Review of Finance (2014) 18: pp. 561–590
doi:10.1093/rof/rft015
Advance Access publication: May 31, 2013

Investing in a Global World*

JEFFREY A. BUSSE[1], AMIT GOYAL[2] and SUNIL WAHAL[3]
[1]*Goizueta Business School, Emory University,* [2]*Swiss Finance Institute, University of Lausanne, and* [3]*WP Carey School of Business, Arizona State University*

Abstract. We examine active retail mutual funds and institutional products with a mandate to invest in global equity markets. We find little reliable evidence of alphas in the aggregate or on average. The right tail of the distribution contains some large alphas. Decomposing stock selection from country selection, we find some evidence of superior stock picking abilities in the extreme right tail. However, simulations suggest that they are produced just as likely by luck as by skill. Persistence tests show little evidence of continuation in superior performance.

JEL Classification: G15, G23

1. Introduction

We study the returns delivered by: (i) US-registered active retail mutual funds that allow individual US investors to invest in global equities, and (ii) active institutional products that allow plan sponsors, endowments, and foundations to invest in global equities.[1] We seek to understand the cross-sectional and time-series distribution of alphas so as to address issues of performance and persistence.

There is ample evidence on the performance and persistence of funds that invest in US equities. Some conclude that there is little alpha to be had, that

* We thank seminar participants at Arizona State University, Emory University, Texas Christian University, University of Pompeu Fabra, University of Rochester, University of St Gallen, University of Utah, an anonymous referee, Michael Goldstein, Andrew Karolyi, Antonio Macias, Laura Starks, and especially Narasimhan Jegadeesh for helpful suggestions. Amit Goyal would like to thank Rajna Gibson for her support through her NCCR-FINRISK project. We thank Zhiyi Qian for excellent research assistance.

[1] Institutional products are not the same as (and should not be confused with) the institutional class of traditional retail mutual funds. Instead, they represent strategies from which separate accounts for plan sponsors are derived with independent Investment Management Agreements (IMAs) between the investment advisor and the plan sponsor. As many separate accounts can be generated by the same strategy, investment advisors report composite returns under the Global Investment Performance Standards agreement (referred to as GIPS compliant returns). Henceforth, we refer to these institutional composites as "institutional funds." We do so only for convenience and emphasize that these are different investment vehicles from standard mutual funds.

what appears to be alpha is likely due to luck, and that there is no persistence. Others offer evidence of skill and argue that it is persistent. Some examples of the former include Carhart (1997), Fama and French (2010), and Busse, Goyal, and Wahal. (2010). On the other hand, Chen, Jegadeesh, and Wermers (2000), Bollen and Busse (2005), Cohen, Coval, and Pastor (2005), Kacperczyk, Sialm, and Zheng (2005, 2008), Kosowski *et al.* (2006), Kacperczyk and Seru (2007), Cremers and Petajisto (2009), Cohen, Polk, and Silli (2010), and Busse and Tong (2012) provide evidence in favor of skill and its persistence.

Given this literature, our examination is of independent interest for at least two reasons. First, the market we study is large but with almost no large sample direct evidence. At the end of 2009, an individual investor domiciled in the USA could invest in over 600 US-registered equity mutual funds that invested ex-USA with >$900 billion in assets. Large institutional investors (plan sponsors, endowments, and foundations) could similarly invest in over 700 institutional investment products whose mandates were either international (ex-USA) or global, with assets >$1.2 trillion. Yet, despite this size and potential importance, this set of investment vehicles has received little academic attention. Second, it is not an uncommon belief that markets outside of the most developed ones are less efficient and, therefore, exploitable by active fund managers (e.g., see Bekaert and Harvey (2002) for academic evidence on relative inefficiency of emerging markets as well as Griffin, Kelly, and Nardari (2010) and references therein). All else equal, this implies that the search for alpha (by academics and practitioners alike) would be more fruitful in international (ex-USA) markets. Even if funds that invest exclusively in the USA do not generate excess performance or display persistence (a not uncontroversial statement given the above-cited and extant literature), it does not logically follow that the same is necessarily true for funds that invest globally. We bring direct evidence to bear on this issue.

We find that the three (four)-factor alphas of equal- and value-weighted portfolios of retail funds are slightly positive (negative), but statistically indistinguishable from zero; gross of fees, on average or on aggregate, there does not appear to be persuasive evidence of risk-adjusted excess returns. One might anticipate systematic differences between retail and institutional funds for a variety of economic and structural reasons. Institutional funds cater to a more sophisticated group of investors that provide relatively patient capital and engage in costly monitoring of portfolios. Del Guercio and Tkac (2002) document significant differences in the flow–performance relationship between retail and institutional funds, and Goyal and Wahal (2008) show that consultants are active in hiring, firing, and retention decisions. Institutional capital may also be more patient, resulting in lower turnover and trading costs, and hence potentially higher gross returns. We find that equal-weighted alphas of institutional funds are in fact higher than their retail counterparts by 0.25% per quarter. However, they are still statistically indistinguishable from zero. Value-weighted institutional alphas are surprisingly more negative than retail funds, indicating that any superior performance of institutional funds arises from those that are relatively small.

To estimate fund-specific alphas, we require a minimum number of returns. As worse performing funds are more likely to die early, this truncates the distribution of funds from the left resulting in a sample that is more likely to have positive alphas.[2] Ignoring this truncation, the right tail shows some very large alphas. For retail and institutional funds, the 95th percentile of alphas using four-factor models is 1.53 and 1.97% per quarter, respectively. By comparison, Kosowski *et al*. (2006) report that the 95th percentile of four-factor alphas for US (domestic equity) funds is 0.35% per quarter. There are at least two potential explanations for the large alphas in our sample. One possibility is that the managers of these funds generate superior performance through stock and/or country selection. Another is that the alphas in the tails are not due to genuine skill but merely luck. If it is luck, it is unlikely to persist.

For each fund, we calculate the difference between the actual fund return and the return that an investor would have earned with the same country weights but earning passive country index returns. This decomposition separates stock versus country selection and shows some statistical evidence in favor of security selection at the extreme right tail of the distribution of alphas. Approximately, 8% of each institutional and retail funds have positive and significant alphas (at the 5% level of significance) that are attributable to stock selection.[3]

Funds can be in these apparently thick right tails because managers are truly skilled in these international markets or because of pure chance. We employ two well-established procedures to separate the two. First, in the spirit of Kosowski *et al*. (2006) and Fama and French (2010), we perform simulations that compare simulated t-statistics under the null of no alpha to actual t-statistics. These simulations show little evidence of skill—funds that

[2] We verify that this is indeed the case in our sample. The simple average annualized return of retail funds that are truncated is 5.93%, whereas the average return of the nontruncated sample is almost twice as large at 10.10%.

[3] We also investigate the country selection process itself and find that country weights are uncorrelated with previous returns. An exception is that when active international funds enter a country, they do so after positive returns.

appear in the tails are likely there by chance. For instance, if one considers the 95th percentile of the cross-sectional distribution of t-statistics of four-factor alphas for institutional and retail funds, the percentage of simulated t-statistics that are greater than actual t-statistics is 8.5 and 7.7%, respectively. Second, we perform persistence tests because if a fund has a large alpha purely by chance, it is unlikely that the alpha will persist. An added advantage is that fund investors are naturally more interested in whether there are *ex ante* rules that allow them to choose subsequent winners over a certain period. Tests of persistence again show differences between institutional and retail funds. Using three-factor models, institutional funds in the top deciles deliver persistently high returns up to 2 years after deciles formation. In contrast, there is no persistence in retail funds. However, just as in US domestic equity funds (Carhart, 1997), accounting for the mechanical effects of momentum removes any evidence of persistence, even in institutional funds. Although the lack of persistence certainly suggests a lack of skill, that need not logically follow. Indeed, in models such as Berk and Green (2004), individual managers can have skill and yet there be no persistence. Yet, the totality of the evidence—average and aggregate alphas, luck versus skill simulations, and the persistence tests—suggests little systematic evidence of skill.

Where does this leave the average investor? All our results are generated using returns that are gross of fees. Expense ratios generally vary between 1 and 2% per year, further eroding any risk-adjusted excess return that an investor might hope for. We cannot and do not claim that an investor should not invest in these funds. Indeed, the diversification benefits offered by these funds may be such that they offset the incremental costs. Our results do suggest, however, that the average investor would be better off diversifying via passive rather than active funds.

