are consistent across the three estimators, we only report results using our preferred choice RD3. In Panel A, we focus on the holdings and trades among US stocks. In Specifications (1)–(6), we use the Vanguard funds as benchmarks. Our findings using funds' holdings returns are consistent with Pastor, Stambaugh, and Taylor (2015) and Zhu (2018), who find diseconomies of scale on the industry and fund level. Both the effect of fund size and active industry size are statistically negative when included together in Specification (3). The regressions using trading return as the dependent variable reveal a similar effect of the size of the industry on performance, and the magnitude of the estimated coefficient is larger. In contrast to the holdings-based regressions, fund size loses its statistical significance when included together with the active industry size, though it still points in the right direction.

In Panel B, we estimate the second-stage regressions jointly across all non-US stocks, while estimating the first-stage regression per market. In Specifications (1)–(6), where we use traded funds as benchmarks (BvB), the estimated coefficients of active industry size and log fund size are not statistically significant. We further estimate the second-stage

---

13 Note that our numbering of these estimators does not match the one in Zhu (2018).

Table IV. Regressions of fund alpha on active mutual fund industry size and fund size

This table presents the results of predictive regressions of monthly fund holding and trading returns of funds in the international sample, specific to a given market, on active industry size (defined as the total equity ownership by all active funds in that market, scaled by the combined market capitalization of all equities in that market) and log of fund size. Fund sizes are inflated to millions of 2014 US dollars using the value of all US stocks in our sample logs and scaled by 10^6 in order to make coefficients easier to read. The RD3 estimator used in these regressions is defined in Section 6.3. In Specification (1)–(6), we use the Vanguard-traded funds as benchmarks and in Specifications (7)–(12) we use DGTW-adjusted returns. In Panel A, we estimate the second-stage regressions separately for the US market, and in Panel B, we estimate the second-stage regressions jointly across all markets except for the US one. In Panels C–G, we estimate the second-stage regressions separately for stocks in the developed Asia-Pacific excl. Japan (APA), Canada (CAN), Emerging Markets (EME), developed Europe (EUR), and Japan (JAP), respectively. We report robust standard errors clustered on the fund and month level. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Benchmarks: traded funds						Benchmarks: DGTW					
	Holdings based			Trades based			Holdings based			Trades based		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: US market												
ActiveIndustrySize	−0.0727**		−0.0739**	−0.1404***		−0.1203***	−0.0299*		−0.0281	−0.1245**		−0.1068*
	(0.0338)		(0.0351)	(0.0459)		(0.0445)	(0.0177)		(0.0178)	(0.0609)		(0.0648)
Log(FundSize)		−0.0016**	−0.0015**		−0.0024**	−0.0020		−0.0007	−0.0006		−0.0023***	−0.0021**
		(0.0007)	(0.0008)		(0.0012)	(0.0013)		(0.0005)	(0.0005)		(0.0008)	(0.0009)
Observations	540,551	524,197	524,197	549,419	529,507	529,507	647,248	616,525	616,525	633,685	607,654	607,654
Panel B: Non-US market												
ActiveIndustrySize	−0.0067		−0.0257	0.0538		0.0474	−0.0146*		−0.0062	−0.1418***		−0.1361***
	(0.0725)		(0.0675)	(0.0639)		(0.0605)	(0.0080)		(0.0096)	(0.0474)		(0.0490)
Log(FundSize)		−0.0011	0.0006		−0.0025	−0.0003		−0.0004	−0.0008		−0.0012	−0.0014
		(0.0022)	(0.0030)		(0.0020)	(0.0026)		(0.0004)	(0.0005)		(0.0014)	(0.0017)
Observations	5,392,353	5,225,690	5,225,690	5,199,348	5,029,315	5,029,315	7,237,489	6,905,809	6,905,809	7,333,846	7,005,968	7,005,968

(continued)

Table IV. Continued

| | Benchmarks: traded funds | | | | | | Benchmarks: DGTW | | | | | |
| | Holdings based | | Trades based | | | | Holdings based | | | Trades based | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
						Panel C: APA market						
ActiveIndustrySize	−0.0840		−0.0611	−0.0960		−0.0894	−0.0047		−0.0015	−0.3692***		−0.3870**
	(0.0609)		(0.0952)	(0.1172)		(0.1459)	(0.0375)		(0.0402)	(0.1398)		(0.1630)
Log(FundSize)		−0.0000	0.0044		−0.0018	−0.0016		−0.0002	0.0047		−0.0030	0.0011
		(0.0023)	(0.0047)		(0.0029)	(0.0035)		(0.0024)	(0.0048)		(0.0020)	(0.0041)
Observations	720,455	699,171	699,171	691,043	669,510	669,510	944,142	903,795	903,795	960,122	919,076	919,076
						Panel D: CAN market						
ActiveIndustrySize	−0.0156		−0.0385	0.0890		0.1483	−0.0709		−0.0507	−0.1503		−0.1507
	(0.2210)		(0.2385)	(0.1532)		(0.1648)	(0.0517)		(0.0440)	(0.1810)		(0.2040)
Log(FundSize)		−0.0016	−0.0021		−0.0049*	−0.0058**		−0.0017	−0.0018		−0.0046*	−0.0049*
		(0.0034)	(0.0039)		(0.0027)	(0.0029)		(0.0034)	(0.0038)		(0.0025)	(0.0026)
Observations	329,166	320,198	320,198	335,319	325,633	325,633	425,710	408,876	408,876	417,755	401,101	401,101
						Panel E: EME market						
ActiveIndustrySize	−0.2212		−0.2687	−0.0151		−0.0275	−0.0610***		−0.0579***	−0.2572**		−0.2496**
	(0.1544)		(0.1701)	(0.0530)		(0.0720)	(0.0173)		(0.0210)	(0.1050)		(0.1197)

(continued)

Table IV. Continued

	Benchmarks: traded funds						Benchmarks: DGTW					
	Holdings based			Trades based			Holdings based			Trades based		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log(FundSize)		0.0043	0.0025		−0.0003	0.0004		−0.0008	−0.0003		0.0001	−0.0009
		(0.0037)	(0.0034)		(0.0036)	(0.0031)		(0.0008)	(0.0008)		(0.0032)	(0.0031)
Observations	940,613	915,553	915,544	896,948	871,421	871,421	1,380,712	1,325,325	1,325,325	1,397,625	1,340,898	1,340,898
						Panel F: EUR market						
ActiveIndustrySize	0.0501		0.0230	0.0781		0.0502	−0.0026		0.0026	−0.1071**		−0.1019*
	(0.0786)		(0.0712)	(0.0737)		(0.0667)	(0.0087)		(0.0094)	(0.0525)		(0.0559)
Log(FundSize)		−0.0033	0.0022		−0.0029	0.0025		−0.0066*	−0.0020		−0.0013	−0.0012
		(0.0047)	(0.0060)		(0.0039)	(0.0051)		(0.0038)	(0.0051)		(0.0020)	(0.0026)
Observations	3,143,602	3,040,963	3,040,963	3,021,010	2,917,730	2,917,730	4,167,421	3,969,328	3,969,328	4,242,430	4,043,954	4,043,954
						Panel G: JAP market						
ActiveIndustrySize	0.0112		0.1089	−0.1310		−0.0584	−0.0313		−0.0208	−0.1680		−0.0687
	(0.2752)		(0.2704)	(0.2589)		(0.2542)	(0.0301)		(0.0335)	(0.2666)		(0.2650)
Log(FundSize)		−0.0111**	−0.0109**		−0.0095**	−0.0098**		−0.0008	−0.0009		−0.0110***	−0.0113***
		(0.0051)	(0.0053)		(0.0040)	(0.0041)		(0.0010)	(0.0010)		(0.0040)	(0.0041)
Observations	258,517	249,805	249,814	255,028	245,021	245,021	319,504	298,485	298,485	315,914	300,939	300,939

regressions separately for each market and report these results in Panels C–G.[14] The US market stands out with significant estimates of diseconomies of scale. Within each of the other regions, the estimated coefficients are not statistically significant although they mostly point in the negative direction, as predicted by theory. There are a few potential explanations for the weaker evidence of diseconomies of scale in international stocks. First, the power of our tests might be low. Funds in our sample hold a relatively smaller fraction of their assets in international stocks, making it harder to estimate the diseconomies of scale parameters. Second, because of the relatively smaller presence of active funds, the overall fund industry outside of the USA might not be sufficiently large for the true impact of decreasing returns to scale to be revealed in our data. This could also explain why the aggregate trading performance of funds in our sample is better in international stocks than it is in domestic stocks.

Another concern is that, for markets other than the USA, the availability of low-cost region-specific benchmark funds is very limited over the sample period. Accordingly, for investments in non-US markets it would be relatively easy for fund managers to obtain positive alpha by having a non-zero value or momentum exposure, as the traded benchmarks may not be correcting for this. Related to this, it is likely that the benchmark returns soak up less variation outside the USA, and therefore result in low power of our tests.

With the above results in mind, in Specifications (7)–(12) we replace the benchmark-adjusted alphas with DGTW-adjusted returns. For the US market in Panel A, the results based on the DGTW-adjustment are very similar to those using traded benchmarks, but for non-US markets the changes in the coefficients for industry size (which determine the presence of diseconomies of scale), as well as their statistical significance, are substantial. Among the non-US stocks in Panel B, we find statistically significant impact of the active industry size on performance. The estimated coefficient using the holding-based regression in Specification (7), –0.0146, is about half of the estimate obtained for the USA (–0.0299). The trades-based coefficient in Equation (10) is larger and statistically even stronger than the one for the US market. As predicted by theory, (almost) all estimated coefficients on Active Industry Size are negative—and the many of them statistically significant—when we estimate the model per region (Panels C–G). Despite the fact that the estimated coefficients on Active Industry Size are negative, corresponding to decreasing returns to scale, there is considerable variation across specifications and across regions. For example, using the holdings-based returns it appears a bit more challenging to separate out the role of fund size and industry size [Specification (9)].

To estimate optimal industry size for regions other than the USA in the next subsection, we rely upon the estimation results based upon the DGTW-adjusted returns. Whereas there is very limited availability of passive funds that track value indices around the world and literally no momentum funds throughout sample period, during the last few years Vanguard (and other fund families) have started to offer passive funds that track regional value and momentum indices. Going forward, an investment set that includes passive exposures to region-specific value and momentum would better represent a relevant benchmark. Therefore, we the currently present diseconomies of scales may be better estimated with a characteristics-based benchmark that includes value and momentum.

---

14 We follow MSCI classification when grouping countries in broad regions.

6.4 Estimates of Optimal Industry Size

Our results in the previous sections raise the possibility that the active US industry has surpassed its optimal size. The rational equilibrium framework reveals that the optimal fund size is jointly determined by the alpha on the first dollar and the coefficient on decreasing return to scale [see Equation (6)]. If we assume that the industry acts like one fund, we can use Equation (6) to estimate an optimal industry size. If we look at this problem from another angle, if all funds in our sample have the same level of skill and are at their optimal size, then Equation (6) specifies the optimal size of the industry. Similarly to Zhu (2018), we back out the parameter *a* from the observed gross alpha, the estimated coefficient on diseconomies of scale, and the empirically observed size of the active industry:

$$\hat{a}_m = \frac{1}{T}\sum_{t=1}^{T}(\alpha_{mt}^g + \hat{b}_m \, \text{AIS}_{m,t}), \tag{11}$$

where $\hat{b}_m$ is taken to be the negative of the estimate of β_1^m from Equation (8) with log fund size omitted. For the US market, $\hat{b}_m$ is based on Specifications (1) and (7) from Table IV, using Vanguard funds and DGTW as benchmarks, respectively. We decide to work with the estimated coefficient based on the holdings-based returns as these are more closely connected to overall fund performance. To remain consistent, the gross alpha estimates are based on the same benchmarks as those underlying the estimation of $\hat{b}_m$.

Results are reported in Table V. Using the two alternative benchmarks to obtain gross alpha and $\hat{b}_m$, we find an optimal size of the US active industry of 6.8–7.4% of the overall stock market. Given the size of the US equity market in 2014, this corresponds to an optimum of nearly \$2.3 trillion. To put this in perspective, actively managed US-domiciled mutual funds in our sample manage \$3 trillion in domestic equity at the end of our sample period. In addition, there is \$0.5 trillion actively managed by funds domiciled outside the USA. This implies that in 2014, there is an excess of nearly \$1.2 trillion that is actively managed. The precision of our estimated optimal industry size is driven by the standard errors of both $\hat{b}_m$ and average gross alpha. We use this to derive standard errors for our estimates of the optimal industry size in Appendix G. Based on 95% confidence intervals, our findings indicate that the active industry in the USA has become significantly larger than its optimal size as the current size of the active industry is outside of the 95% confidence bounds.