Although our funds are nominally designated for US investors, the results are useful for a broader global audience. Institutional funds are not restricted to US-domiciled investors and are easily transfigured for non-US sponsors. Similarly, although our retail funds are registered in the USA (under the 1940 Investment Company Act), there is little that prevents fund families from marketing a single portfolio strategy using different funds in different countries. Creating such clone funds is not uncommon, merely requiring registration and compliance with local regimes and well within the capabilities of global fund management companies. From an economic perspective, two funds with the same underlying portfolio but registered in two different countries are similar if not identical.

Our article fits into the broad literature on the performance of delegated asset managers. As mentioned earlier, the vast majority of that literature is

US-focused, with some recent notable exceptions. Khorana, Servaes, and Tufano (2005, 2009) try to understand the size and fee structure of the mutual fund industry around the world. There are numerous studies that examine the performance of funds in non-US countries (see, e.g., Blake and Timmermann, 1998; Panetta and Cesari, 2002) with mixed conclusions. Earlier papers also investigate performance and persistence for US-based international funds but are limited by data quantity and quality. For example, Cumby and Glen (1990) find no evidence of superior performance for a sample of fifteen funds between 1982 and 1988 using a market model. Droms and Walker (2001) look at winner–winner transition probabilities in funds between 1977 and 1996 using annual returns. They find persistence in 1 year returns but none thereafter, and do not control for the mechanical effects of momentum. There are three recent studies with a broader international perspective. Ferreira *et al.* (2013) study domestic and international equity funds across twenty-seven countries, but their interest is in scale issues and cross-sectional determinants of performance differences. Didier, Rigobon, and Schmukler (2011) seek to understand why US funds that invest overseas invest in a relatively small number of securities. Cremers *et al.* (2011) study funds from thirty countries and find that closet indexing is common and that explicit indexing is rare.

The remainder of the article is organized as follows. Section 2 describes the mutual fund and institutional data as well as our procedures for factor construction. Section 3 contains our results. Section 4 concludes.

2. Data

Our analysis requires data from multiple sources. In this section, we describe the basic elements of the data and our sampling procedures as well as descriptive statistics for these data.

2.1 INSTITUTONAL FUNDS

Data on global institutional products come from Informa Investment Solutions (IIS), a firm that provides data, services, and consulting to plan sponsors, investment consultants and investment managers. We refer to these products as “funds” in the text but they are more appropriately thought of as strategies from which multiple products (funds, separate accounts, etc.) are derived. They are distinct from the institutional class of mutual funds reported in traditional mutual fund databases, such as the Center for Research in Security Prices (CRSP) Survivor-Bias-Free Mutual Fund

Database and Morningstar, and are often referred to as composites to denote the idea that one strategy can give birth to multiple products. The products themselves are sold to pension funds, endowments, foundations, and large trusts. Our analysis is conducted on these composites so there is no multiple counting of derived strategies in the data. Although the data are self-reported, plan sponsor oversight, the threat of actual regulatory audits, monitoring by investment consulting firms, and GIPS compliance requirements ensure accuracy (see www.gipsstandards.org). The sample period is from 1991 to 2009.

The database contains descriptions of investment style, which we use to identify international equity funds and to assign benchmark indices. This eliminates domestic equity, fixed income, and balanced funds. We also eliminate all passive funds, leaving us with a sample of 1,218 institutional funds.

For this sample, we gather quarterly returns that are net of trading costs but gross of fees. We also collect annual fee schedules, portfolio turnover, and assets under management. Asset information is not available for all funds. The data also contain self-reported quarterly country weights for each fund, although the weight may have a reporting delay (i.e., the weight may not be at the end of the quarter but rather during the quarter). These country weights are available for the period 1996–2009. Some data items are cross-sectional, such as investment style and manager-identified benchmark.

2.2 RETAIL FUNDS

We start with all mutual funds in the CRSP Mutual Fund database from 1987 to 2009. To this universe, we apply a sequence of filters. We retain only funds for which the Lipper Asset Code is equity. This eliminates fixed income funds (including money market and convertible bond funds), real estate funds, and balanced funds. One way to separate international/global equity funds from domestic equity funds is to employ the “product type” field maintained by CRSP that explicitly identifies international funds. However, an examination of the raw data shows that this field is frequently missing and omits many international/global equity funds. Therefore, we manually examine fund objectives (labeled “Lipper Class Names” by CRSP) to extract a relevant list of funds. Roughly speaking, it includes all funds identified with “global” and “international” monikers and some regional funds. These procedures generate a list of international equity funds. From this list, we further remove funds that CRSP identifies as being variable annuity or index funds. Unfortunately, these programmatic filters are still inadequate. Therefore, based on the fund’s name, we eliminate levered funds, and any remaining index funds. This leaves us with a group of funds whose investment objective is to invest ex-USA or globally.

Our country momentum and persistence tests require each fund to have a benchmark index. For each fund, we use a sequential process to assign a benchmark to each fund. First, we use the fund's prospectus benchmark as reported in the Morningstar Direct database. Second, we examine documents filed by the fund with the Securities and Exchange Commission (SEC), typically in the Electronic Data-Gathering, Analysis, and Retrieval (EDGAR) filing system or through commercial reference databases. Third, if the above procedures are unable to assign a benchmark, we use the Lipper Class Name provided by CRSP to assign a benchmark that approximates that class name. In the majority of cases, the benchmark indices established using these procedures are Morgan Stanley Capital International (MSCI)-based. In cases where a fund designates a non-MSCI index as its benchmark, we map it to an equivalent MSCI index. This allows us to use a manageable number of indices and also enables comparisons with institutional funds.

The above process delivers a sample of 1,019 funds for our sample period. For these funds, we extract data on returns, total net assets (TNA), 12b-1 fees, management fees, expense ratios, and portfolio turnover from CRSP. As these data items are at the fund class level, we aggregate them to a fund level using a two-step process. CRSP does not match fund classes with parent funds for these funds, so we manually assign multiple classes of a fund to a unique fund identifier. Frequently, such assignment is obvious from the fund's name (e.g., Aberdeen International Equity Fund Class A Shares and Aberdeen International Equity Fund Class B Shares). In some cases, however, this again requires background research from public sources (typically the fund's website, Lipper, or Morningstar). Once unique fund identifiers are assigned, we calculate TNA-weighted averages across all fund classes for the variables of interest.

We obtain country weight information from the Morningstar Direct database. The weights are reported monthly for each fund class. As before, we map each fund class to unique fund identifiers. As a check on these self-reported weights, we also design a web crawler to access mandatory weights reported on Form N-Q to the SEC. These data are only available after 2003, and filings are at the end of the 1st and 3rd quarter. Nonetheless, they allow us to verify the accuracy of the self-reported weights.

2.3 FACTORS

We conduct performance evaluation using alphas from factor models. There is, however, little agreement about the right asset pricing model. Fama and French (1998) demonstrate the failure of the international capital asset pricing model (CAPM) and propose a two-factor model that includes a value factor. Griffin (2002) asks whether these factors should be local or global. Hou, Karolyi, and Kho (2011) find that momentum and value factors capture much of the common variation in global stock returns.

There is also little consensus about the right way to construct factors. Fama and French (1998) and Griffin (2002) construct world factors using country-specific breakpoints and then employ country (total market capitalization) weights to aggregate across countries. Hou, Karolyi, and Kho (2011) construct factors using global, rather than country-specific, breakpoints. Finally, Fama and French (2011) examine variation in value and momentum premiums across size groups by constructing factors using regional size breakpoints.

Some of the variation in construction methods is undoubtedly due to views about market integration and/or the precise nature of the tests. Unlike the above authors, we have no asset pricing purpose and therefore take a stripped down approach to performance assessment. Given the weight of the above-cited literature, we focus our attention on three- and four-factor alphas. With respect to factor construction, we follow the approach used by Ken French as detailed on his website and largely following Fama and French (1998). That is, we first construct factors for each country using country-specific breakpoints, and then aggregate these country-specific factors to developed and emerging markets, respectively, based on the country's market capitalization. That way, for example, small stocks in Thailand are separated from large stocks in Thailand; using global breakpoints, most stocks in Thailand would enter into a small stock portfolio. We believe our approach makes both economic and practical sense: many portfolio managers make a distinction between small and large stocks within a country.