For international markets, we use the pooled estimate of $\hat{b}_m$ from Specification (7) in Panel B of Table IV to reduce noise. The optimal active industry size varies between 2.9% and 7.9% across the five geographical areas. Relating this to the actual size of the active industry at the end of our sample period, it shows that for Canada and Japan, the actual size is roughly equal to its optimum. For Europe, the optimal size is roughly two-third of its actual size, similar to our findings for the USA. For emerging markets and the Asia-Pacific region, however, the actual industry size is only about half of the optimal size, suggesting that equilibrium forces, from the side of investors or fund managers, are likely to push further growth in these markets (and reduce net alpha going forward). In a rational expectations equilibrium, investors will chase investment opportunities with positive net present value, while fund managers, in the aggregate, will reallocate across geographical markets or adjust their degrees of active management across the globe, so as to maximize the total amount they extract from financial markets (Berk and van Binsbergen, 2017).

Of course, our estimates on the optimal size of the industry need to be interpreted with caution. As reflected in the 95% confidence intervals, our estimates of the optimal active industry size in international markets are relatively imprecise because of the difficulty of

Table V. Optimal active industry size around the world

This table presents the calculations of optimal industry size, separately for the USA, developed Asia-Pacific excl. Japan (APA), Canada (CAN), Emerging Markets (EME), developed Europe (EUR), and Japan (JPN) active mutual fund industries. In column Benchmarks, we indicate whether the estimated coefficient on diseconomies of scales $\hat{b}$ is from regressions using Vanguard Indices or DGTW as benchmarks. To remain consistent, the gross alpha estimates are estimated using the same benchmarks as $\hat{b}$. We first average alphas across all funds (using TNAs in equities in the given market as the weight) and then report time-series averages, expressed in percentages per month. Optimal industry size is estimated from Equation (6), where alpha on the first dollar is calculated according to Equation (11). The confidence intervals are determined using the variance of the optimal industry size estimator as defined in Appendix G. We further set a lower bound of the optimal industry size to 0. For comparison, we also provide the size of the active industry for each region at the end of our sample.

Market	Benchmarks:	Gross alpha (in % per month)	$\hat{b}$	Optimal active industry size			Active industry size (end of sample)
				Estimate	95% confidence interval		
USA	Vanguard funds	0.05	0.073	0.068	0.042	0.094	0.114
USA	DGTW	0.06	0.030	0.074	0.038	0.110	0.114
APA	DGTW	0.17	0.015	0.079	0.000	0.191	0.039
CAN	DGTW	0.07	0.015	0.060	0.000	0.243	0.069
EME	DGTW	0.14	0.015	0.067	0.000	0.206	0.038
EUR	DGTW	0.05	0.015	0.044	0.002	0.086	0.065
JPN	DGTW	0.04	0.015	0.029	0.000	0.095	0.027

accurately estimating gross alphas. In addition, they are sensitive to the functional form used to estimate diseconomies of scale. Moreover, by assuming that the whole industry acts like one fund, we oversimplify the nature of active investing. Yet, the findings that the active fund industry has grown beyond its optimal in the USA are consistent with the evidence of poor trading performance. Moreover, our findings are consistent with Leippold and Rueegg (2018) who find that with the possible exception of the US market, the Berk and Green equilibrium describes the data well. As our estimates provides only a first glance at the important issue of how large the active industry should be, we leave it for future research to provide a more thorough investigation on the best way to describe diseconomies of scale on the industry level and derive an optimal industry size.

6.5 Potential Reasons Why US Equity Markets Have Become More Crowded

There are a number of potential reasons why the US equity market have become more crowded over time. French (2008) documents the gradual displacing of retail investors with more sophisticated institutional investors over a few decades. This trend naturally leads to increased competition. Khandani and Lo (2011) document a drastic increase in the number of hedge funds involved in arbitrage activities on equity markets and the subsequent decline of the expected returns of typical quant equity strategies. Technological advances may allow high frequency traders to detect the informed orders of institutional investors and thus lower their expected profits (Menkveld, 2016). Next, the unconventional monetary policy of the FED lead pension funds to increase their allocations to equity markets
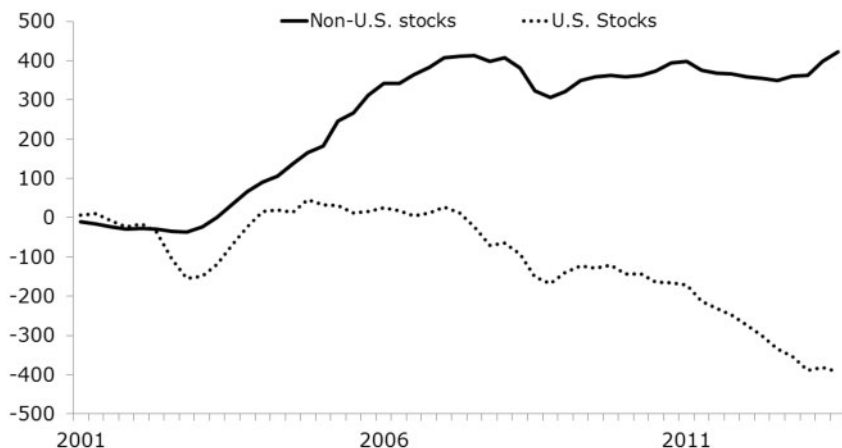
Figure 1. Cumulative capital allocation by active mutual funds domiciled in the USA. This figure presents the cumulative capital allocated by US funds in the international sample to non-US and US stocks. Numbers are expressed in \$bln.

(Boubaker *et al.*, 2017). Concurrent with these developments, the number of public companies in the USA has been falling.[15] As a result, the previous landscape of a limited amount of active money chasing attractive investment opportunities may have shifted toward one with an excessive amount of money chasing deteriorating investment opportunities.

In addition, there are at least a few regulatory changes that could have increased the crowdedness of financial markets. Agarwal *et al.* (2015) study the impact of a 2004 regulatory change that mandated more frequent portfolio disclosure. Their intuition is that when funds disclose more information, other market participants can trade on the same information and thus increase the competition informed funds face. Regulation Fair Disclosure (Reg FD) was promulgated in 2000 and limited the selective access to firm-specific information that mutual funds enjoyed at the time. This may lead to fewer information signals for mutual funds and thus increased competition. In line with this argument, Bhojraj, Cho, and Yehuda (2012) find that mutual fund performance decreased as a response to Reg FD. A couple of regulatory changes may have increased the execution costs of active funds and thus lowered the profits from their information signals. Bollen and Busse (2006) show that following decimalization, market depth declined and increased the trading costs of active funds while not affecting the trading costs of passive funds. Chung and Chuwonganant (2012) observe increases in trading costs following Regulation National Market System (NMS) too.

6.6 Capital Allocation Decisions

The rational expectations equilibrium framework has another interesting implication. If there is too much capital managed in the USA, we would expect funds to respond to the intense competition by diversifying across the rest of the world. In Figure 1, we report the

---

15 See, for example, Grullon, Larkin, and Michaely (2019) for broad evidence, and Doidge, Karolyi, and Stulz (2009, 2013) for cross-listings of international stocks on US exchanges and the number of domestic IPOs, respectively.

cumulative investments of US domiciled funds in crowded US stocks versus less crowded non-US stocks. During the period 2000–14, US funds have bought $400 billion in foreign stocks, while withdrawing a nearly identical amount from US equities. Because of this capital shift, the total assets under management of US funds among US equity has decreased from 91% to 71% during our sample period. Consistent with those findings, the size of the active industry in the US market at the end of our sample is lower than its sample mean (11% vs. 13%). These findings may indicate that the active US industry is declining in order to move closer to its optimal size.

6.7 Stock-Level Regressions: Industry Size and Herding

The diseconomies of scale regressions in Section 6.3 can establish patterns within markets, driven by time-series changes in the active industry size. In addition, the potentially detrimental impact of intense competition between fund managers can be identified cross-sectionally across and within markets. To investigate this, we regress quarterly stock returns on a measure of active fund trading (changes in fractional holdings, ΔFracHold), the size of the active fund industry in a given country (AIS), and their interaction. Results are reported in the first two columns of Table VI and include a wide range of controls. Following Gompers and Metrick (2001), we add lagged active fund ownership (FracHold) as a proxy for institutional demand. Standard errors are clustered at the stock level, though results remain consistent when we additionally cluster on the time dimension. All of our specifications include country-, time-, and industry-fixed effects. Hence, our regression analysis isolates the effect of trading by active funds on performance while taking into account any possible influences of country-level characteristics studied in previous research (e.g., Khorana, Servaes, and Tufano, 2005; Ferreira et al., 2013).

In Specification (1), we find that trading by active funds is statistically significantly associated with negative subsequent returns (t-statistic = −4.97). This confirms the central results of Panel A in Table II, while allowing us to control for a wide variety of other characteristics: trades correlate negatively to subsequent returns. Adding active industry size to the regression, and—most importantly—its interaction with changes in fractional holdings, allows us to explore how the relation between changes in holdings and subsequent stock returns varies across markets (countries) with different importance of the active fund industry, as well as over time [Specification (2)]. Consistent with our central hypothesis that the trading performance is poorer for markets that are more crowded, we find a significantly negative relationship with AIS interacted with changes in fractional holdings. The interaction term between changes in fractional holdings and active industry size enters the equation with a coefficient of −3.634 (t-statistic = −4.50). This implies that a one standard deviation increase in ownership by active funds in markets where funds hold only 1% of all assets is associated with 6 bp lower returns in the subsequent quarter. A similar in magnitude trading in markets where active funds own 5% of all assets leads to a subsequent reduction in performance of 29 bp. In Specification (2) we further find a positive coefficient on AIS. Thus, markets with a larger fund presence may still offer high investment returns, as long as funds engage in less trading. This finding is consistent with recent evidence by Cremers and Pareek (2016) and Lan, Moneta, and Wermers (2018), who show that more patient positions are characterized with positive abnormal returns. The message from these results is that trades correlate negatively to subsequent returns, and more so if the active industry size in a country is larger.

Table VI. Regressions of international quarterly stock returns on trading, active mutual fund industry size, and stock-level herding

This table presents the results of predictive regressions of quarterly stock returns on trading, active industry size, and stock-level herding, among the sample of international stocks during the period 2001–14. The dependent variable in each specification is stock return in local currency in quarter $t+1$. Depending on the specification, we include changes in fractional holdings by active funds (ΔFracHold) in quarter t; signed LSV measure taking the value of LSV if ΔFracHold is greater than zero and negative LSV otherwise; and active industry size (AIS), defined as total equity ownership by all active funds in that market, scaled by the combined market capitalization of all equities in that market. All specifications include industry-, country-, and time-fixed effects. Control variables are defined in Table I. All variables are winsorized at the 0.05% level and we divide the coefficient on PRICE by 1,000. We estimate coefficients using pooled regressions and report robust standard errors clustered on the stock level. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	(1)	(2)	(3)	(4)
ΔFracHold$_t$	−0.179***	0.159**	−0.166***	0.183**
	(0.036)	(0.080)	(0.040)	(0.083)
AIS * ΔFracHold$_t$		−3.634***		−3.702***
		(0.807)		(0.810)
AIS		0.220***		0.220***
		(0.021)		(0.021)
LSV^s			−0.008**	−0.008*
			(0.004)	(0.004)
FracHold$_{t-1}$	0.066***	0.062***	0.066***	0.062***
	(0.006)	(0.006)	(0.006)	(0.006)
BTM	0.007***	0.007***	0.007***	0.007***
	(0.000)	(0.000)	(0.000)	(0.000)
SIZE	−0.009***	−0.009***	−0.009***	−0.009***
	(0.000)	(0.000)	(0.000)	(0.000)
RET	0.009***	0.009***	0.009***	0.009***
	(0.002)	(0.002)	(0.002)	(0.002)
TURN	−0.036***	−0.035***	−0.036***	−0.035***
	(0.004)	(0.004)	(0.004)	(0.004)
VOL	0.026***	0.026***	0.026***	0.026***
	(0.002)	(0.002)	(0.002)	(0.002)
PRICE	−0.021	−0.016	−0.021	−0.016
	(0.014)	(0.014)	(0.014)	(0.014)
DY	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
ANALYSTS	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
ILLIQ	−0.013	−0.015	−0.013	−0.015
	(0.050)	(0.050)	(0.050)	(0.050)
MOM	0.024***	0.024***	0.024***	0.024***
	(0.001)	(0.001)	(0.001)	(0.001)
MSCI	0.008***	0.008***	0.008***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	793,981	793,981	793,981	793,981
R^2	0.154	0.154	0.154	0.154

The perverse effect of industry size is closely related to herding. Typically, herding is defined as the tendency of funds to follow the contemporaneous trades of their peers. We employ the widely used LSV herding measure, introduced by Lakonishok, Shleifer, and Vishny (1992). Herding is likely to be stronger in countries with larger active fund presence. However, in contrast to active industry size, LSV is a stock-level variable that allows us to exploit differences in stock level crowdedness and link it to subsequent returns. The LSV herding measure is based on the premise that if funds follow each other into and out of the same stocks over the same time interval, then funds would be primarily buyers or sellers of those stocks over that period. Specifically, the LSV herding measure for stock i in quarter t is defined as

$$LSV_{i,t} = |p_{i,t} - E_t(p_{i,t})| - E_{i,t}|p_{i,t} - E_t(p_{i,t})|, \tag{12}$$

where $p_{i,t}$ refers to the relative number of traders for stock i in quarter t, calculated as the number of funds buying stock i in quarter t divided by the sum of the number of funds buying stock i in quarter t and the number of funds selling stock i in quarter t. $E_t(p_{i,t})$ refers to the cross-sectional average of $p_{i,t}$ in quarter t. If institutions follow each other into and out of the same securities within the same quarter, $p_{i,t}$ will differ much from $E_t(p_{i,t})$ and the LSV herding measure will be positive. If, however, funds do not follow each other into and out of the same security within the same quarter, then $|p_{i,t} - E_t(p_{i,t})|$ tends to zero and consequently the LSV measure will be low. $E_{i,t}|p_{i,t} - E_t(p_{i,t})|$ is a stock-time specific adjustment factor which accounts for the fact that simply by chance, the number of buyers will be higher or lower than the number of sellers.[16] The LSV herding measure has been widely used in previous research (e.g., Grinblatt, Titman, and Wermers, 1995; Wei, Wermers, and Yao, 2014).

In Specification (3) of Table VI, we include a signed version of the LSV measure taking the value of LSV if ΔFracHold is greater than zero and negative LSV otherwise. This allows us to test whether stock level herding is negatively associated with subsequent returns. The results support this conjecture—the coefficient on signed LSV is negative and statistically different from zero, albeit marginally. In Specification (4) we include both LSV and the interaction between AIS and ΔFracHold. The size of the active industry appears to be a statistically stronger predictor of future returns than stock-level herding. In countries where the active fund industry is more important, their trading returns tend to be poorer.

7. Alternative Performance Measures and Evaluation Methods

7.1 Alternative Definitions of Aggregate Trades

We investigate the robustness of the trading performance using alternative definitions of the aggregate buys and sales portfolios. Results are reported in Table VII. First, we define buys (sales) as stocks with institutional demand higher (lower) than the cross-sectional average, where a stock's institutional demand is defined as the number of funds buying the stock relative to all funds trading the stock (in each direction). The patterns are similar and the economic magnitude and statistical significance of the findings is even stronger. For

16 Following LSV, the adjustment factor is calculated under the null hypothesis of no herding. To compute it, we assume that the number of institutional investors buying a security i follows a binomial distribution with probability $E_t(p_{i,t})$; see Lakonishok, Shleifer, and Vishny (1992) for further details.

Table VII. Gross alphas of the stocks traded by active mutual funds: alternative definitions of trades

This table presents the gross performance of the aggregate trades of mutual funds in the international sample using alternative definitions to trades. We define buys (sales) in two different ways: stocks where the number of funds buying the stock relative to all funds trading the stock is higher (lower) than the cross-sectional average during quarter *t* ("relative number of traders"); and stocks with an increase (decrease) in weight in the aggregate mutual fund portfolio during quarter *t* ("changes in aggregate Weight"). Next, we weigh stocks in the buys (sales) portfolio using aggregate volume bought (sold) during the quarter. Gross trading performance is calculated as the difference in performance between the buys and the sales. We track the excess return of the aggregate trading portfolio during the next 3 months and repeat the calculations. Next, we estimate the benchmark-adjusted trading performance using the Vanguard index funds as an alternative investment set. We report monthly alphas with standard errors in parentheses, separately for all, US, and non-US stocks as well as for all funds, US funds, and funds domiciled outside of the USA. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Domicile	Relative number of traders				Changes in aggregate weight			
	Stock location				Stock location			
	All	USA	Non-USA	Difference	All	US	Non-USA	Difference
All	0.12	−0.28***	0.54	−0.82**	0.09	−0.25	0.46	−0.72*
	(0.32)	(0.10)	(0.33)	(0.34)	(0.37)	(0.16)	(0.38)	(0.39)
USA	0.02	−0.31**	0.73	−1.04*	−0.05	−0.27	0.43	−0.69
	(0.37)	(0.12)	(0.58)	(0.59)	(0.42)	(0.18)	(0.64)	(0.65)
Non-USA	0.23	−0.21**	0.44	−0.64**	0.26	−0.22*	0.48	−0.70**
	(0.42)	(0.09)	(0.32)	(0.32)	(0.46)	(0.13)	(0.34)	(0.35)

instance, the monthly underperformance by all funds in our sample amounts to 0.28% per month in US stocks. Notably, however, the trading performance outside of the USA appears to be more positive than the one reported in Table II. In addition, we define buys (sales) as stocks with an increase (decrease) in weight in the aggregate holdings portfolio. The results are again largely consistent and similar in economic magnitude to the findings when trades are defined using changes in fractional holdings as in Table II, albeit with lower statistical significance.

7.2 Trading Performance Using DGTW-Adjusted Returns

In addition to benchmark-adjusted returns, we assess trading performance using DGTW-adjusted returns. We use this to establish the performance of the aggregate trades over different horizons. The results of the trading performance over quarterly and yearly holding horizons are reported in Table VIII. Overall, results in Panel A are consistent with those in Table II—the stocks funds purchase underperform the stocks they sell by 0.28% per quarter, after adjusting for size, book-to-market, and momentum. The underperformance, however, is statistically indistinguishable from zero. Among US stocks, the poor trading record is statistically significant and amounts to −0.61% per quarter. This negative trading return does not reverse over the course of the year and even increases to −1.90%. In the aggregate, trades among US stocks significantly underperform trades among non-US stocks. The

Table VIII. The gross alpha of the stocks traded by active mutual funds using DGTW returns

This table presents the gross performance of the aggregate trades of mutual funds in the international sample using DGTW-adjusted returns. We define buys (sales) in three different ways: stocks with aggregate increases (decreases) in fractional holdings during quarter t ("changes in fractional holdings," Panel A); stocks where the number of funds buying the stock relative to all funds trading the stock is higher (lower) than the cross-sectional average during quarter t ("relative number of traders," Panel B); and stocks with an increase (decrease) in weight in the aggregate mutual fund portfolio during quarter t ("changes in aggregate weight," Panel C). Next, we weigh stocks in the buys (sales) portfolio using aggregate volume bought (sold) during the quarter. Trading performance is calculated as the difference in performance between the buys and the sales. We track the risk-adjusted trading returns (DGTW returns) during the following one quarter and 1 year. We repeat the calculations every quarter and obtain a time-series of trading returns. We calculate aggregate trading returns separately for all, US, and non-US stocks as well as for all funds, US-domiciled funds, and funds domiciled outside of the USA. We report time-series means with standard errors in parentheses. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Domicile	Quaterly Dgtw Ret				Yearly Dgtw Ret			
	Stock location				Stock location			
	All	USA	Non-USA	Difference	All	USA	Non-USA	Difference
Panel A: Changes in fractional holdings								
All	−0.28	−0.61**	0.19	−0.80***	−1.10	−1.90**	0.12	−2.01***
	(0.29)	(0.29)	(0.38)	(0.29)	(1.02)	(0.82)	(1.07)	(0.64)
USA	−0.25	−0.48*	0.27	−0.75**	−1.02	−1.61**	0.49	−2.11***
	(0.25)	(0.29)	(0.34)	(0.33)	(0.80)	(0.77)	(0.77)	(0.73)
Non-USA	−0.40	−0.77**	−0.22	−0.55	−1.14	−1.69**	−0.85	−0.84
	(0.35)	(0.39)	(0.39)	(0.35)	(1.07)	(0.79)	(1.10)	(0.75)
Panel B: Relative number of traders								
All	−0.59*	−0.97***	0.24	−1.21***	−2.03*	−3.42***	0.08	−3.50***
	(0.31)	(0.29)	(0.38)	(0.34)	(1.14)	(0.86)	(1.22)	(0.76)
USA	−0.48*	−0.77**	0.01	−0.77*	−1.94**	−3.33***	0.23	−3.56***
	(0.25)	(0.30)	(0.34)	(0.40)	(0.89)	(0.89)	(0.85)	(0.95)
Non-USA	−0.50	−0.97***	−0.26	−0.71***	−1.31	−2.01***	−1.12	−0.88
	(0.38)	(0.34)	(0.40)	(0.27)	(1.14)	(0.73)	(1.16)	(0.86)
Panel C: Changes in aggregate weight								
All	−0.53	−0.83*	0.02	−0.85**	−1.50	−2.51**	0.12	−2.63***
	(0.44)	(0.45)	(0.50)	(0.34)	(1.42)	(1.22)	(1.40)	(0.72)
USA	−0.49	−0.77*	0.20	−0.98**	−1.10	−2.27*	0.46	−2.73***
	(0.35)	(0.41)	(0.39)	(0.38)	(0.96)	(1.17)	(0.96)	(0.81)
Non-USA	−0.63	−0.85*	−0.51	−0.33	−0.97	−1.55	−0.90	−0.66
	(0.52)	(0.52)	(0.54)	(0.35)	(1.43)	(1.08)	(1.39)	(0.83)

quarterly difference in returns between trades in US and non-US stocks amounts to –0.75% (*t*-statistic = –2.8), increasing to –2.01% (*t*-statistic = –3.1) after 1 year. Yet, the return among non-US trades does not significantly beat the benchmark portfolio of stocks with similar characteristics, both over the short and long term. Using an alternative definition of trades (Panels B and C), we report similar results. The economic magnitude and statistical significance of the findings appear to be even stronger.

7.3 Robustness to Factor Regressions

Since Hou, Karolyi, and Kho (2011) point that the international evidence on which factors and characteristics price stocks may differ from US findings, it is further important to examine the robustness of our findings to alternative performance measurements. In addition, there is uncertainty as to what asset model is used by investors to assess performance (see Barber, Huang, and Odean, 2016; Berk and van Binsbergen, 2016). Therefore, we obtain risk factors for the developed regions of North-America, Europe, Asia-Pacific, and Japan from Ken French's website. We estimate separately alphas for each region and then weigh them using average volume traded in those markets monthly estimated alphas are reported in Table IX and are largely consistent with the results in Table II. Alphas of the US trading portfolios based on changes in fractional holdings are typically significantly smaller than alphas of international trading portfolios. When momentum is included as an additional factor, however, performance among non-US stocks appears to be closer to that among US stocks. These results are consistent when gross alphas are defined using relative number of traders and changes in aggregate weight.

8. Results from the US Market Using Longer Time-Series

The poor trading performance among US stocks stands in stark contrast to earlier work by Chen, Jegadeesh, and Wermers (2000), who find that stocks funds buy outperform the stocks they sell. In order to reconcile our findings with previous research, we complement the recent international sample with a sample of domestic US funds that stretches back to 1980. The returns of the aggregate trades of domestic US equity mutual funds are summarized in Table X. Prior to 2000, we find results similar to those of Chen, Jegadeesh, and Wermers (2000)—the aggregate trading performance is positive and statistically different from zero both in the short and long run. In contrast, following 2000, funds lose money through trading. When gross alpha is defined via changes in fractional holdings, the difference in trading performance amounts to –1.62% (*t*-statistic =−3.2) in the subsequent quarter increasing to –2.51% (*t*-statistic =−2.5) in the year following trading. These reversals in performance are statistically and economically stronger when trading is defined using the relative number of traders and changes in aggregate weight. In unreported results, we find consistent results when performance is assessed using factor portfolios.