We construct separate factors for developed and emerging markets and use both in factor models. There are several advantages to doing so. This geographic classification reflects the way that many practitioners think about performance issues and design portfolios; international funds are frequently categorized and characterized as "Developed Market" funds, "Emerging Market" funds, or "EAFE" funds. Despite that, the sample of funds is heterogeneous in terms of the countries that they invest in. Some of this is by choice, our sample includes funds that, for example, invest in East Asia (which includes both developed markets such as Singapore as well as emerging markets such as Indonesia). Another source of heterogeneity comes from variations in definitions from different index providers or index reclassifications over time. As an example of the former, MSCI classifies South Korea as an emerging market, but Financial Times Stock Exchange (FTSE) classifies it as a developed market. Even within index providers, classifications change over time. In May 2010, for example, Israel was reclassified by MSCI from an emerging market to a developed market. Although active funds may be benchmark sensitive (although certainly less so than index funds), they need not be slaves to such definitions. As a result of this, and because our sample of funds is inclusive of funds with wide geographic mandates, we include factors for both developed and emerging markets. For both these regions, we construct a market factor (MktmRf), a size factor (SMB), a value factor (HML), and a momentum factor (UMD). Further details on factor construction are provided in Appendix A.

We do not show detailed factor returns to avoid redundancy with other studies. By way of data description, however, the average annual excess market return for developed and emerging markets in our sample period is 4.34 and 12.28%, with t-statistics of 1.33 and 2.39, respectively. The average size, value, and momentum premiums are 1.83, 4.94, and 6.47% for developed markets and 12.53, 8.17, and 4.93% for emerging markets, respectively.

2.4 DESCRIPTIVE STATISTICS

2.4.a Institutional funds

Panel A of Table I shows the number of institutional funds, average and total assets, average portfolio turnover rates, and fee schedules in each year. Generally, the number of funds and total assets increase over time. In 2008, concomitant with the market break, both the number of funds and total assets decline. Despite that decline, the average annual growth rate in the number of global institutional funds (total assets) over this sample period is 7.9% (13.4%). At the end of the time series, there are 777 funds with >$1.2 trillion invested.

Average portfolio turnover varies from a minimum of 52% per year in 1991 to a maximum of 83% in 2000, but averages about 65%. Fees reported in the panel are constructed from fee schedules based on investments of $10 million, $50 million, and $100 million. These schedules are nominal and represent an upper limit; actual fees are typically lower and are individually negotiated by plan sponsors. There appears to be a slight increase in average stated fees over time but the increase is gradual. Not surprisingly, fees are lower for large investment mandates reflecting expected quantity discounts. The stated fees are higher than those for domestic equity funds. For instance, in 2009, the average $10 m, $50 m, and $100 m fees in our global funds are 0.87, 0.78 and 0.72%, respectively. The equivalent fees for domestic (USA)

Table I. Descriptive statistics

The sample of institutional funds consists of all products reported by the data provider, IIS, with an investment style designated as global or international equity between 1991 and 2009. The sample of retail funds consists of all active global or international equity mutual funds in the CRSP Mutual Fund Database between 1987 and 2009. For retail funds, class-level information is aggregated to fund level using TNA-weighted averages. Assets are in millions of dollars, turnover is in percent per year, and fees are in percent per year.

	Number of funds	Total assets	Average assets	Average turnover	Fees: $10 million	Fees: $50 million	Fees: $100 million
Panel A: institutional funds							
1991	211	141,571	757	52.6	0.87	0.75	0.65
1992	253	89,844	399	55.8	0.85	0.72	0.64
1993	308	174,652	677	57.9	0.86	0.72	0.64
1994	364	242,453	758	53.9	0.89	0.73	0.64
1995	429	309,099	818	54.5	0.89	0.74	0.65
1996	488	428,753	1,028	54.0	0.87	0.73	0.64
1997	541	648,428	1,400	63.2	0.88	0.73	0.64
1998	588	698,818	1,387	67.9	0.87	0.74	0.65
1999	632	1,005,410	1,831	73.1	0.86	0.74	0.66
2000	655	916,673	1,622	83.2	0.86	0.75	0.67
2001	693	841,721	1,431	80.5	0.89	0.75	0.67
2002	713	823,683	1,344	71.4	0.87	0.76	0.68
2003	719	1,189,642	1,871	70.5	0.86	0.76	0.68
2004	748	1,558,364	2,354	67.5	0.85	0.76	0.68
2005	778	1,873,300	2,692	66.1	0.85	0.76	0.69
2006	813	2,387,371	3,226	66.5	0.87	0.76	0.69
2007	850	2,746,766	3,464	66.5	0.87	0.77	0.71
2008	842	1,409,145	1,874	72.6	0.87	0.77	0.72
2009	777	1,201,898	2,568	66.0	0.87	0.78	0.72

	Number of funds	Total assets	Average assets	Average turnover	Fees: 12b1	Fees: ExpRatio	Fees: Management
Panel B: retail funds							
1987	76	15,541	219	76.0	–	1.19	–
1988	89	16,142	199	93.7	–	1.45	–
1989	99	20,401	219	75.5	–	1.65	–
1990	121	22,090	203	71.3	–	1.83	–
1991	144	29,969	225	88.0	–	1.76	–
1992	188	36,467	217	12.0	–	1.53	–
1993	243	83,237	391	71.6	0.24	1.74	–
1994	319	117,437	397	69.9	0.25	1.69	–
1995	378	147,302	414	62.4	0.29	1.69	–

(continued)

Table I. Continued

	Number of funds	Total assets	Average assets	Average turnover	Fees		
					12b1	ExpRatio	Management
1996	440	216,432	530	73.9	0.31	1.78	–
1997	517	277,940	583	70.0	0.30	1.81	–
1998	571	311,388	548	62.3	0.26	1.70	–
1999	612	450,917	753	81.6	0.25	1.72	0.82
2000	614	418,435	677	96.8	0.52	1.70	0.86
2001	619	336,047	549	104.0	0.55	1.73	0.86
2002	584	282,951	490	104.8	0.56	1.82	0.82
2003	541	411,711	772	81.8	0.53	1.81	0.80
2004	508	551,858	1,086	85.2	0.50	1.67	0.82
2005	527	741,398	1,381	77.2	0.48	1.63	0.83
2006	564	1,024,810	1,801	66.6	0.47	1.57	0.87
2007	618	1,254,134	2,090	60.8	0.46	1.49	0.83
2008	663	660,525	1,002	63.0	0.47	1.48	0.90
2009	655	917,756	1,412	84.2	0.56	1.51	0.78

equity reported by Busse, Goyal, and Wahal (2010) are 0.81, 0.69, and 0.64%, respectively.

2.4.b. Retail funds

Panel B of Table I shows that the sample of retail mutual funds also experiences growth over time. The number of funds (total assets under management) grows from seventy-six ($15 billion) in 1987 to 655 ($918 billion) in 2009, representing a 10% (20%) annual increase. Together with institutional funds, at the end of 2009, these investments represent ~4.5% of the total world market capitalization of $45 trillion (World Federation of Exchanges). Average turnover over the entire sample period is 75%, which is higher than for institutional funds. This is consistent with the expectation that retail fund flows have higher variability than institutional flows. The last three columns of the panel show actual 12b-1 fees, expense ratios, and management fees for funds in each year. CRSP does not record 12b-1 and management fees until 1993 and 1999, respectively, but expense ratios are available for the entire time series. Expense ratios generally increase up to the early 2000s and decline thereafter. Expense ratios are between 1.5 and 2.0%. In contrast, the Investment Company Institute (2010) reports average expense ratios for all stock funds declining from 1.28% in 2000 to 0.99% in 2008.

3. Results

3.1 METHODOLOGY

We calculate abnormal performance based on alphas from factor models. We use factors corresponding to the CAPM, Fama and French (1993) three-factor model, and the Carhart (1997) four-factor model to provide as much information as possible to the reader. But we conduct the majority of inferences using a four-factor approach. To do so, we estimate the following time-series regressions:

$$R_{\mathrm{it}} - R_{\mathrm{ft}} = \alpha_{\mathrm{i}} + \beta'_{\mathrm{id}} F_{\mathrm{dt}} + \beta'_{\mathrm{ie}} F_{\mathrm{et}} + u_{\mathrm{it}}, \tag{1}$$

where F's are factors, and subscripts d and e refer to developed and emerging markets, respectively. R_{ft} is the US 1-month Treasury Bill rate.

The dependent variables in these regressions consist of (i) equal-weighted returns of all funds, (ii) value-weighted returns of all funds where the weight is the assets under management in the previous year for institutional funds and in the previous quarter for mutual funds, or (iii) individual fund returns. Equal-weighted returns give us a sense of what the average fund delivers. Value-weighted returns provide information about what the industry as a whole delivers.

We use the highest frequency of data available; for retail mutual funds we estimate regressions using monthly returns, whereas for institutional funds we use quarterly data. Retail funds are net of fees, so to compare them to institutional funds, we add one-twelfth of the annual expense ratio back to the monthly return. For consistency, we report all alphas in percent per quarter, even for regressions using monthly frequency.

3.2 AVERAGE AND AGGREGATE PERFORMANCE

Table II shows factor models using equal- and value-weighted fund returns. Panels A and B show models for institutional and retail funds, respectively. We report both factor loadings and alphas. The numbers in parentheses next to the coefficients are *t*-statistics.

Panel A shows that the sum of market betas on developed and emerging markets for institutional funds is close to one. The coefficients on SMB indicate a size tilt toward smaller firms in developed markets but not in emerging markets. In addition, lower SMB loadings for regressions with value-weighted aggregate returns than those using equal-weighted average returns suggest that larger funds invest in higher market capitalization securities. For developed markets, loadings on the value factor are

Table II. Average and aggregate alphas and factor loadings

We form portfolios from individual funds. Portfolios are both equal- and value-weighted (we value weight based on asset size from December of the previous year). Returns are gross of fees. We report alphas and betas from time-series regressions of these portfolios on different factor models. We use three combinations of factors—the one-factor model corresponds to just the market factor (MktmRf), the three-factor model corresponds to the three Fama and French (1993) factors (MktmRf, SMB, and HML), and the four-factor model adds the momentum factor to the mix (MktmRf, SMB, HML, and UMD). Separate factors are included for developed (D) and emerging (E) markets. Regression frequency is quarterly for institutional funds and monthly for retail mutual funds. All alphas are in quarterly percent. Numbers in parentheses are *t*-statistics. The sample period is 1991–2009 for institutional funds and 1987–2009 for retail funds.

	Equal-weighted average			Value-weighted aggregate		
Panel A: institutional funds						
Alpha	0.405	0.481	0.232	0.131	0.037	−0.136
	(1.89)	(2.04)	(0.99)	(0.65)	(0.17)	(−0.60)
MktmRf(D)	0.757	0.768	0.799	0.818	0.835	0.857
	(21.11)	(21.32)	(22.99)	(24.10)	(24.87)	(25.41)
SMB(D)		0.116	0.140		0.067	0.083
		(2.22)	(2.84)		(1.37)	(1.75)
HML(D)		0.060	0.109		0.121	0.156
		(1.40)	(2.57)		(3.05)	(3.78)
UMD(D)			0.087			0.060
			(2.63)			(1.87)
MktmRf(E)	0.250	0.243	0.255	0.194	0.193	0.201
	(11.65)	(11.08)	(12.11)	(9.54)	(9.43)	(9.84)
SMB(E)		0.000	−0.006		0.008	0.004
		(−0.01)	(−0.24)		(0.30)	(0.15)
HML(E)		−0.066	−0.083		−0.045	−0.057
		(−2.14)	(−2.84)		(−1.58)	(−2.02)
UMD(E)			0.011			0.009
			(0.28)			(0.23)
Adjusted R^2	96.20	96.51	96.97	96.45	96.83	97.02
Panel B: retail funds						
Alpha	0.038	0.070	−0.074	0.116	0.148	−0.002
	(0.16)	(0.31)	(−0.33)	(0.48)	(0.64)	(−0.01)
MktmRf(D)	0.753	0.809	0.827	0.756	0.808	0.823
	(29.28)	(33.61)	(34.50)	(29.15)	(32.71)	(33.17)
SMB(D)		0.268	0.267		0.253	0.251
		(8.76)	(8.93)		(8.05)	(8.13)
HML(D)		0.052	0.079		0.047	0.070
		(1.51)	(2.28)		(1.32)	(1.95)
UMD(D)			0.074			0.058
			(3.87)			(2.95)
MktmRf(E)	0.211	0.170	0.173	0.188	0.148	0.152
	(12.97)	(11.25)	(11.70)	(11.49)	(9.59)	(9.95)
SMB(E)		−0.045	−0.043		−0.049	−0.048
		(−2.89)	(−2.85)		(−3.03)	(−3.02)
HML(E)		−0.016	−0.013		−0.006	−0.008
		(−0.77)	(−0.62)		(−0.26)	(−0.38)
UMD(E)			−0.036			0.003
			(−1.65)			(0.15)
Adjusted R^2	92.34	94.10	94.37	91.81	93.46	93.66

significant only using value-weighted aggregate returns implying that smaller funds have no distinct preference for value or growth stocks. Momentum loadings are small and significant only for developed markets. The time-series fit of these models is quite good with adjusted R^2 of ~97% for the four-factor model.

In retail funds, the sum of market betas is again close to one. But there are other interesting differences in loadings from institutional funds. In aggregate (VW) or on average (EW), retail funds appear to tilt their portfolios toward smaller firms in developed markets and larger firms in emerging markets. This could be because information on small capitalization securities in emerging markets is hard to come by or because managers are concerned about liquidity problems. Compared with retail funds, the coefficient on HML is systematically higher in institutional funds (and frequently statistically significant), suggesting that value strategies are much more popular with plan sponsors than with retail investors. As with institutional funds, momentum loadings for retail funds are positive for developed markets but insignificant for emerging markets. Model fit, as measured by adjusted R^2, is lower in all specifications for retail funds than it is for institutional funds. This might be a reflection of less diversified portfolios for retail funds or it might reflect that retail funds hold some fraction of their assets in cash. Didier, Rigobon, and Schmukler (2011) report that the median number of stocks held by such retail funds is 95, suggesting that the former is the case.

Our primary interest is in alphas. The equal-weighted three-factor alpha for all institutional funds is 0.48% per quarter with a *t*-statistic of 2.04. Adding momentum shrinks alpha by about half to 0.23% and drops the *t*-statistic to 0.99. Alphas for value-weighted returns do not fare any better. The three-factor alpha is 0.04% per quarter with a *t*-statistic of 0.17 and adding momentum drops it to −0.14% (*t*-statistic = −0.60).

The alphas using equal-weighted retail fund returns are systematically lower than those for equal-weighted institutional funds. For instance, the four-factor alpha for institutional funds was 0.23% in Panel A; for retail funds, it is −0.07% (Panel B). Using value-weighted returns, retail funds fare a bit better than institutional funds with a four-factor alpha of −0.00% compared to −0.14%, but both have high standard errors.

We perform a number of unreported tests that are available upon request. We estimate factor models for retail funds using 36 months of data, rolling forward each month so as to generate a time series of factor loadings and alphas. Our interest in doing this stems from the possibility that funds may make risk systematic choices that co-vary with business cycles. Plots of time varying equal- and value-weighted alphas and factor loadings show no systematic patterns. However, rolling regressions are estimated relatively imprecisely and it is hard to disentangle the estimation error from true time variation in coefficients. Therefore, we also estimate conditional factor models in the spirit of Ferson and Schadt (1996). These models allow for conditional betas and unconditional alphas, or conditional betas and conditional alphas. For conditioning information, we employ the lagged dividend yield of the CRSP value-weighted index, a lagged measure of the term structure (defined as the difference between yields on long-term government bonds and the 3-month Treasury bill; data from Ibbotson), and a lagged measure of the default spread (defined as the difference between yields on BAA- and AAA-rated corporate bonds; data from Ibbotson). The alphas from conditional models are generally lower than those from the unconditional models reported in Table II with relatively large standard errors. In general, time variation and conditioning information does not change our broad inferences.

Regardless of the differences between institutional and retail funds, the data do not speak highly of the ability of the average fund to deliver alpha relative to even the simplest factor models. Even without accounting for momentum, simple adjustments for market exposure, size, and value effects show that the average fund is unable to deliver positive risk-adjusted returns. Moreover, as evidenced by the similar alphas for value-weighted returns, the industry as a whole does not deliver risk-adjusted excess returns.

It is useful to keep in mind that our alphas are generated from returns that are gross of fees. At the end of 2009, the average annual fee for institutional funds with a mandate of \$50 million was 0.78% and for retail funds, the average expense ratio was 1.50%. The value-weighted three-factor alpha for institutional funds is 0.037%, implying an annual alpha of 0.14%, well below the average annual fee for such funds. Similarly, the value-weighted three-factor alpha of retail funds (0.15%) implies an annual alpha of 0.60%, about half of the average annual fee. Thus, net of fees, investors in aggregate receive alphas that are at best zero, but most likely negative, consistent with the idea that if there are rents to be had, they are extracted by fund managers.

3.3 THE DISTRIBUTION OF PERFORMANCE

The inability of the average fund, or of the industry as a whole, to deliver alpha does not mean some funds cannot and do not do so. No active fund wants to regard itself as average and no active investor wants to think of herself as picking the average fund. More importantly, the cross-sectional distribution of performance across funds is critical to economic questions of market efficiency and active management.

In Table III, we tabulate various percentiles of alphas based on individual fund regressions for institutional and retail funds. We also show the percentage of alphas that are positive (or negative) and statistically significant. The former gives us a sense of the overall distribution, whereas the latter gives us a sense of the statistical significance of these alphas.