The longer time-series and richer stock- and fund-level data allow us to investigate this finding in more detail. This helps us not only to better understand the dramatic change in trading performance in the US sample, but also the cross-country differences in the more recent international sample. Consequently, we investigate the secular trend in the tendency of mutual funds to trade in herds as a possible driver for their deteriorating trading performance. We measure time-series trends in herding using average LSV as well as the intertemporal herding measure of Sias (2004). The LSV herding measure captures a temporal dimension of fund herding, that is, the tendency of funds to trade in the same direction as

Table IX. The performance of the stocks traded by active mutual funds—gross monthly alphas using factor regressions

This table presents the performance of the aggregate trades of mutual funds in the international sample, using factor models for performance measurement. We define buys (sales) in three different way: stocks with aggregate increases (decreases) in fractional holdings during quarter t ("changes in fractional holding"); stocks with an increase (decrease) in weight in the aggregate mutual fund portfolio during quarter t ("relative number of traders"); and stocks with an increase (decrease) in weight in the aggregate mutual fund portfolio during quarter t ("changes in aggregate weight"). We weigh stocks in the buys (sales) portfolio using aggregate volume bought (sold) during quarter t. Trading performance is calculated as the difference in performance between the buys and the sales, separately for the stock regions of North America, Europe, Asia-Pacific, and Japan. We track the raw returns of trades portfolio during each of the following 3 months. We repeat the calculations every quarter and obtain a time-series of trading returns. We calculate CAPM, Fama–French 3, Fama–French 3+Momentum, and Fama–French five-factor alphas using developed factor returns from Ken French's website. We calculate alphas separately for all, US, and non-US stocks as well as for all funds, funds domiciled in the USA, and funds domiciled outside of the USA. When stocks come from more than one region, we weigh alphas using average traded volume. We report monthly alphas with standard errors in parentheses. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

| | CAPM | | | FF3 | | | FF3+MOM | | | FF5 | | |
| | Stock location | | | Stock location | | | Stock location | | | Stock location | | |
	USA	Non-USA	Difference	USA	Non-USA	Difference	USA	Non-USA	Difference	USA	Non-USA	Difference
Domicile						Changes in fractional holdings						
All	−0.27	0.06	−0.32**	−0.39***	−0.06	−0.33**	−0.51***	−0.32**	−0.19	−0.62***	−0.31**	−0.32**
	(0.17)	(0.14)	(0.14)	(0.15)	(0.13)	(0.14)	(0.11)	(0.12)	(0.13)	(0.14)	(0.13)	(0.15)
USA	−0.24	0.07	−0.30**	−0.37***	−0.07	−0.30**	−0.46***	−0.20	−0.26*	−0.58***	−0.27**	−0.31**
	(0.16)	(0.14)	(0.14)	(0.13)	(0.12)	(0.14)	(0.11)	(0.13)	(0.13)	(0.13)	(0.12)	(0.15)
Non-USA	−0.34**	−0.09	−0.26*	−0.37**	−0.12	−0.25	−0.51***	−0.38***	−0.13	−0.57***	−0.32**	−0.25
	(0.16)	(0.11)	(0.16)	(0.16)	(0.11)	(0.15)	(0.11)	(0.10)	(0.14)	(0.16)	(0.13)	(0.17)

(continued)

Table IX. Continued

| | CAPM | | | FF3 | | | FF3 + MOM | | | FF5 | | |
| | Stock location | | | Stock location | | | Stock location | | | Stock location | | |
	USA	Non-USA	Difference	USA	Non-USA	Difference	USA	Non-USA	Difference	USA	Non-USA	Difference
Domicile												
				Relative number of traders								
All	−0.43***	−0.01	−0.42***	−0.56***	−0.12	−0.44***	−0.67***	−0.38***	−0.29**	−0.72***	−0.30**	−0.42***
	(0.15)	(0.13)	(0.15)	(0.13)	(0.12)	(0.14)	(0.10)	(0.10)	(0.13)	(0.14)	(0.12)	(0.15)
USA	−0.39***	−0.03	−0.36***	−0.53***	−0.17	−0.36**	−0.60***	−0.32***	−0.28**	−0.65***	−0.38***	−0.27*
	(0.14)	(0.12)	(0.14)	(0.12)	(0.11)	(0.14)	(0.10)	(0.11)	(0.14)	(0.12)	(0.10)	(0.14)
Non-USA	−0.41***	−0.09	−0.32**	−0.46***	−0.13	−0.32**	−0.58***	−0.39***	−0.20	−0.62***	−0.32***	−0.30*
	(0.16)	(0.10)	(0.14)	(0.16)	(0.10)	(0.14)	(0.12)	(0.08)	(0.14)	(0.16)	(0.11)	(0.16)
				Changes in aggregate weight								
All	−0.25	0.03	−0.29**	−0.38*	−0.10	−0.28**	−0.56***	−0.48***	−0.08	−0.68***	−0.40**	−0.28*
	(0.22)	(0.17)	(0.14)	(0.20)	(0.16)	(0.14)	(0.14)	(0.14)	(0.13)	(0.20)	(0.16)	(0.15)
USA	−0.27	0.05	−0.31**	−0.40**	−0.11	−0.30**	−0.56***	−0.32**	−0.24*	−0.68***	−0.37***	−0.31*
	(0.21)	(0.16)	(0.16)	(0.19)	(0.15)	(0.15)	(0.14)	(0.15)	(0.14)	(0.18)	(0.14)	(0.16)
Non-USA	−0.29	−0.13	−0.16	−0.35*	−0.21	−0.14	−0.57***	−0.58***	0.02	−0.68***	−0.46***	−0.22
	(0.21)	(0.14)	(0.16)	(0.21)	(0.14)	(0.16)	(0.12)	(0.11)	(0.13)	(0.21)	(0.16)	(0.17)

Table X. The performance of the stocks traded by active mutual funds in the US sample—quarterly and yearly DGTW returns

This table presents the performance of the aggregate trades of mutual funds in the US sample. We define buys (sales) in three different way: stocks with aggregate increases (decreases) in fractional holdings during quarter t ("changes in fractional holdings"); stocks where the number of funds buying the stock relative to all funds trading the stock is higher (lower) than the cross-sectional average during quarter t ("relative number of traders"); and stocks with an increase (decrease) in weight in the aggregate mutual fund portfolio during quarter t ("changes in aggregate weight"). We weigh stocks in the buys (sales) portfolio using aggregate volume bought (sold) during quarter t. Trading performance is calculated as the difference in performance between the buys and the sales. We track the risk-adjusted trading returns (DGTW returns) during the following one quarter. We repeat the calculations every quarter and obtain a time-series of trading returns. We calculate aggregate trading returns separately for the subperiods 1980–2000 and 2001–12. We report time-series means with standard errors in parentheses. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Quarterly DGTW Ret			Yearly DGTW Ret		
1980–2000	2001–12	Difference	1980–2000	2001–12	Difference
Changes in fractional holdings					
0.94***	–0.68*	–1.62***	1.33*	–1.17*	–2.51**
(0.33)	(0.38)	(0.51)	(0.73)	(0.67)	(0.99)
Relative number of traders					
1.09***	–0.95**	–2.04***	1.56*	–1.57*	–3.13***
(0.39)	(0.44)	(0.59)	(0.83)	(0.82)	(1.16)
Changes in aggregate weight					
0.90**	–1.27**	–2.17***	1.47*	–1.38	–2.85**
(0.35)	(0.55)	(0.65)	(0.80)	(1.01)	(1.29)

other funds during the same time period. On the other hand, the Sias (2004) captures the intertemporal dimension of fund herding, defined as the tendency of funds to trade in the same direction as other funds in the previous time interval. More specifically, the intertemporal herding measure is calculated as the estimated coefficient from a cross-sectional regression (using all stocks i in quarter t) of the relative number of traders on its lagged value:

$$p_{i,t} = \beta_t p_{i,t-1} + \epsilon_{i,t}, \tag{13}$$

where we standardize $p_{i,t}$ to have mean zero and unit variance in order to compare coefficients across time. The estimated coefficient β_t captures the tendency of funds to follow their trades across two consecutive quarters. Hence, if funds follow their trades, we would expect the estimated β_t coefficient to be positive.

We investigate secular trends in herding among mutual funds in the US sample in Table XI. In Panel A, we report the average LSV measure in the two time periods. The mean LSV scores in our study are consistent with previous research (e.g., Lakonishok, Shleifer, and Vishny, 1992; Grinblatt, Titman, and Wermers, 1995). On average, we find a

Table XI. Changes in herding over time among active mutual funds in the US sample

This table presents the results of the average cross-sectional herding and the extent of inter-temporal herding among funds in the US sample, separately for the subperiods of 1980–2000 and 2001–12 as well as for the difference between the two sub-periods ("Difference"). In Panel A, we report the average LSV (1992) measure across all stock quarters. In Panel B, Model 1, we present the results from the average cross-sectional regressions of standardized relative number of traders (RelNumTraders) in quarter $t+1$ on its lagged value in quarter t, where relative number of traders is defined as the fraction of funds buying a stock divided by the total number of funds trading that stock. We standardize the variable by subtracting the cross-sectional mean and dividing by the cross-sectional standard deviation. In Model 2, we decompose the estimated slope from Model 1 in the part that comes from funds following their own trades and the part that comes from funds following other funds' trades. Standard errors are reported in parentheses. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	1980–2000	2001–12	Difference
Panel A: Cross-sectional herding			
LSV mean	0.025***	0.031***	0.006***
	(0.00026)	(0.00029)	(0.00039)
Panel B: Intertemporal herding			
Model 1			
RelNumTraders	0.108***	0.231***	0.123***
	(0.008)	(0.020)	(0.022)
Model 2			
RelNumTraders^{own}	0.048***	0.079***	0.031***
	(0.006)	(0.011)	(0.012)
RelNumTraders^{others}	0.061***	0.152***	0.091***
	(0.006)	(0.013)	(0.014)

slightly positive LSV score of 0.025 prior to 2000. However, we find a significant increase in the average LSV score of 0.006 in the second half of the sample, indicating that temporal herding has increased over time. In Model 1 of Panel B, we present the average slope coefficients of Equation (13). Similarly to Sias (2004), we find that funds exhibit positive intertemporal herding—there is a positive correlation of the fraction of funds buying stock i in quarter t with the fraction of funds buying stock i in quarter t–1. This positive association is significantly stronger after 2000—the estimated slope coefficient in Equation (2) increases from 0.108 to 0.231 across the two periods. Thus, the increase in intertemporal herding shown in Panel B is consistent with the increase in cross-sectional herding documented in Panel A.

The slope coefficient in Equation (13) can further be decomposed in the part that comes from funds following their own trades and the part that comes from funds following other funds' trades.[17] The results, reported in Model 2 of Panel B, indicate that the increase in the average intertemporal herding is due to both increased tendency of funds to follow their own trades, as well as an increased tendency of funds to follow other funds' trades. However, most of the increase in herding is due to funds following each other—the part of

---

17 See Sias (2004) for the exact derivation of the two coefficients.

the average β_t that comes from institutions following other funds' trades increases from 0.061 to 0.152, while the component of the average β_t that comes from institutions following their own trades increases from 0.048 to 0.079.

In Table XII, we regress subsequent stock returns on trading and herding, utilizing the rich time-series of the US sample. All specifications include a wide range of controls as well as stock industry and time-fixed effects. In Specification (1), we find that overall during the 1980–2012 period, trades by active funds lead to subsequent positive returns. In Specification (2), we include the interaction of a dummy variable D2001, taking the value of 1 if the sample date is from year 2001 or later, with trading by active funds.[18] Consistent with the results in Table II, we find a negative coefficient of –0.133 (*t*-statistic =–4.2) on the interaction term. Thus, mutual fund trading in the US sample following 2000 destroys value: on average, a one standard deviation increase in institutional ownership is associated with 1.36% lower returns in the following quarter. In Specification (3), we find a negative coefficient of –0.034 (*t*-statistic =–4.8) on the interaction between D2001 and signed LSV. Thus, the trading losses following 2000 are stronger among stocks with more pronounced herding—stocks where funds herd more have significantly poorer performance in the subsequent quarter.

To be consistent with the international sample, Specifications (4) and (5) add an interaction of AIS with ΔFracHold. As this variable exhibits little independent variation over time for any given stock, it is not surprising to see that it enters the models insignificantly. The fixed effects and ΔFracHold alone explain >98% of its variation and thus leave very little variation in this interaction variable to explain stock returns. The coefficients for the interaction terms between D2001 and changes in fractional holdings, and between D2001 and signed LSV, are hardly affected in these final two columns.

9. Conclusion

In this paper, we study the investment skills of actively managed mutual funds from sixteen domicile countries investing in forty-two equity markets over the period 2001–14, in relation to returns to scale in the industry. In US equity mutual funds achieve particularly poor trading performance: after benchmark adjustment, the stocks they buy underperform those they sell by 0.31% per month (*t*-statistic =–2.8). In non-US equity, their trades perform better, achieving an insignificant gross monthly return of –0.04%. Exploring the investment environment for the mutual fund industry around the world, we find evidence of diseconomies of scale in the US equity market. Internationally, the statistical significance depends on the choice of market and also indicate the presence of diseconomies of scale. Importantly, mutual funds achieve particularly low returns when they trade equities in markets with a larger-scale active fund industry. This result suggests a link between negative trading performance and diseconomies of scale in active fund management.

Building upon the theoretical models of Berk and Green (2004) and BvB, we use these results to derive estimates of the optimal size of the active industry across different international regions. This allows us to address an important asset allocation question: how much money to allocate to indexing versus active investing across these regions. Keeping the limitations and imprecision of our estimates into account, there is strong evidence that

---

18 Note that the effects of D2001 in Specification (2)–(5) and AIS in Specifications (4) and (5) are subsumed by the time-fixed effects and hence not reported.