Ideally, we would like to estimate regressions for each fund and examine the full distribution of coefficients. But this requires a long time series of returns for both dead and live funds. We, therefore, require a minimum number of observations before we estimate the factor model regressions for individual funds. As dead funds are more likely to underperform, by requiring a minimum time series of returns, we induce an upward bias in the distribution of alphas.

For retail funds, we require a minimum of thirty-six monthly returns to estimate factor models. As our returns data for institutional funds are quarterly, we require twenty quarterly returns.[4] These restrictions reduce the number of institutional (retail) funds from 1,218 (1,019) to 907 (849), a truncation of about 26% (17%). Thus, the upward bias induced by truncation is similar and large for both samples. We find that the average annualized gross returns of mutual funds with <36 months of data are 5.93%, whereas those with more than 36 months of data are 10.10%. The bias is less pronounced in the sample of institutional funds—average annualized gross returns of institutional funds with <twenty quarters of data are 10.16%, whereas those with more than twenty quarters of data are 11.51%. We bear this in mind when examining the tails of the distribution.

The median three- and four-factor model alphas for institutional and retail funds are similar in magnitude to those using equal-weighted returns in Table II. The tails of the distribution are far more interesting. The 95th percentile of institutional fund four-factor alphas is an impressive 1.97% per quarter. For retail funds it is 1.53%. By way of comparison, Kosowski *et al.* (2006) report that the annual four-factor alpha of the 95th percentile of US domestic equity mutual funds is 1.3%. Using a 5% level of significance as a cutoff criterion and again focusing on four-factor alphas, 7.94% of institutional funds (7.54% of retail funds) have positive and significant alphas. Some of this right skewness could be due to the truncation discussed above. But regardless, the data suggest that the right tail of the distribution

4 A four-factor model requires us to estimate eight parameters, not including the intercept, so a minimum number of observations is necessary to generate meaningful estimates.

Table III. Distribution of individual fund alphas

We report percentiles of alphas from time-series regressions of individual funds on factor models. We use three combinations of factors—the one-factor model corresponds to just the market factor (MktmRf), the three-factor model corresponds to the three Fama and French (1993) factors (MktmRf, SMB, and HML), and the four-factor model adds the momentum factor to the mix (MktmRf, SMB, HML, and UMD). Separate factors are included for developed and emerging markets. Only institutional funds with more than twenty quarters of data and retail funds with more than 36 months of data are included in the estimation. Regression frequency is quarterly for institutional funds and monthly for mutual funds. All alphas are in percent per quarter. The sample period is 1991–2009 for institutional funds and 1987–2009 for retail funds.

	Institutional funds			Retail funds		
	One-factor	Three-factors	Four-factors	One-factor	Three-factors	Four-factors
Percentiles of alphas						
5	−0.753	−1.112	−1.414	−1.286	−1.415	−1.641
10	−0.426	−0.752	−1.095	−0.808	−0.875	−1.198
25	−0.006	−0.266	−0.441	−0.333	−0.361	−0.563
50	0.436	0.228	0.060	0.167	0.145	−0.040
75	0.865	0.837	0.584	0.685	0.779	0.593
90	1.453	1.719	1.293	1.147	1.440	1.107
95	1.964	2.367	1.971	1.668	1.893	1.530
Percentage of negative and significant alphas						
at 10%	4.52	6.73	9.48	6.71	9.89	14.49
at 5%	1.98	3.53	5.40	4.00	6.60	8.60
at 1%	0.44	0.33	0.88	0.82	2.83	4.12
Percentage of positive and significant alphas						
at 10%	19.85	17.53	11.91	14.13	16.84	10.84
at 5%	12.13	11.91	7.94	8.60	12.25	7.54
at 1%	4.74	5.51	2.87	2.59	4.95	2.47
Number of funds		905			849	

generates impressive performance. Ignoring issues of model error and statistical inference for the moment, and focusing solely on magnitudes, a quarterly alpha between 1 and 2% cannot be ignored.[5] From an economic perspective, these alphas justify the expense ratios reported in Table I. In the next section, we attempt to understand the source of this abnormal performance.

[5] As a specification check, we also include lagged values of factors in the models in Tables II and III. We do not report values in tables but in the vast majority of cases, there are only minor variations in alphas.

3.3.a. Country versus security selection

Consider a simple performance attribution of fund returns. Letting w_{ict} (w_{bct}) be the weights in country c at time t by fund i (benchmark b), R_{ict} be the returns generated by fund i in country c at time t, and R_{bct} be the benchmark market return in country c at time t, it is easily verified that:

$$\begin{aligned} R_{it} - R_{bt} &= \sum_c w_{ict-1} R_{ict} - \sum_c w_{bct-1} R_{bct} \\ &= \sum_c (w_{ict-1} - w_{bct-1}) R_{bct} + \sum_c w_{ict-1}(R_{ict} - R_{bct}). \end{aligned} \tag{2}$$

The 1st term on the right-hand side is the contribution to excess returns due to country selection, whereas the 2nd term is the contribution due to security selection. To determine whether the alphas in the right tail come from security selection by fund managers, we calculate this security selection component for fund i as follows:

$$\mathrm{SR}_{it} = R_{it} - \sum_c w_{ict-1} R_{bct}, \tag{3}$$

where we use the country index return from MSCI as R_{bct} and weights reported by the funds themselves as w_{ict-1}.

We then estimate time-series factor models on this security selection return. Country weights for institutional funds are available only from 1996. Even after this time period, this information is missing for many funds. Further imposing the requirement of at least twenty quarters shrinks the sample to 345 funds. For retail mutual funds, we can only estimate models for 414 funds. Table IV shows the distribution of results for various percentiles of alphas and/or excess returns. We also show the percentage of positive/negative and significant alphas at various levels of significance.

Once again, the data show some right skewness with the right tail containing some large alphas. The 95th percentile of four-factor alphas for institutional (retail) funds is 1.43% (1.67%). Moreover, at the 5% level of significance, there are 8.12% of institutional funds with significantly positive alphas using the four-factor model. The equivalent percentage for retail funds is lower, at 5.56%.

On the surface, these statistics suggest some successful stock picking ability. But we urge caution for two reasons. First, the left tail is likely truncated because of data requirements. Second, with 345 institutional funds and 414 retail funds in the sample, the actual number of funds with superior stock selection ability is far from overwhelming. Assessing statistical significance at the 5% level, there are approximately twenty-five each

Table IV. Security selection

We report percentiles of alphas from time-series regressions of individual funds on different factor models. We use three combinations of factors—the one-factor model corresponds to just the market factor (MktmRf), the three-factor model corresponds to the three Fama and French (1993) factors (MktmRf, SMB, and HML), and the four-factor model adds the momentum factor to the mix (MktmRf, SMB, HML, and UMD). Separate factors are included for developed and emerging markets. We calculate synthetic benchmark returns as the product of actual fund country weights and country index returns. The dependent variable is the difference between the actual fund return and this synthetic benchmark return and is equal to the component of the return due to security selection at the country level. Only institutional funds that have more than twenty quarters of data and retail funds with more than 36 months of data are included in the factor models. Regression frequency is quarterly, and alphas are reported in percent per quarter. The sample period is 1996–2009 for institutional funds and 1991–2009 for retail funds.

	Institutional funds			Retail funds		
	One-factor	Three-factors	Four-factors	One-factor	Three-factors	Four-factors
Percentiles of alphas						
5	−0.459	−0.748	−0.599	−1.686	−1.644	−1.821
10	−0.255	−0.489	−0.468	−1.187	−1.075	−1.246
25	0.011	−0.158	−0.179	−0.651	−0.575	−0.578
50	0.277	0.312	0.224	−0.131	−0.041	−0.081
75	0.693	0.784	0.557	0.414	0.501	0.487
90	1.116	1.511	1.143	0.927	1.121	1.061
95	1.480	2.112	1.433	1.224	1.521	1.672
Percentage of negative and significant alphas						
at 10%	3.48	4.35	4.35	12.80	13.04	13.53
at 5%	2.03	1.74	1.74	7.97	8.45	7.73
at 1%	0.87	0.29	0.29	2.17	2.17	2.42
Percentage of positive and significant alphas						
at 10%	18.26	24.06	14.20	8.45	9.90	8.45
at 5%	9.57	13.62	8.12	4.59	7.49	5.56
at 1%	3.77	4.06	2.61	1.21	3.38	3.38
Number of funds		345			414	

(8.1% × 345 for institutional funds and 5.56% × 414 for retail funds) institutional and retail funds with apparently superior stock selection ability. Of course, subtracting stock selection from realized returns, the "left-over" portion is country selection. Next, we turn to the issue of how funds choose countries.