Table XII. Regressions of US quarterly stock returns on trading, active mutual fund industry size, and stock-level herding

This table presents the results of predictive regressions of quarterly stock returns on trading, active industry size, and stock-level herding, among the sample of US stocks during the period 1980–2012. The dependent variable in each specification is stock return in USD in quarter $t+1$. Depending on the specification, we include changes in fractional holdings by active funds (ΔFracHold) in quarter t; a dummy variable D2001 taking the value of 1 if the sample year is 2001 or higher and zero otherwise; signed LSV measure taking the value of LSV if ΔFracHold is greater than zero and negative LSV otherwise; and active industry size (AIS), defined as total equity ownership by all active funds in the US market, scaled by the combined market capitalization of all equities in that market. All specifications include industry- and time-fixed effects. Control variables are defined in Table I. All variables are winsorized at the 0.05% level and we divide the coefficient on PRICE by 1,000. We estimate coefficients using pooled regressions and report robust standard errors clustered on the stock level. * denotes significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	(1)	(2)	(3)	(4)	(5)
ΔFracHold$_t$	0.021**	0.033*	0.026**	0.074	−0.014
	(0.009)	(0.017)	(0.013)	(0.054)	(0.055)
D2001 * ΔFracHold$_t$		−0.133***	−0.116***	−0.131***	−0.118***
		(0.032)	(0.030)	(0.032)	(0.032)
LSV^s			0.027***		0.027***
			(0.004)		(0.004)
D2001 * LSV^s			−0.034***		−0.035***
			(0.007)		(0.007)
AIS * ΔFracHold$_t$				−0.271	0.262
				(0.401)	(0.389)
FracHold$_{t-1}$	0.002	0.004	0.003	0.004	0.003
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
BTM	0.018***	0.018***	0.018***	0.018***	0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
SIZE	−0.024***	−0.024***	−0.024***	−0.024***	−0.024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
RET	0.039***	0.039***	0.039***	0.039***	0.039***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
TURN	−0.050***	−0.050***	−0.050***	−0.050***	−0.050***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
VOL	−0.008***	−0.008***	−0.008***	−0.008***	−0.008***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
PRICE	0.356***	0.354***	0.353***	0.354***	0.353***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
DY	−0.002*	−0.002*	−0.002*	−0.002*	−0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ANALYSTS	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ILLIQ	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MOM	0.033***	0.033***	0.033***	0.033***	0.033***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	541,836	541,836	541,836	541,836	541,836
R^2	0.152	0.152	0.152	0.152	0.152

for the US market, the active fund industry has surpassed its optimal size. While we also find decreasing returns to scale outside the USA, it does not appear that the active industry has reached its optimal size, perhaps with the exception of Europe. This suggests that fund managers, in the aggregate, would increase the amount of money they extract from financial markets by decreasing their level of active trading in the USA and reallocating their investments toward international markets. In a similar spirit, rational investors are expected to learn about international funds and their performance, increasing their flows to funds that invest in international markets where the active industry size is still relatively small. In the future, we plan to investigate psychological and institutional forces that might inhibit fuller international diversification and hamper the move toward the long-run equilibrium.

Appendix A

Overview Construction of the International and US Samples

In this Appendix, we provide an overview of the construction of the International and US samples used in this study. In Appendices B–E, we provide a more detailed overview of the stock selection procedures and the methodology for constructing characteristic-adjusted portfolios for the International Sample.

A.1 International Sample

The international sample builds upon the portfolio holdings data available in Factset. Factset currently provides the reported holdings of >90,000 funds located in eighty-nine domiciles. The database covers active and passive mutual funds, insurances, pension funds, and other funds and includes both alive and defunct funds. Similarly to Thomson Reuters, Factset acquires the quarterly positions of US funds via forms N-Q and N-CSR, available on SEC's EDGAR system, or directly from active management companies. In some countries, such as Spain and Sweden, fund holdings data are provided by the regulatory authority of the country or the mutual fund association. Positions of Canadian funds are obtained from the Interim and Annual Financial Statements on Canada's SEDAR system. For other countries, Factset obtains portfolio holdings data from the websites of, or from communications with, the respective asset management companies.

Similar to Chuprinn, Massa, Schumacher (2015), we exclude fund reports before 2001 because the coverage of Factset prior to this year is limited. Following Elton, Gruber, and Blake (2001) and Chen et al. (2004), we exclude funds with net assets of <15 million USD, as their data are potentially biased. Next, we drop any fund report in which a single security constitutes >25% of the total assets of the fund. We further drop index funds and select funds classified by Factset as either open-ended or offshore. Note that Factset classifies funds from Luxembourg and Ireland as offshore rather than open-ended funds. Thus, this selection criterion ensures our sample covers funds from those major fund domiciles. Next, we keep only funds that hold at least 50 stock holdings (equities and/or depository receipts) in their portfolios. This way we capture only funds with active equity components and exclude funds that may hold equities for diversification purposes only. This procedure also ensures we drop funds from countries with lax portfolio reporting regulations, such as Australia where funds are required to report only their top 10 portfolio holdings. Some portfolio reports contain likely data errors because the TNA value appears to bounce back

close to its original value after a spike in either direction. We exclude such reports. Specifically, we exclude fund portfolio reports across two periods where reported TNAs increase/decrease by a factor of >9 (quarter $q–1$ vs. quarter q) and which is subsequently reversed by a factor of at least 4.5 in the opposite direction (quarter q vs. quarter $q+1$) while the increase/decrease across both periods does not exceed 4.5 in the original direction (quarter $q+1$ vs. quarter $q–1$).[19] We further include fund reports only if the same fund has another report available in one of the previous two quarters in order to be able to calculate changes in holdings. This way we exclude domiciles with infrequent portfolio disclosure, for which it is hard to approximate trading decisions. For example, most Singaporean funds in Factset report only once a year and are therefore dropped. As a result of this choice, for 80% of the portfolio holding reports of funds, there exists a portfolio holding report in the previous quarter. For the remaining 20%, we use a lagged portfolio report that is two quarters old. In some cases, there exist more than one report per quarter or the report does not refer to end-of-quarter positions (i.e., February rather than March). In such cases, we always choose the portfolio snapshot closest to the end of the quarter and use the reported holdings as if they were reported at the end of the quarter. Because coverage of some countries may be scarce, we only include domiciles with at least 20 funds present during at least 75% of the time. For example, this data selection procedures result in the exclusion of funds from China, which are present in Factset only between 2008 and 2011, and funds from Japan, which appear to be a very small number in Factset (most years <20). This yields 322,628 unique fund-quarterly report date observations.

We match the reported fund holdings with stock specific information from Worldscope and Datastream using CUSIP, ISIN, and SEDOL identifiers. Because idiosyncratic shocks in stocks from these countries are potentially not easy to diversify, we drop countries that are not members of the Standard & Poor's BMI indices for developed and emerging markets (Europe, the Americas, Japan, and the Asia Pacific). Moreover, fund holdings in such stocks are negligible and data quality is likely to be low. We further follow the data-cleaning procedures prescribed in Ince and Porter (2006), Schmidt et al. (2011), and Dyakov and Wipplinger (2020). Specifically, we exclude (a) stock issues with $>20\%$ difference in market capitalization between Datastream and Factset, (b) stocks where a single fund is reported to own $>25\%$ of the shares, and (c) stocks with some key missing information in either Factset or Datastream.[20] We also note a potentially spurious pattern among some of the stock holdings: Some funds increase their reported holdings by a factor of, for example, 100 only to decrease their holdings by a similar factor in the next reporting period. Such changes are apparent data errors and we exclude them using the same screen for individual holdings as the one used for large reversals of total reported assets mentioned above.

A detailed overview on the stock-level country selection, the merging of Factset with Datastream, and the cleaning of stock information from Datastream is available in Appendices B–E.

---

19 For example, suppose that this quarter TNA is \$9 mln. If the previous quarter TNA was \$1 mln, then the increase was by a factor of 9. If next quarter TNA dropped to \$2 mln, then the subsequent drop is by a factor of 4.5. Between the previous and next quarter, however, the increase is only by a factor of 2 (\$2 mln vs. \$1 mln). Hence, the TNA "bounces" from \$1 mln up to \$9 mln and then back to \$2 mln and the portfolio snapshot is excluded.

20 Stocks with key missing information from Datastream and Worldscore are also not included in the benchmark portfolios and also not part of the return-predictive regressions.

A.2 US Sample

The US sample is built upon the Thomson Financial/CDA database which covers quarterly/semi-annual holdings of mutual funds, as reported to the SEC or voluntarily reported by the funds. We select funds with an investment objective code of growth, aggressive growth, and growth and income. We further exclude all index funds by deleting funds that have the strings INDEX, INDE, INDX, S&P, or MSCI in their names. We link the Thomson Financial/CDA database to the CRSP Mutual Fund Database using the MFLINKS tool provided by WRDS. The final dataset covers funds included in both mutual fund databases, for which we have two consecutive quarterly (or semi-annual) reports in Thomson Financial/CDA. Since most actively managed US equity funds offer different share classes to investors, we sum the net assets over different share classes and take asset-weighted share class averages of different attributes such as returns and expense ratios.

Appendix B: Region Assignment

This Appendix outlines the selection of equity markets, part of the International Sample. Note that this is different from the selection of fund domiciles, outlined in Appendix A. Specifically, we select countries from Europe, the Americas, and the Asia Pacific region for which Worldscope Constituent Lists are available in Datastream. We further restrict our sample to countries which are also members of the Standard & Poor's BMI indices for developed (DEV) and emerging markets (EMs). Next, we assign countries to one of the following seven regions, based on broad geographical location and level of development:

- Developed North America (NAM)
- Developed Europe (EUR)
- Japan (JPN)
- Developed Asia Pacific (APA)
- Emerging Europe (EM_EUR)
- Emerging Asia Pacific region (EM_APA)
- Emerging Latin America (EM_LAM)

Table BI lists the selected forty-two countries and their, respectively, assigned regions. In Appendices C–E, we use the region selection in order to assign stocks to benchmark portfolios and compute characteristic-adjusted returns.

Appendix C: Stock Selection

In this Appendix, we provide an overview of the selection of stocks used in the International Sample. The price and balance sheet data of all stocks part of the international sample stems from Datastream and Worldscope. The methodology builds upon Ince and Porter (2006) and Schmidt et al. (2011) and closely follows Dyakov and Wipplinger (2020). Further note that mutual funds, part of the Factset database, do not hold every possible stock in the universe of forty-two equity markets. However, we collect company level and main issue level information for every stock in that universe because of (a) the calculation of benchmark portfolios and (b) the regressions of stock returns on changes in fractional holdings, both of which require information on companies not part of the sample of stocks held by mutual funds.

Table BI. Stock selection

This table present the inputs for the stock-level selection procedures, part of the International Sample. Region refers the developed/emerging geographic regions to which countries are assigned. Developed regions include North-America (NAM), Europe (EUR), Japan (JPN), Asia-Pacific ex Japan (APA). EMs include Emerging Europe (EM_EUR), Emerging Asia-Pacific (EM_APA), Latin America (EM_LAM). ISO refers to the three digit Country Code with the currency in which Worldscope reports balance sheet data in parenthesis (CUR), WSNationCode is a numeric country code from Worldscope, LEGOR refers to the Legal Origin of a country from Djankov et al. (2008): (E)nglish, (F)rench, (G)erman, and (S)candinavian. We also report the Worldscore and Research and Dead Constituent Lists used to retrieve the initial list of Datastream Codes.

Region	Country	ISO (CUR)	WS-Code	LEGOR	Worldscope constituent lists	Research and dead constituent lists	Size-BTM-Mom.	Selection
NAM	Canada	CAN (CAD)	124	E	WSCN1-2-3-4-5-6, WSCOPECN	DEADCN1-2, FCDNX, FCDNX1-2-3-4-5, FTORO, FTORO1, FTORO2, FTORO3, FTORO4, FTORO5-U, FVANC, FVANC1, FVANC2, FVANC3, FVANC4, FVANC5	5-5-5	1
NAM	USA	USA (USD)	840	E	WSUS1-2-3-4-5-6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-21-22	DEADUS, DEADUS1-2-3-4-5-6, FAMERA-B-C-D-E-F-G-H-I-J-K-L-M-N-O-P-Q-R-S-T-U-V-W-X-Y-Z, FUSAA-B-C-D-E-F-G	5-5-5	1
EUR	Austria	AUT (EUR)	40	G	WSCOPEOE	DEADOE, FOST	5-5-5	1
EUR	Belgium	BEL (EUR)	56	F	WSCOPEBG	DEADBG, FBDO, FBEL	5-5-5	1
EUR	Denmark	DNK (DKK)	208	S	WSCOPEDK	DEADDK, FDEN	5-5-5	1
EUR	Finland	FIN (EUR)	246	S	WSCOPEFN	DEADFN, FFIN, FNALL	5-5-5	1
EUR	France	FRA (EUR)	250	F	WSCOPEFR, WSFR1, WSFR2	DEADFR, FFDOM, FFRA, FFRA1, FFRA2, FFRA3	5-5-5	1
EUR	Germany	DEU (EUR)	280	G	WSBD1, WSBD2, WSCOPEBD	DEADBD1-2-3-4-5-6, FGER1, FGER2, FGERDOM, FGERINT, FGFO1-2-3-4-5-6-7-8	5-5-5	1
EUR	Greece	GRC (EUR)	300	F	WSCOPEGR	DEADGR, FGREE	5-5-5	1
EUR	Ireland	IRL (EUR)	372	E	WSCOPEIR	DEADIR, FIRL	5-5-5	1
EUR	Italy	ITA (EUR)	380	F	WSCOPEIT	DEADIT, FITA	5-5-5	1

(continued)

Table BI. Continued

Region	Country	ISO (CUR)	WS-Code	LEGOR	Worldscope constituent lists	Research and dead constituent lists	Size-BTM-Mom.	Selection
EUR	Luxembourg	LUX (EUR)	442	F	WSCOPELX	DEADLX, FLUX	5-5-5	1
EUR	The Netherlands	NLD (EUR)	528	F	WSCOPENL	DEADNL, FHOL	5-5-5	1
EUR	Norway	NOR (NOK)	578	S	WSCOPENW	DEADNW, FNOR	5-5-5	1
EUR	Portugal	PRT (EUR)	620	F	WSCOPEPT	DEADPT, FPOR, FPSM	5-5-5	1
EUR	Spain	ESP (EUR)	724	F	WSCOPEES	DEADES, FSPN	5-5-5	1
EUR	Sweden	SWE (SEK)	752	S	WSCOPESD	DEADSD, FSWD	5-5-5	1
EUR	Switzerland	CHE (CHF)	756	G	WSCOPESW	DEADSW, FSWA, FSWA1, FSWA2, FSWS	5-5-5	1
EUR	UK	GBR (GBP)	826	E	WSCOPEUK, WSUK1-2-3-4-5-6	DEADUK, DEADUK1-2-3-4-5-6-7, FBRIT	5-5-5	1
JPN	Japan	JPN (JPY)	392	G	WSCOPEJP, WSJP1-2-3-4-5-6	DEADJP, FJAPA-B-C, FJAP2A, FJAP2B, FJAPAIM, FNAGOYA, FOSAKA, FQBOARD, FSAPPORO, FTOKYO, FUKUOKA	4-4-4	1
APA	Australia	AUS (AUD)	36	E	WSAU1, WSAU2, WSAU3, WSCOPEAU	DEADAU, FAUS, FAUSA, FAUSB	3-3-3	1
APA	Hong Kong	HKG (HKD)	344	E	WSCOPEHK, WSHK1, WSHK2	DEADHK, FHK1, FHK2, FHKQ	3-3-3	1
APA	New Zealand	NZL (NZD)	554	E	WSCOPENZ	DEADNZ, FNWZ	3-3-3	1
APA	Singapore	SGP (SGD)	702	E	WSCOPESG	DEADSG, FSIN, FSINQ	3-3-3	1
APA	South Korea	KOR (KRW)	410	G	WSCOPEKO, WSKO1, WSKO2, WSKO3	DEADKO, FKOR, FKOR1, FKOR2	3-3-3	1
EM_EUR	Czech Republic	CZE (CZK)	203	G	WSCOPECZ	DEADCZ, FCZECH, FCZECH1	2-2-2	2
EM_EUR	Hungary	HUN (HUF)	350	G	WSCOPEHN	DEADHU, FHUN	2-2-2	2
EM_EUR	Poland	POL (PLN)	617	G	WSCOPEPO	DEADPO, FPOL	2-2-2	2
EM_EUR	Russia	RUS (RUB)	643	F	WSCOPERS	DEADRU, FRUS	2-2-2	2

(continued)

Table BI. Continued

Region	Country	ISO (CUR)	WS-Code	LEGOR	Worldscope constituent lists	Research and dead constituent lists	Size-BTM-Mom.	Selection
EM_EUR	Turkey	TUR (TRY)	796	F	WSCOPETK	DEADTK, FTURK	2-2-2	2
EM_APA	China	CHN (CNY)	156	G	WSCH1-2-3-4, WSCOPECH	DEADCH, FCHINA, FCHSHA, FCHSHB, FCHSZA, FCHSZB, FCHSZCN	3-3-3	2
EM_APA	India	IND (INR)	356	E	WSCOPEIN, WSIN1, WSIN2, WSIN3	DEADIND, FINDIA	3-3-3	2
EM_APA	Indonesia	IDN (IDR)	366	F	WSCOPEID	DEADID, DEADIDN, FINO	3-3-3	2
EM_APA	Malaysia	MYS (MYR)	458	E	WSCOPEMY, WSMY1, WSMY2	DEADMY, FMAL, FMALF, FMALQ1, FMALQ2	3-3-3	2
EM_APA	Philippines	PHL (PHP)	608	F	WSCOPEPH	DEADPH, FPHI	3-3-3	2
EM_APA	Taiwan	TWN (TWD)	760	G	WSCOPETA, WSTA1, WSTA2, WSTA3	DEADTW, FTAIQ	3-3-3	2
EM_APA	Thailand	THA (THB)	764	E	WSCOPETH	DEADTH, FNVDRTH, FNVTH, FTHA, FTHAF, FTHAQ, FTHAQ1, FTHAQ2	3-3-3	2
EM_LAM	Brazil	BRA (BRL)	76	F	WSCOPEBR	DEADBRA, FBRA, FBRA1, FBRA2	2-2-2	2
EM_LAM	Chile	CHL (CLP)	152	F	WSCOPECL	DEADCHI, FCHILE	2-2-2	2
EM_LAM	Colombia	COL (COP)	175	F	WSCOPECB	DEADCO, FCOL, FCOLO, FCOLBG	2-2-2	2
EM_LAM	Mexico	MEX (MXN)	484	F	WSCOPEMX	DEADME, FMEX, FMEXFO	2-2-2	2
EM_LAM	Peru	PER (PEN)	597	F	WSCOPEPE	DEADPE, FPERU, FPERUC, FPERUT	2-2-2	2

C.1 Company Issue Selection

The stock-level analysis in the paper is carried on the company level. Since companies have multiple issues, we carry a number of selection procedures in order to obtain a single primary security for every company in the forty-two equity markets.

First, we select candidate Datastream Codes (DSCD) from the Research, Dead, and Worldscope constituent lists from Datastream in February 2015 for the countries selected in Appendix B. The download order is (1) Worldscope, (2) Research lists, and (3) Dead lists. This way we avoid missing defunct codes that may been moved to the dead lists during the download time.

Next, we perform the following data-selection procedures:

1. We select issues at the end of June from 1999 to 2014 which are either listed in the Worldscope lists of Datastream or listed as major securities in Datastream (Datastream item MAJOR = "Y")[21];
2. Furthermore, we consider only issues that refer to equities (Datastream item TYPE = "EQ") and are neither depository receipts nor preferred equity; issues that have at least 2 years of balance sheet data available in Wordscope; have a 13-month return at the end of June that is not zero;
3. We require that the stock is assigned to an Industry Classification Benchmark (ICB) industry subsector (Datastream item ICBSUC) and that Datastream contains information on the exchange (Datastream item ISOMIC);
4. We remove penny stocks by requiring that the unadjusted stock price is >1 in local currency (Datastream item UP >1).
5. Last, we consider only issues where the exchange is either in the same region as the Worldscope country of the stock or it is both the primary and major listing according to Datastream (Datastream items ISINID = "P" and MAJOR = "Y").

This selection is not sufficient to obtain a single primary security for each company at all times because Datastream may report multiple listings and share classes for a company. In order to select a primary issue among the issues selected above at each point in time, we rank the Datastream equity listings (Datastream identifier DSCD) for each Worldscope permanent identifier (Worldscope item 06105): We prioritize securities in the following order:

1. Securities classified as major securities in Datastream (Datastream item MAJOR = "Y");
2. Securities that are contained in the Worldscope lists (see Table BI);
3. Securities with non-zero trading volume in the prior month (Datastream item UVO > 0);
4. Securities classified as primary listing for the corresponding International Securities Identification Number (ISIN, Datastream item ISINID = "P").

For each Worldscope identifier, we select only the top-ranked Datastream identifier at each point in time and repeat the identifier ranking procedure every month. Table BI provides technical details on the Research, Dead, and Worldscope constituent lists from Datastream used in this study.

---

21 Datastream sometimes reassigns the major label for dead securities. For example, dead securities are sometimes replaced by new securities after re-emergence of a company from bankruptcy or if a preferred stock remains in issue.

C.2 Return Screens

First we select only dates between the first and last month where all unpadded unadjusted prices are available (undocumented Datastream items UP#S and UP#T).[22]

Ince and Porter (2006) point to the low accuracy of reported price and return indices (RIs) in Datastream—price indices (PIs) and total RIs generally contain only 1–2 decimal digits, which can lead to substantial inaccuracies in returns for declining stocks or stocks with long histories. Therefore, we calculate returns from unadjusted prices (Datastream item UP), unadjusted dividends (Datastream item UDDE), and capital adjustment indices (Datastream item CAI) whenever feasible. We also address Datastream reporting capital adjustment indices with a similar low accuracy by using a rational approximation to the ratio of capital adjustment indices. We use the rational approximation with the lowest absolute integer denominator >100, which yields the same value as $\frac{CAI_t}{CAI_{t-1}}$ when rounded to two decimal digits.

There are small differences between returns calculated from RIs from Datastream and returns calculated using the unadjusted values. Typically, the returns from unadjusted values are more reliable because of the numerical accuracy issues mentioned above. However, when there are large differences, RIs yield more reliable values because these contain less errors in capital adjustments and dividends. We therefore check if returns are within rounding errors of returns obtained from RIs.

We ensure that returns R_t satisfy

$$1 + R_t < \frac{RI_t + a}{RI_{t-1} - a} \qquad \text{when } RI_{t-1} > a, \qquad (C.1)$$

$$1 + R_t > \frac{RI_t - a}{RI_{t-1} + a} \qquad \text{when } RI_t > a, \qquad (C.2)$$

$$1 + R_t < 9.9 \qquad \text{when } RI_{t-1} \leq a, \qquad (C.3)$$

$$1 + R_t > 0 \qquad \text{when } RI_t \leq a, \qquad (C.4)$$

with $a = 0.005$ (maximal error from rounding to two decimal digits) or we fall back to calculating returns from the RI before cleaning for data errors.

Last, we closely follow Ince and Porter (2006) and Schmidt et al. (2011) and employ the following dynamic screens. We set returns that exceed 890% to missing and remove large return reversals by setting both R_t and R_{t-1} to missing whenever $(1 + R_{t-1})(1 + R_t) - 1 < 0.5$, while R_t or R_{t-1} exceed 300%. For daily returns which are required for the calculation of the Amihud (2002) illiquidity measure, we apply a similar filter that checks for return reversals within 5 weeks.[23]

---

22 UP#S and UP#T are unadjusted prices, unpadded while alive and unpadded while dead, respectively. This replaces the static screens at the beginning and end of the sample of Ince and Porter (2006) in order to remove prices of inactive or dead securities.

23 We obtain the set of daily returns where there is at least one return reversal within 5 weeks (35 calendar days equivalent to 25 trading days). That is we collect the set of daily returns $\{R_u, R_d\}$ for which the following conditions hold for $d - 25 \leq u \leq d + 25$:

$$R_d > 3, \qquad (C.5)$$

Appendix D: Linking Mutual Fund Holdings in the International Sample to Datastream and Worldscope

Factset identifies holdings information at the security and listing level. In our study, we match every reported holding to a primary company issue. This implies that we match primary, secondary listings, and depository receipts at the company level. Therefore, we first obtain static identifiers from Datastream for the constituent lists of Table BI as well as identifiers from a Datastream search for all ISIN, SEDOL, and CUSIP codes reported in Factset. We then obtain a list of all Datastream Codes which can be linked to a Factset reported mutual fund position via either the ISIN, SEDOL, or CUSIP Codes of a security or via any of these identifiers to the underlying of a depository receipt.[24]

We are able to obtain a raw link to a Datastream Code or Worldscope Permanent ID for 95.86% and 94.87% of all equity- and depository receipt-reporting date combinations in Factset, respectively. When we additionally apply the screens of Appendix C, most notably the requirement of available Wordscope data items, we are able to match 75.70% of all position-reporting date combinations to a security in one of the benchmark portfolios (see Appendix E.2 for procedures on calculating benchmark-adjusted portfolios for stocks in the International Sample).

Appendix E: Benchmark Portfolios

E.1 Characteristic-Adjusted Portfolios for the US Sample

In order to measure the risk-adjusted returns of stocks, we follow the methods prescribed by Daniel et al. (1997), Wermers (1999), and Wermers (2003), and compare the performance of stocks with that of stocks with similar size, book-to-market, and momentum characteristics (also known as DGTW returns). In the DGTW methodology, at the end of each June stocks are allocated to five size quintiles based on their market capitalization. Within each size quintile, stocks are further ranked in five quintiles based on their book-to-market ratios, yielding a total of 25 size and book-to-market sorted portfolios. Next, stocks within each of the twenty-five portfolios are further subdivided in five additional portfolios, based on their prior 12 month return. This procedure results in 125 stock portfolios. The benchmark returns are then computed as the returns of the 125 portfolios in the next 12 months, after which the portfolios are updated. We obtain the stock allocation and the returns of

$$(1 + R_u)(1 + R_d) - 1 \; < \; 0.5 \tag{C.6}$$

First, whenever both returns of $\{R_u, R_d\}$ are assigned to exactly one unique pair within that set, we only set R_u and R_d to missing. This covers cases where a single return reversal occurs: for example, due to a capital adjustment recorded at a wrong date. Second, when either is part of multiple pairs, we set all returns from $min(u, d)$ to $max(u, d)$ to missing as there are multiple reversals within the period.