3.3.b. Country momentum

There are a myriad of reasons that might cause a fund to under- or overweight a country relative to market capitalization weights. For instance, risk management reasons might cause portfolio managers to cap a portfolio's exposure to a country. More directly, country bets on performance might cause a fund to place a larger or smaller weight on a country. One candidate that lends itself to empirical testing is past country returns (country momentum). To test whether funds select countries based on past returns, we adopt the method of Grinblatt, Titman, and Wermers (1995) and calculate the following measure of trading on country momentum:

$$\mathrm{ITM}_{it}(k,l) = \sum_{c}(w_{ict} - w_{ict-1})(R_{bct-k} - R_{bt-k}), \tag{4}$$

where R_{bct-k} is the country index return for country c at time $t-k$, and R_{bt-k} is the average return of the MSCI benchmark for fund i. This measure of country momentum is, thus, a cross-product of country weight changes and previous period country returns.

As the reporting of country weights for institutional funds is quarterly, we use $l=1$ for both retail and institutional funds. We allow k to take on values of zero through four ($k=4$ corresponds to an across-country momentum strategy based on annual returns). Different values of k allow us to examine the importance of different past returns on the decision to change portfolio holdings. Generally, when $k=0$, the return is contemporaneous with weight changes. The consequence is that, strictly speaking, we cannot interpret our measure as reflecting momentum trading. This is certainly true for retail funds where weights are identified at month end. However, for institutional funds, there is potentially an observation delay; weights reported at the end of the quarter might reflect weights during the quarter. Therefore, we opt to report results for $k=0$ for both sets of funds so that the reader has full information.

Badrinath and Wahal (2002) show that the entry and exit of countries to and from the portfolio can distort our measure in Equation (4). Across all fund-quarter observations, entry and exit constitute 6.4 and 6.2% of country weight changes. We, therefore, report separate estimates for "entered countries" (when a fund first establishes a position in country c), "exited countries" (when a fund sells its entire position), "adjustments to existing countries" (when the country weight is nonzero at the beginning and end of period l) and for the entire portfolio.

We calculate the ITM_{it} measure for each fund in each period. We then calculate the cross-sectional mean in each period. We finally report the time-series means and medians of these measures for institutional funds (Panel A) and retail funds (Panel B) in Table V. As ITM measures across periods are unlikely to be independent, we calculate t-statistics using a Fama and MacBeth (1973) approach, generating standard errors from the

Table V. Changes in country weights

We report the Grinblatt, Titman, and Wermers (1995) measure of momentum trading, $\mathrm{ITM}_{it}(k,l)=\sum_c(w_{ict}-w_{ict-1})(R_{bct-k}-R_{bt-k})$. We use $l=1$, which implies that weight changes are calculated over quarterly intervals. We calculate separate values for each fund's entire portfolio, for countries that it entered during the quarter ($w_{ict-1}=0$), for countries that it exited during the quarter ($w_{ict}=0$), and for countries where the fund was invested during the entire period. We first calculate the ITM_{it} measure for each fund in each period and then calculate the cross-sectional mean in each period. We finally report the time-series means and medians of these measures. Fama and MacBeth (1973) *t*-statistics are reported in parentheses next to means. The percentage of funds with positive time-series mean estimates of the ITM measure is reported in parentheses next to medians. The number of funds that are used in the calculation is reported next to the category of funds. All ITM measures are multiplied by 100. The sample period is 1996–2009 for institutional funds and 1991–2009 for retail funds.

	Entire portfolio		Entered countries		Exited countries		Adjustments to existing countries	
	Mean (*t*-stat)	Median (% positive)	Mean (*t*-stat)	Median (% positive)	Mean (*t*-stat)	Median (% positive)	Mean (*t*-stat)	Median (% positive)
Panel A: institutional funds								
Number of funds = 751								
$k=0$	0.675	0.541	0.064	0.047	0.011	0.009	0.600	0.478
	(10.80)	(90.50)	(6.06)	(75.10)	(1.43)	(49.30)	(10.74)	(91.70)
$k=1$	−0.034	−0.021	0.033	0.025	−0.020	−0.017	−0.047	−0.015
	(−0.95)	(46.10)	(3.83)	(69.80)	(−2.21)	(34.90)	(−1.40)	(43.30)
$k=4$	−0.144	−0.157	0.106	0.113	−0.095	−0.113	−0.155	−0.172
	(−1.61)	(35.60)	(5.16)	(73.50)	(−5.39)	(23.70)	(−1.86)	(32.50)
Panel B: retail funds								
Number of funds = 567								
$k=0$	0.545	0.445	0.048	0.031	0.020	−0.003	0.478	0.379
	(9.68)	(96.30)	(4.97)	(78.30)	(1.81)	(52.20)	(9.88)	(96.10)
$k=1$	0.018	−0.017	0.031	0.025	−0.009	−0.013	−0.004	−0.022
	(0.50)	(52.90)	(2.98)	(75.10)	(−1.11)	(38.80)	(−0.11)	(50.10)
$k=4$	−0.003	−0.077	0.102	0.096	−0.071	−0.078	−0.034	−0.095
	(−0.03)	(47.60)	(4.91)	(83.20)	(−4.14)	(20.10)	(−0.40)	(40.90)

time-series averages. We also report the percentage of funds with positive time-series mean estimates of the ITM. All ITM measures are multiplied by 100.

Both institutional and retail funds establish positions in countries after positive returns. For example, for institutional funds, the average value of $\mathrm{ITM}_{it}(k,l)$ at $k=4$ is 0.106 with a *t*-statistic of 5.16. Similarly, for retail funds, the corresponding average is 0.102 with a *t*-statistic of 4.91. The averages for developed and emerging market funds are reliably positive. These funds also exit after positive returns. The average $\mathrm{ITM}_{it}(k,l)$ at

$k = 4$ for institutional (retail) funds is -0.095 (-0.071) with a t-statistic of -5.39 (-4.14). Finally, adjustments to countries to which funds already have exposure do not appear to be reliably related to past country returns. As the momentum estimates for entry are similar in magnitude to the contrarian estimates for exit, they essentially offset each other. The net effect, combining entry, exit, and adjustments, is that entire portfolio country weight changes are unrelated to past country returns. The implication is that the choice of investing in a country is not related to country momentum.

3.4 TRUE ALPHAS: LUCK VERSUS SKILL

The evidence thus far shows little superior performance on average or in aggregate, but substantial alphas in the right tail. Of course, it is entirely possible that these alphas are not true in the sense that they are not a reflection of skill but merely due to luck. We use simulation approaches to disentangle the two. We use the approach of Kosowski *et al.* (2006), as modified by Fama and French (2010), bootstrapping returns under the null of zero alphas with 5,000 simulation draws. We base inference on the entire cross-section of simulated t-statistics. We prefer to conduct inference based on t-statistics rather than alphas because, by definition, t-statistics control for the precision of the alpha. In doing so, we sample the fund and factor returns jointly to account for common variation in fund returns not accounted for by factors.

Table VI shows percentiles of actual and average simulated t-statistics as well as the percentage of simulation draws that produce a t-statistic greater than the corresponding actual value. If we conduct statistical inference based on a 5% cutoff, evidence of skill would come from fewer than 5% of the simulations being greater than actual. The results are presented only for four-factor alphas. As in the calculation of individual fund alphas, we require a minimum of twenty quarters (36 months) of observations for institutional (retail) funds.

Consider institutional funds. At the 95th percentile, the percent of simulated t-statistics is greater than the actual in 8.5% of cases. At the 99th percentile, the evidence against skill is even stronger as 13.5% of simulated t-statistics are greater than the actual. Similarly, in retail funds, the percent of simulated t-statistics is greater than the actual in 7.7% of cases at the 95th percentile and in 2.9% of cases at the 99th percentile. By way of comparison, for US equity funds with an assets under management of at least \$5 million and using a four-factor model on gross returns, Fama and

Table VI. Skill versus luck in individual fund alphas

Performance is measured using the four-factor model, where the four factors are MktmRf, SMB, HML, and UMD. Separate factors are included for developed and emerging markets. The table shows percentiles of actual and average simulated *t*-statistics of alphas from these models. We also show the percentage of simulation draws that produce a *t*-statistic greater than the corresponding actual value. Only those institutional funds that have more than twenty quarters of observations and those mutual funds that have more than 36 months of observations are included in the estimation. Regression frequency is quarterly for institutional funds and monthly for mutual funds. The sample period is 1991–2009 for institutional funds and 1987–2009 for retail funds.