24 We consider both securities and securities underlying a depository receipt (linked via Datastream item ADRP) since a mutual fund may hold a depository receipt. Worldscope also uses different company level identifiers for depository receipts when compared with the company level identifier of the underlying security.

the benchmark portfolios from Russ Wermers' webpage[25] and calculate benchmark-adjusted stock returns as returns in excess of the return of the relevant benchmark portfolio.

E.2 Characteristic-Adjusted Portfolios for the International Sample

For international stocks, we follow the approach proposed in Dyakov and Wipplinger (2020) that extends Wermers (1999) and Wermers (2003) to international stocks. Because global markets are not integrated (Fama and French, 2012) and risk premiums could be related to local factors (Griffin, 2002), we construct benchmark-adjusted portfolios separately for stocks belonging to broad geographical regions. Specifically, benchmark-adjusted portfolios are constructed for the developed regions of North-America, Europe, Japan, Asia-Pacific, as well as the EM regions of Asia-Pacific, Europe, and Latin America. Due to the relatively small number of stocks in some of these regions, we often resort to a number of stock buckets different form the original $5 \times 5 \times 5$ used by DGTW. The final portfolio breakdown in North America and Developed Europe splits stocks in 125 portfolios ($5 \times 5 \times 5$), 64 portfolios in Japan ($4 \times 4 \times 4$), 27 portfolios in Developed Asia-Pacific and EMs Asia-Pacific ($3 \times 3 \times 3$), and 8 portfolios in Emerging Markets Europe and Emerging Markets Latin America ($2 \times 2 \times 2$). Below we outline the construction of the benchmark portfolios in more details.

We begin the construction of the portfolios at the end of each June, when we assign companies from each region to a portfolio of companies with similar size, book-to-market, and momentum characteristics. We keep the assignment until next June, when we rebalance the portfolios.

E.2.1 Size and Momentum Characteristics

For the size characteristic, we obtain the market value reported in Datastream (Datastream item MV) as of the last trading day of each security in June. Then we use the exchange rates from Worldscope to convert market values into US Dollars. Our measure of Momentum is the 11-month total return in local currency (source: Datastream) of a security from June of the previous year till May.

E.2.2 Industry-Adjusted Book-To-Market Ratio

We calculate raw book-to-market ratios from the most recent fiscal year-end report as of December each year. We measure book equity (BE) as the sum of Total Common Equity (Worldscope item 03501) and Deferred Taxes (Worldscope item 035263). Because multiple share classes are common for international stocks, we use the values reported as Worldscope's Year-end Market capitalization (Worldscope item 08001) as market value of equity (ME). We use this item because it aggregates across share classes.[26] However,

---

25 The DGTW benchmarks are available on http://terpconnect.umd.edu/~wermers/ftpsite/Dgtw/cover page.htm.

26 Note that the market value used to calculate size in Section E.2.1 is derived differently from the market value used in calculating the book-to-market ratio. Size in this study is a security-level variable and book-to-market ratio is a company level variable. We use different approaches because of multiple share classes: We require that book value and market value are using the same base for book-to-market ratios. For size, we chose to go for a security level variable because it is not possible to obtain only the relevant share classes from Datastream and we do require size to reflect capital changes occurring after the last fiscal-year end.

Worldscope reports this value as of different dates compared with the fiscal year-end depending on country and we standardize all market values to the end of each December using returns from Datastream.

E.2.3 Industry Group Assignment

Log book-to-market ratio is calculated as $\text{btm}_{i,t}^{(j)} = \log(\text{BE}_{i,t}/\text{ME}_{i,t})$. Following Wermers (2003), we use a similar adjustment to the book-to-market ratio:

$$\text{btm}_{i,t}^{\text{adj}} = \frac{\text{btm}_{i,t}^{(j)} - \text{btm}_t^{(j)}}{\sigma_t^{(j)}}, \tag{E.1}$$

where $\text{btm}_{i,t}^{(j)}$ is the individual logarithmic book-to-market of a stock i belonging to industry group (j) and $\text{btm}_t^{(j)}$ is the logarithm of the aggregate book-to-market ratio of all $N_t^{(j)}$ stocks in group (j) at time t.

We further generalize Wermers (2003) who assigns stocks to groups based on SIC codes in the same industry to an international setting. Our group assignments additionally reflect that global markets are not fully integrated (Fama and French, 2012) and that accounting standards may differ among legal traditions. Because the sample size is limited for some countries we use the following assignment rules to allocate each stock i to an industry group (j).

1. We first attempt to match a stock i to a group (j) based on having the same country development status (developed vs. emerging Worldscope nation), the same country legal origin (English, German, French, Scandinavian), the same geographical region (North America, Europe, Asia Pacific, Japan, and Latin America), and having the same Industry Classification (ICB) Level 3 (forty-one sectors) as reported by Datastream.
2. If a group does not contain sufficient number of stocks, we relax the ICB restriction to Level 2 codes (nineteen supersectors);
3. Next, we drop the legal origin restriction; and
4. Last, we drop the geographical region restriction.

We always keep the assignment to developed or emerging countries until a stock can be matched to a group containing at least 20 stocks in the same industry. Any cases that cannot be matched are dropped.

E.2.4 Robust Adjustment

Another concern we address relates to the presence of outliers. Outliers are common among book-to-market ratios and can potentially affect the ranking across industries based on Equation (19) because an inflated standard deviation will pull all non-outliers in an industry toward the median. This already matters when outliers are only present among small companies because standard deviation is not a robust estimator of scale. The aggregate book-to-market ratio is more robust to outliers among smaller companies. However, for international companies, outliers may occur even among large stocks because of lower data quality for international firms.[27] Therefore, we apply robust alternatives for both the standard

---

27 Of particular concern is the possibility that our data provider has not collected information on all share classes as multiple classes of common stock occur more frequently among international stocks compared with US stocks. This would lead to overestimated aggregate book-to-market levels for an industry which would in turn lead to a general underestimation of adjusted book-to-market ratios for all stocks in the same group.

deviation and aggregate book-to-market level. Our industry adjustment replaces $\mathrm{btm}_t^{(j)}$ in Equation (19) with $\widetilde{\mathrm{btm}}_t^{(j)}$ by trimming stocks that are outliers from the aggregate book-to-market level and replacing standard deviation with the scale-τ estimate of Yohai and Zamar (1988) while keeping all observations. The industry adjusted book-to-market ratio becomes:

$$\widetilde{\mathrm{btm}}_{i,t}^{(j)} = \frac{\mathrm{btm}_{i,t}^{(j)} - \widetilde{\mathrm{btm}}_t^{(j)}}{\tau_t^{(j)}}, \tag{E.2}$$

where

$$\widetilde{\mathrm{btm}}_t^{(j)} = \ln\left(\widetilde{\mathrm{BTM}}_t^{(j)}\right) = \ln\left(\frac{\sum_k w_{k,t}^{(j)} \mathrm{BE}_{k,t}^{(j)}}{\sum_k w_{k,t}^{(j)}}\right) - \ln\left(\frac{\sum_k w_{k,t}^{(j)} \mathrm{ME}_{k,t}^{(j)}}{\sum_k w_{k,t}^{(j)}}\right), \tag{E.3}$$

with weights

$$w_{k,t}^{(j)} = \begin{cases} 0 & \text{if } |\mathrm{btm}_{k,t}^{(j)} - \widetilde{\mu}_{k,t}^{(j)}| \geq c\mathrm{MAD}_{k,t}^{(j)} \\ 1 & \text{otherwise} \end{cases} \tag{E.4}$$

and where MAD is the median absolute deviation from the median $\widetilde{\mu}_{k,t}^{(j)}$. The trim parameter c is set to 4.5 which is equal to the trim parameter in the calculation of the τ estimate of scale, for which we follow Maronna and Zamar (2002).[28] The advantage of this procedure is that the aggregate group book-to-market ratio remains unaltered when no outliers are present while the scale-tau estimator retains high efficiency and consistency for a Gaussian distribution even in the presence of outliers.

E.2.5 Portfolio Sorts
In order to assign stocks to characteristic-based benchmark portfolios we sort stocks first on size, then robust industry-adjusted book-to-market ratio, and lastly on momentum.

Importantly, we chose different numbers of splits per region (Table EI). This is especially important as it guarantees that each portfolio has sufficient securities assigned at any time and that there are equal amounts of splits according to size, robust industry-adjusted book-to-market ratios, and momentum within a region. For example, we split North-American stocks into 125 portfolios ($5\times 5\times 5$) as in Wermers (2003) but use 27 portfolios ($3\times3\times3$) for the developed Asian-Pacific region, and only 8 portfolios ($2\times2\times3$) for EMs Europe because these regions contain fewer listed companies.

Specifically, we

1. Sort primary securities on size, defined as the market value in June from Datastream in US dollars. Then, we assign stocks to a number of size portfolios based on their rank and Worldscope region. For developed markets, we first assign micro-capitalization stocks below a threshold of 1% of its region's total aggregate market capitalization to the lowest size portfolio and assign equal numbers of securities to the remaining size portfolios based on size rank. For emerging markets, we directly assign securities based on size rank.

---

28 We obtain the scale-τ estimate from winsorized residuals around a bi-weighted mean as in Maronna and Zamar (2002, p. 7) and rescale it by a constant factor of approximately 1.04 for consistency with the standard deviation of a normal distribution. Our implementation follows the normal-consistent estimator in the package "robustbase" for the statistical software R.

Table EI. Portfolios per region

The table shows the number of portfolios for each sorting step in June used in assigning securities to portfolios. Portfolio breakdown refers to the number of Size×Book-To-Market × Momentum splits per region.

Region		Portfolio breakdown	Total number of portfolios	(Avg.)
Developed markets				
North America	NAM	5×5×5	125	5,075
Europe	EUR	5 × 5×5	125	4,497
Japan	JPN	4 × 4×4	64	3,445
Asia-Pacific	APA	3 × 3×3	27	1,942
Emerging markets				
Emerging Asia-Pacific	EM_APA	3 × 3×3	27	5,034
Emerging Europe	EM_EUR	2 × 2×2	8	541
Emerging Latin America	EM_LAM	2 × 2×2	8	521

2. Each size portfolio is then subdivided into portfolios based on robust industry-adjusted book-to-market ratios (Section E.2.2) from the most recent fiscal year-end as of December.
3. Lastly, each of these sub-portfolios is further divided based on the 11-month return through May of the ranking year.

This ranking procedure is repeated at the end of each June and all the constituents are reassigned at this point in time. The return of a benchmark portfolio is the value-weighted return of its constituents (the highest ranked issue of Appendix C1) in local currency.

Appendix F: Details on Recursive Demeaning

The recursive demeaning estimation entails a two-stage estimation procedure. In the first stage, we regress forward-demeaned fund size, per market, upon its backward-demeaned version as well as forward-demeaned active industry size. In the second stage, we regress forward-demeaned trading performance on the fitted values from the first-stage regression as well as forward-demeaned active industry size.

$$\bar{q}_{f,t-1} = \psi^m + \rho_1^m \overline{\mathrm{AIS}}_{t-1}^m + \rho_2^m \underline{q}_{f,t-1} + v_{ft}^m, \tag{F.1}$$

$$\bar{r}_{ft}^m = \beta_1^m \overline{\mathrm{AIS}}_{t-1}^m + \beta_2^m \bar{q}_{f,t-1}^* + \vartheta_{ft}^m. \tag{F.2}$$

where $\bar{q}_{f,t-1}^*$ is the fitted value from Equation (F.1). Following Zhu (2018), Equation (F.1) includes an intercept as a zero intercept restriction imposes the unrealistic assumption that fund size fluctuates across a constant mean and further decreases goodness of fit. We refer to this estimator as RD1.

Alternatively, the forward-demeaned fund size variable can be instrumented using lagged fund size rather than lagged backward-demeaned size (Zhu, 2018). The new estimator, which we denote by RD2, is based on using fitted values from

$$\bar{q}_{f,t-1} = \psi^m + \rho_1^m \overline{\mathrm{AIS}}_{t-1}^m + \rho_2^m q_{f,t-1} + v_{ft}^m, \tag{F.3}$$

The instruments used in RD1 and RD2 are both exogenous because they are uncorrelated with innovations dated t or later. It is not a priori obvious which of these is the better instrument. Therefore, we combine the two in a third estimator, which we denote by RD3, which is based on the following reduced form:

$$\bar{q}_{f,t-1} = \psi^m + \rho_1^m \overline{\text{AIS}}_{t-1}^m + \rho_2^m \underline{q}_{f,t-1} + \rho_3^m q_{f,t-1} + v_{ft}^m. \tag{F.4}$$

As results are consistent across the three estimators, in Table IV we present results using Equation (F.4) only.