	Institutional funds			Retail funds		
	Actual	Simulated	%(Sim > Act)	Actual	Simulated	%(Sim > Act)
1	−2.63	−3.33	22.0	−3.33	−2.51	95.2
5	−2.00	−2.08	46.8	−2.43	−1.71	94.9
10	−1.56	−1.58	50.9	−1.83	−1.32	90.2
20	−0.92	−1.03	41.2	−1.21	−0.87	81.8
25	−0.69	−0.84	38.4	−0.98	−0.70	78.6
40	−0.20	−0.36	35.1	−0.40	−0.27	63.9
50	0.10	−0.08	31.1	−0.06	−0.02	54.5
60	0.41	0.19	27.8	0.34	0.24	38.2
75	0.98	0.66	19.4	0.95	0.66	22.3
80	1.18	0.84	18.3	1.15	0.83	19.4
90	1.85	1.35	11.0	1.74	1.28	11.3
95	2.44	1.80	8.5	2.22	1.66	7.7
99	3.56	2.86	13.2	3.29	2.42	2.9

French (2010) report that at the 95th percentile, there are 11.33% of cases in which simulated *t*-statistics are greater than actual.

The simulations show that there seems to be no evidence of skill in the right rail of the distribution. The implication is that the noticeably large alphas in the right tail in Table III are most likely there due to chance.

3.5 PERFORMANCE PERSISTENCE

Our measures of performance so far are all *ex post*. From a practical viewpoint, it is perhaps more important to assess whether there are *ex ante* rules based on past performance which can be used to select funds that have superior subsequent performance. In this section, we investigate persistence in fund returns.

Persistence tests can also be viewed as an alternative way to think about the alphas in the right tail; if funds simply have large alphas because of luck in 1 year, they should not continue to have large alphas in future years. The downside of persistence tests is that, as the ranking of funds is based on short-term performance, we may not find evidence of persistence because the allocation of funds to winners and losers may be based on noisy realized returns. However, as we will see shortly, we find significant patterns in three-factor alphas (which are substantially reduced by the addition of the momentum factor). These patterns allay the fears that we are sorting solely on noise.

We sort funds into deciles using benchmark-adjusted returns during a ranking period and examine returns in a postranking period. The 1st decile contains the worst-performing funds, and the last decile contains the best-performing funds. After sorting into performance deciles, we compute equal-weighted decile returns over subsequent evaluation horizons. We then roll forward in time, producing a nonoverlapping time series of concatenated returns. With these returns, we estimate factor models as described earlier. We use four different postranking horizons: one-quarter, 1st-, 2nd-, and 3rd year.

When the ranking period is 1 year, the postranking period ranges from 1 year to 3 years (portfolios are rebalanced annually). We examine returns in the 1st-, 2nd-, and 3rd year of the postranking period. When the ranking period is less than a year, we use one-quarter returns for ranking institutional funds and 1-month returns for ranking retail funds; portfolios are rebalanced quarterly and monthly, respectively. Table VII shows the alphas from these models for various combinations of funds and factors.

There is evidence of short-term persistence for both institutional and retail funds. Funds in the top decile continue to generate high returns in the subsequent quarter for both groups. For institutional (retail) funds, the next quarter four-factor alpha for the highest decile is 0.65% (0.40%) with a t-statistic of 1.76 (1.20). Beyond the 1st quarter, the estimates of three-factor alpha are high and statistically significant for the 1st postranking year for institutional funds (but not for retail funds). Indeed, in the 1st postranking year, the three-factor alpha of the top deciles of institutional funds is 1.18% per quarter (t-statistic = 2.38). However, incorporating momentum shrinks this alpha to 0.35% per quarter (t-statistic = 1.03). In retail funds, there is no evidence of persistence in the 1st postranking year, with or without including the mechanical effects of momentum. Moreover, looking beyond the 1st postranking year, there is no evidence of persistence, in institutional or retail funds, using three- or four-factor models.

Table VII. Persistence in performance

We sort funds into quintiles according to the benchmark-adjusted return during a ranking period and hold these portfolios for a postranking period. For institutional funds, when the ranking period is one quarter, the postranking period is also one quarter (portfolios are rebalanced quarterly). For mutual funds, when the ranking period is 1 month, the postranking period is also 1 month (portfolios are rebalanced monthly). For all funds, when the ranking period is 1 year, the postranking period ranges from 1 to –3 years (portfolios are rebalanced annually). The 1st quintile contains the worst-performing funds and the last quintile contains the best-performing funds. We use three combination of factors—the one-factor model corresponds to just the market factor (MktmRf), the three-factor model corresponds to the three Fama and French (1993) factors (MktmRf, SMB, and HML), and the four-factor model adds the momentum factor to the mix (MktmRf, SMB, HML, and UMD). Separate factors are included for developed and emerging markets. Regression frequency is quarterly for institutional funds and monthly for mutual funds. All alphas are in quarterly percent, and *t*-statistics are reported in parentheses next to alphas. The sample period is 1991–2009 for institutional funds and 1987–2009 for retail funds.

	Institutional funds			Retail funds		
	One-factor	Three-factors	Four-factors	One-factor	Three-factors	Four-factors
One quarter/1 month						
1	0.158	0.404	0.531	−0.809	−0.930	−0.644
	(0.48)	(1.05)	(1.33)	(−2.51)	(−2.76)	(−1.93)
3	0.164	0.280	0.194	−0.242	−0.283	−0.255
	(0.84)	(1.25)	(0.83)	(−0.98)	(−1.15)	(−1.02)
5	0.329	0.409	0.279	−0.093	−0.132	−0.221
	(1.55)	(1.73)	(1.15)	(−0.38)	(−0.56)	(−0.93)
7	0.424	0.514	0.212	0.169	0.219	−0.018
	(1.85)	(1.93)	(0.86)	(0.64)	(0.87)	(−0.07)
10	1.196	1.481	0.651	0.935	1.142	0.403
	(2.64)	(2.87)	(1.76)	(2.09)	(2.90)	(1.20)
10−1	1.038	1.078	0.119	1.744	2.072	1.047
	(1.99)	(1.85)	(0.27)	(3.55)	(4.33)	(2.79)
1st year						
1	0.017	−0.286	−0.107	−0.337	−0.350	−0.212
	(0.06)	(−0.92)	(−0.33)	(−1.28)	(−1.30)	(−0.75)
3	0.293	0.451	0.398	−0.233	−0.202	−0.305
	(1.35)	(1.83)	(1.53)	(−0.98)	(−0.87)	(−1.27)
5	0.430	0.616	0.475	−0.018	0.018	−0.168
	(1.92)	(2.38)	(1.78)	(−0.07)	(0.07)	(−0.68)
7	0.338	0.474	0.084	0.181	0.251	−0.083
	(1.36)	(1.64)	(0.33)	(0.64)	(0.94)	(−0.31)
10	0.747	1.189	0.353	0.482	0.671	−0.182
	(1.67)	(2.38)	(1.03)	(1.11)	(1.77)	(−0.53)
10−1	0.730	1.475	0.459	0.819	1.022	0.030
	(1.30)	(2.39)	(1.05)	(2.00)	(2.63)	(0.09)

(continued)

Table VII. Continued

	Institutional funds			Retail funds		
	One-factor	Three-factors	Four-factors	One-factor	Three-factors	Four-factors
2nd year						
1	0.753	0.573	0.591	0.272	0.152	−0.091
	(2.57)	(1.88)	(1.84)	(0.94)	(0.52)	(−0.31)
3	0.379	0.335	0.187	0.093	−0.012	−0.241
	(1.83)	(1.39)	(0.78)	(0.38)	(−0.05)	(−1.00)
5	0.440	0.406	0.227	0.147	0.094	−0.079
	(2.11)	(1.60)	(0.91)	(0.59)	(0.39)	(−0.32)
7	0.317	0.231	0.055	0.015	0.030	−0.258
	(1.59)	(0.96)	(0.23)	(0.06)	(0.12)	(−0.99)
10	0.589	0.904	0.564	0.115	0.381	−0.064
	(1.69)	(2.22)	(1.50)	(0.32)	(1.18)	(−0.20)
10−1	−0.163	0.331	−0.027	−0.157	0.229	0.028
	(−0.43)	(0.87)	(−0.08)	(−0.54)	(0.79)	(0.09)
3rd year						
1	0.746	0.661	0.448	0.187	0.168	−0.217
	(2.81)	(2.13)	(1.53)	(0.64)	(0.60)	(−0.78)
3	0.468	0.347	0.157	0.049	0.028	−0.334
	(2.16)	(1.37)	(0.64)	(0.18)	(0.10)	(−1.24)
5	0.327	0.325	0.119	0.280	0.290	0.103
	(1.40)	(1.17)	(0.46)	(1.10)	(1.17)	(0.41)
7	0.288	0.252	0.122	0.182	0.269	−0.019
	(1.30)	(0.98)	(0.46)	(0.65)	(1.03)	(−0.07)
10	0.214	0.420	0.140	−0.093	0.116	−0.287
	(0.72)	(1.19)	(0.41)	(−0.29)	(0.40)	(−1.00)
10−1	−0.531	−0.241	−0.307	−0.280	−0.052	−0.071
	(−1.79)	(−0.66)	(−0.82)	(−1.09)	(−0.19)	(−0.25)

4. Conclusions

Investors access global equity markets using (retail) mutual funds and/or through institutional funds. Using factor models that include size, value, and momentum, we find little evidence of superior performance in actively managed funds. This is true on average and in aggregate. The tails contain some large alphas, but here the superior performance appears to come from very few funds. Regardless, even in the tails, it is difficult to argue that superior performance comes from skill—there is little evidence of stock picking ability, and simulations suggest that funds in the right tail are largely there due to luck. Finally, there is virtually no evidence of persistence.