Appendix G: Variance of Optimal Industry Size

Empirically, the optimal industry size in Equation (5) is determined by the estimated coefficient on decreasing returns to scale and the estimated average gross alpha. Using Equations (6) and (11), we can express the optimal industry size AIS* as:

$$\text{AIS}^* = \frac{\hat{\alpha}_m^g}{2\hat{b}_m} + \frac{1}{2}\text{AIS}_m, \tag{G.1}$$

where $\hat{\alpha}_m^g$ denotes the average gross alpha per market, AIS_m is the average size of the active industry per market, and $\hat{b}_m$ is the estimated coefficient measuring decreasing returns to scale. Conditional upon average industry size, the variance of AIS* is drive by the (co)variances of $\hat{\alpha}_m^g$ and $\hat{b}_m$. Let us denote

$$\text{var}(\hat{\alpha}_m^g, \hat{b}_m) = \begin{pmatrix} \text{var}_a & \text{cov}_{ab} \\ \text{cov}_{ab} & \text{var}_b \end{pmatrix}.$$

Using a Taylor approximation, the variance of Equation (27) is given by

$$\text{var}\left(\widehat{\text{AIS}}^*\right) = \begin{pmatrix} \dfrac{1}{2b} & -\dfrac{\alpha_g}{2b^2} \end{pmatrix} \begin{pmatrix} \text{var}_a & \text{cov}_{ab} \\ \text{cov}_{ab} & \text{var}_b \end{pmatrix} \begin{pmatrix} \dfrac{1}{2b} \\ -\dfrac{\alpha_g}{2b^2} \end{pmatrix}.$$

This can be written as

$$\text{var}\left(\widehat{\text{AIS}}_*\right) = \frac{1}{4b^2}\text{var}_a + \frac{\alpha_g^2}{4b^4}\text{var}_b - \frac{\alpha_g}{2b^3}\text{cov}_{ab}. \tag{G.2}$$

Given that b_m is estimated using 2SLS after forward-demeaning (producing zero means per fund/market), it is reasonable to assume it has zero covariance with the level of average gross alpha. This allows us to estimate the variance of the estimated optimal industry size as

$$\widehat{\text{var}}\left(\widehat{\text{AIS}}^*\right) = \frac{1}{\hat{b}^2}\left[\frac{1}{4}\widehat{\text{var}}_a + \frac{\hat{\alpha}_g^2}{4\hat{b}^2}\widehat{\text{var}}_b\right]. \tag{G.3}$$

We use this to determine 95% confidence intervals, where we truncate the lower bound to zero if a negative value arises. These exceptional cases correspond to regions where an optimal industry size does not make sense and no economically meaningful equilibrium exists, for example when there are constant returns to scale in combination with a nonzero alpha on the first dollar.

References

Agarwal, V., Mullally, K., Tang, Y., and Yang, B. (2015): Mandatory portfolio disclosure, stock liquidity, and mutual fund performance, *The Journal of Finance* 70, 2733–2776.

Amihud, Y. (2002): Illiquidity and stock returns: cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Baker, M., Litov, L., Wachter, J. A., and Wurgler, J. (2010): Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements, *Journal of Financial and Quantitative Analysis* 45, 1111–1131.

Barber, B. M., Huang, X., and Odean, T. (2016): Which factors matter to investors? Evidence from mutual fund flows, *Review of Financial Studies* 29, 2600–2642.

Barras, L., Scaillet, O., and Wermers, R. (2010): False discoveries in mutual fund performance: measuring luck in estimated alphas, *The Journal of Finance* 65, 179–216.

Berk, J. and Green, R. (2004): Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

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

Berk, J. B. and van Binsbergen, J. H. (2016): Assessing asset pricing models using revealed preferences, *Journal of Financial Economics* 119, 1–23.

Berk, J. B. and van Binsbergen, J. H. (2017): Mutual funds in equilibrium, *Annual Review of Financial Economics* 9, 147–167.

Bhojraj, S., Cho, Y. J., and Yehuda, N. (2012): Mutual fund family size and mutual fund performance: the role of regulatory change, *Journal of Accounting Research* 50, 647–684.

Bollen, N. P. B. and Busse, J. A. (2006): Tick size and institutional trading costs: evidence from mutual funds, *Journal of Financial and Quantitative Analysis* 41, 915–937.

Boubaker, S., Gounopoulos, D., Nguyen, D. K., and Paltalidis, N. (2017): Assessing the effects of unconventional monetary policy and low interest rates on pension fund risk incentives, *Journal of Banking and Finance* 77, 35–52.

Carhart, M. M. (1997): On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Chakrabarty, B., Moulton, P. C., and Trzcinka, C. (2017): The performance of short-term institutional trades, *Journal of Financial and Quantitative Analysis* 52, 1403–1428.

Chen, H.-L., Jegadeesh, N., and Wermers, R. (2000): The value of active mutual fund management: an examination of the stockholdings and trades of fund managers, *The Journal of Financial and Quantitative Analysis* 35, 343–368.

Chen, J., Hong, H., Huang, M., and Kubik, J. D. (2004): Does fund size erode mutual fund performance? The role of liquidity and organization, *The American Economic Review* 94, 1276–1302.

Chung, K. H. and Chuwonganant, C. (2012): Regulation NMS and market quality, *Financial Management* 41, 285–317.

Chuprinn, O., Massa, M., and Schumacher, D. (2015): Outsourcing in the international mutual fund industry: an equilibrium view, *The Journal of Finance* 70, 2275–2308.

Coval, J. D. and Moskowitz, T. J. (1999): Home bias at home: local equity preference in domestic portfolios, *The Journal of Finance* 54, 2045–2073.

Cremers, K. J. M. and Petajisto, A. (2009): How active is your fund manager? A new measure that predicts performance, *The Review of Financial Studies* 22, 3329–3365.

Cremers, M. and Pareek, A. (2016): Patient capital outperformance: the investment skill of high active share managers who trade infrequently, *Journal of Financial Economics* 122, 288–306.

Dahlquist, M., Engström, S., and Söderlind, P. (2000): Performance and characteristics of Swedish mutual funds, *Journal of Financial and Quantitative Analysis* 35, 409–423.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997): Measuring fund performance with characteristic-based benchmarks, *The Journal of Finance* 52, 1035–1058.

Dasgupta, A., Prat, A., and Verardo, M. (2011): Institutional trade persistence and long-term equity returns, *The Journal of Finance* 66, 635–653.

Djankov, S., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2008): The law and economics of self-dealing, *Journal of Financial Economics* 88, 430–465.

Doidge, C., Karolyi, G. A., and Stulz, R. M. (2009): Has New York become less competitive than London in global markets? Evaluating foreign listing choices over time, *Journal of Financial Economics* 91, 253–277.

Doidge, C., Karolyi, G. A., and Stulz, R. M. (2013): The U.S. left behind? Financial globalization and the rise of IPOs outside the U.S., *Journal of Financial Economics* 110, 546–573.

Duan, Y., Hu, G., and McLean, D. R. (2009): When is stock picking likely to be successful? Evidence from mutual funds, *Financial Analysts Journal* 65, 55–66.

Dyakov, T. and Wipplinger, E. (2020): Institutional ownership and future stock returns: an international perspective. *Review of International Finance*, forthcoming.

Edelen, R., Evans, R., and Kadlec, G. (2007): Scale effects in mutual fund performance: the role of trading costs. Working paper.

Edelen, R., Evans, R., and Kadlec, G. (2013): Shedding light on "invisible" costs: trading costs and mutual fund performance, *Financial Analyst Journal* 69, 33–44.

Elton, E. J., Gruber, M. J., and Blake, C. R. (2001): A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *The Journal of Finance* 56, 2415–2430.

Elton, E. J., Gruber, M. J., Das, S., and Hlavka, M. (1993): Efficiency with costly information: a reinterpretation of evidence from managed portfolios, *The Review of Financial Studies* 6, 1–22.

Fama, E. F. and French, K. R. (2012): Size, value, and momentum in international stock returns, *Journal of Financial Economics* 105, 457–472.

Ferreira, M. A., Keswani, A., Miguel, A. F., and Ramos, S. B. (2013): The determinants of mutual fund performance: a cross-country study, *Review of Finance* 17, 483–525.

Ferreira, M. A. and Matos, P. (2008): The colors of investors' money: the role of institutional investors around the world, *Journal of Financial Economics* 88, 499–533.

Ferreira, M. A., Matos, P., Pereira, J. P., and Pires, P. (2017): Do locals know better? A comparison of the performance of local and foreign institutional investors, *Journal of Banking and Finance* 82, 151–164.

French, K. (2008): Presidential address: the cost of active investing, *The Journal of Finance* 63, 1537–1573.

Gompers, P. and Metrick, A. (2001): Institutional investors and equity prices, *Quarterly Journal of Economics* 116, 229–259.

Graham, J. R., Harvey, C. R., and Huang, H. (2009): Investor competence, trading frequency, and home bias, *Management Science* 55, 1094–1106.

Griffin, J. M. (2002): Are the Fama and French factors global or country specific?, *The Review of Financial Studies* 15, 783–803.

Grinblatt, M. and Titman, S. (1989): Mutual fund performance: an analysis of quarterly portfolio holdings, *The Review of Business* 62, 393–416.

Grinblatt, M., Titman, S., and Wermers, R. (1995): Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behavior, *American Economic Review* 85, 1088–1105.

Grullon, G., Larkin, Y., and Michaely, R. (2019): Are U.S. industries becoming more concentrated?, *Review of Finance*, 23, 697–743.

Hou, K., Karolyi, G. A., and Kho, B.-C. (2011): What factors drive global stock returns?, *The Review of Financial Studies* 24, 2527–2574.

Huij, J. and Verbeek, M. (2009): On the use of multifactor models to evaluate mutual fund performance, *Financial Management* 38, 75–102.

Ince, O. S. and Porter, R. B. (2006): Individual equity return data from Thomson Datastream: handle with care!, *Journal of Financial Research* 29, 463–479.

Jiang, H. and Verardo, M. (2018): Does herding behavior reveal skill? An analysis of mutual fund performance, *The Journal of Finance* 73, 2229–2269.

Kacperczyk, M., Sialm, C., and Zheng, L. (2008): Unobserved actions of mutual funds, *The Review of Financial Studies* 21, 2379–2416.

Khandani, A. E. and Lo, A. W. (2011): What happened to the quants in August 2007? Evidence from factors and transactions data, *Journal of Financial Markets* 14, 1–46.

Khorana, A., Servaes, H., and Tufano, P. (2005): Explaining the size of the mutual fund industry across the world, *Journal of Financial Economics* 78, 145–185.

Lakonishok, J., Shleifer, A., and Vishny, R. W. (1992): The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23–43.

Lan, C., Moneta, F., and Wermers, R. (2018): Holding horizon: a new measure of active investment management. Working paper (https://ssrn.com/abstract=2517215).

Leippold, M. and Rueegg, R. (2018): Is active investing a zero-sum game? Working paper.

Lewellen, J. (2011): Institutional investors and the limits of arbitrage, *Journal of Financial Economics* 102, 62–80.

Maronna, R. A. and Zamar, R. H. (2002): Robust estimates of location and dispersion for high-dimensional datasets, *Technometrics* 44, 307–317.

Menkveld, A. J. (2016): The economics of high-frequency trading: taking stock, *Annual Review of Financial Economics* 8, 1–24.

Pastor, L. and Stambaugh, R. F. (2012): On the size of the active management industry, *Journal of Political Economy* 120, 740–781.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2015): Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2017): Do funds make more when they trade more?, *The Journal of Finance* 72, 1483–1528.

Petajisto, A. (2013): Active share and mutual fund performance, *Financial Analyst Journal* 69, 73–93.

Puckett, A. and Yan, X. S. (2011): The interim trading skills of institutional investors, *The Journal of Finance* 66, 601–633.

Schmidt, P. S., von Arx, U., Schrimpf, A., Wagner, A. F., and Ziegler, A. (2011): On the construction of common size, value and momentum factors in international stock markets: a guide with applications. Working paper.

Sias, R. W. (2004): Institutional herding, *The Review of Financial Studies* 17, 165–206.

Stein, J. (2009): Presidential address: sophisticated investors and market efficiency, *The Journal of Finance* 64, 1517–1548.

Wei, K. D., Wermers, R., and Yao, T. (2014): Uncommon value: the investment performance of contrarian funds, *Management Science* 61, 2394–2414.

Wermers, R. (1999): Mutual fund herding and the impact on stock prices, *The Journal of Finance* 54, 581–622.

Wermers, R. (2003): Is money really "smart"? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working paper.

Yohai, V. J. and Zamar, R. H. (1988): High breakdown-point estimates of regression by means of the minimization of an efficient scale, *Journal of the American Statistical Association* 83, 406–413.

Zhu, M. (2018): Informative fund size, managerial skill and investor rationality, *Journal of Financial Economics* 130, 114–134.