Could one have anticipated the above results from the rather voluminous literature on US domestic equity mutual funds (and, to some extent, hedge funds)? We believe such a presumption is premature. The efficiency of less-developed international markets is an open empirical question. Therefore, one cannot automatically assume that just because domestic equity managers are or are not able to deliver superior performance, that ex-USA and global investment vehicles would or would not be able to do so. It is to know this answer that we bring evidence to bear.

Appendix A: Factor Construction

A.1 DATA ISSUES

We obtain a time series of market and accounting information for a broad cross-section of global firms from Datastream. We start with an unconstrained universe of all firms on forty-four countries (excluding USA) in the MSCI All Country Index between 1991 and 2009.[6] This universe includes both live as well as dead stocks, ensuring that the data are free of survivorship bias. To this, we apply the following sequence of filters that are derived from the extensive data investigations conducted by Ince and Porter (2006), Hou, Karolyi, and Kho (2011), and Griffin, Kelly, and Nardari (2010):

(1) We require that stocks have data from Datastream and Worldscope. The former is the source for market data, whereas the latter contains necessary accounting data.
(2) We only retain issues that are listed as equity and require that they be from the firm's primary exchange. The latter requirement, along with another Datastream supplied field, serves to remove duplications.
(3) We eliminate all nonlocal firms with a Geography Group (GEOG) code different from the local market.
(4) When a security dies, the postdeath time series contains returns that are marked as zero, which we eliminate.
(5) We employ a text search algorithm to eliminate securities that are not common stock. This ensures that preferred stock, trusts, warrants, rights, REITS, closed-end funds, ETFs, and depository receipts (GDRs and ADRs) not caught by the above screens are eliminated from the data.

[6] We obtained the US factors directly from Ken French's website.

(6) We compute returns using the return index (which includes dividends) supplied by Datastream. As both return indexes and market capitalization are provided in local currency, we convert them to US$ equivalents using the conversion function built into Datastream.
(7) We set returns to missing for a stock when it rises by 300% or more during 1 month and drops by 50% or more the following month (or falls and subsequently rises). We also treat as missing returns greater than the top 0.1% or less than the bottom 0.1% of the returns of all stocks in a country over time.

After this data cleansing, we impose some basic data requirements to calculate factor returns. For a country to have a nonmissing factor for a particular month, we require at least five stocks per portfolio used to create the factor. For example, typical double-sorted factors that use a 3×2 sort require at least thirty stocks (six portfolios multiplied by five stocks).

A.2 FACTOR CONSTRUCTION

We use dollar-denominated returns throughout our analysis, which presumes that investors can costlessly hedge deviations from purchasing power parity or ignore deviations. To establish size breakpoints, we sort all stocks in a country into two groups, small (S) and big (B), based on their market capitalization as of June of year t. These values are used to generate breakpoints corresponding to small (S) and big (B) portfolios. The small portfolio corresponds to the 25th percentile. Book-to-market ratios are computed by dividing book value for the fiscal year ending in year $t-1$ by market capitalization at the end of December of year $t-1$. These values are used to generate breakpoints corresponding to low (L), medium (M), and high (H) portfolios. The low and high portfolios correspond to the 30th and 70th percentile, respectively, with medium being between those two. For momentum portfolios, stocks in each country are sorted into terciles based on returns from $t-12$ to $t-2$. As with value, stocks are sorted into loser (L), neutral (N), and winner (W) portfolios, again using the 30th and 70th percentiles as breakpoints.

For each size, value, and momentum portfolio, we compute monthly value-weighted returns from July of year t through June of year $t+1$. We compute country-specific factor returns as the difference in average returns across size groups. After computing the above country-specific factor returns, we generate developed and emerging market factors by combining country-specific factors using each country's total market capitalization at the end of the previous year.

References

Badrinath, S. G. and Wahal, S. (2002) Momentum trading by institutions, *Journal of Finance* **57**, 2449–2478.

Bekaert, G. and Harvey, C. R. (2002) Research in emerging markets finance: Looking to the future, *Emerging Markets Review* **3**, 429–448.

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

Blake, D. and Timmermann, A. (1998) Mutual fund performance: Evidence from the UK, *Review of Finance* **2**, 57–77.

Bollen, N. P. B. and Busse, J. A. (2005) Short-term persistence in mutual fund performance, *Review of Financial Studies* **18**, 569–597.

Busse, J. A., Goyal, A., and Wahal, S. (2010) Performance and persistence in institutional investment management, *Journal of Finance* **65**, 765–790.

Busse, J. A. and Tong, Q. (2012) Mutual fund industry selection and persistence, *Review of Asset Pricing Studies* **2**, 245–274.

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

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

Cohen, R. B., Coval, J. D., and Pastor, L. (2005) Judging fund managers by the company they keep, *Journal of Finance* **60**, 1057–1096.

Cohen, R. B., Polk, C., and Silli, B. (2010) Best ideas. Working paper, London School of Economics.

Cremers, K. J. M., Ferreira, M., Matos, P., and Starks, L. (2011) The mutual fund industry worldwide: Explicit and closet indexing, fees, and performance. Working paper, University of Notre Dame, available at SSRN: http://ssrn.com/abstract=1830207, accessed April 17, 2013.

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

Cumby, R. E. and Glen, J. D. (1990) Evaluating the performance of international mutual funds, *Journal of Finance* **45**, 497–521.

Del Guercio, D. and Tkac, P. A. (2002) The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds, *Journal of Financial and Quantitative Analysis* **37**, 523–557.

Didier, T., Rigobon, R., and Schmukler, S. (2011) Unexploited gains from international diversification: Patterns of portfolio holdings around the world, *Review of Economics and Statistics*. forthcoming.

Droms, W. G. and Walker, D. A. (2001) Performance persistence of international mutual funds, *Global Finance Journal* **12**, 237–248.

Fama, E. F. and French, K. R. (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* **33**, 3–56.

Fama, E. F. and French, K. R. (1998) Value versus growth: The international evidence, *Journal of Finance* **53**, 1975–1999.

Fama, E. F. and French, K. R. (2010) Luck versus skill in the cross section of mutual fund returns, *Journal of Finance* **65**, 1915–1947.

Fama, E. F. and French, K. R. (2011) Size, value, and momentum in international stock returns, *Journal of Financial Economics* **105**, 457–472.

Fama, E. F. and MacBeth, J. D. (1973) Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* **81**, 607–636.

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

Ferson, W. and Schadt, R. (1996) Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* **51**, 425–462.

Goyal, A. and Wahal, S. (2008) The selection and termination of investment managers by plan sponsors, *Journal of Finance* **63**, 1805–1847.

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

Griffin, J. M., Kelly, P. J., and Nardari, F. (2010) Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets, *Review of Financial Studies* **23**, 3225–3277.

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

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

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.

Investment Company Institute. (2010) *2010 Investment Company Fact Book*, Investment Company Institute, Washington, D.C.

Kacperczyk, M. and Seru, A. (2007) Fund manager use of public information, *Journal of Finance* **62**, 485–528.

Kacperzyck, M., Sialm, C., and Zheng, L. (2005) On the industry concentration of actively managed equity mutual funds, *Journal of Finance* **60**, 1983–2011.

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

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

Khorana, A., Servaes, H., and Tufano, P. (2009) Mutual fund fees around the world, *Review of Financial Studies* **22**, 1279–1310.

Kosowski, R., Timmermann, A., Wermers, R., and White, H. (2006) Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* **61**, 2551–2595.

Panetta, F. and Cesari, R. (2002) The performance of Italian equity funds, *Journal of Banking and Finance* **26**, 99–126.