Does Size Matter?

Assets under management a questionable criterion.

Gregory C. Allen

After performance, assets under management, or AUM, is probably the most common criterion institutional investors use to narrow down a universe of investment products. The reasons for this are straightforward.

First, AUM is an easy number to collect and compare. Like performance, it is standard across a broad range of investment products and asset classes, and thus enjoys decent coverage in the commercially available manager search databases. This fact allows a sponsor employing a minimum AUM screen to be highly inclusive in the early stages of a search. It also allows a sponsor to eliminate a broad crosssection of managers using one single indisputable number. This makes for a very efficient process that is also likely to be seen as fair.

The second reason AUM is such a popular criterion is that asset management firms with larger asset bases are generally seen as less risky. This is true for a variety of reasons. Larger firms tend to be better capitalized; they typically carry higher levels of insurance; they have more resources to devote to back-office functions including reporting and compliance; and perhaps most important, they have other large institutional clients. This last point allows a sponsor to significantly reduce maverick risk, the risk of being simultaneously wrong and alone.

One question rarely asked is whether the use of AUM as a screening criterion is likely to have an impact, positive or negative, on return. I examine the historical impact of assets under management on performance using a robust database that includes over 5,000 institutional investment products. After cleaning up the data, and segregating the products by asset class and style, I examine the differences in excess returns across different-sized products by grouping

Gregory C. Allen
is the president and director of research at Callan Associates, Inc., in San Francisco.
allen@callan.com

them into quartiles based on AUM. I then use a simple Student-t test to determine if any differences in excess return observed across size quartiles are persistent and statistically significant. Finally, I examine the impacts of survivorship bias and back-fill bias on the conclusions of the analysis.

My findings are not surprising, except perhaps in their consistency and their magnitude. In the more liquidity-constrained asset classes (small-cap equities and high-yield bonds), the price of avoiding maverick risk by choosing larger managers has historically been significant and quite high. In the less liquidity-constrained asset classes (large-cap equities and investment-grade fixed-income), there has been a much smaller penalty for selecting larger managers.

Overall, the empirical findings support the idea that any meaningful reduction in risk in an efficient market generally comes at the expense of return.

DATASET

The dataset is drawn from a broad database of institutional separate account products built and maintained by Callan Associates, Inc., an institutional investment consulting firm. This database represents virtually all the products that have been marketed to U.S. tax-exempt institutional investors over the last 20 years.

The rules established by Callan over 20 years ago were designed to allow the database to support a performance measurement and reporting process for large complex institutional portfolios. Because Callan has remained the steward of this database since its inception, virtually all the data remain intact as originally reported. This includes data on non-surviving or defunct asset management products and firms.

The rates of return in this database are reported gross of fees, and represent the composite returns for institutional separate account products or commingled funds. The assets under management numbers represent the total market value of assets managed in each product at each quarter end date. Mutual funds are not included in this universe as their returns are reported net of fees, and they are often managed to satisfy objectives different from the needs of institutional investors.

No attempt has been made to audit the rate of return or assets under management data in this dataset, but the data have been subjected to a number of reliable algorithms designed to highlight inconsistencies and catch data entry errors. In the data collection process, managers were also given the opportunity to review and correct the entire return history for all products every quarter. In general, returns since about 1990 are asserted to be AIMR- (now GIPS-) compliant.

The data I use in this analysis span the 20-year period ended June 30, 2006, including 4,992 actively managed products, all with at least a three-year AUM and performance record. As of June 30, 1986, the beginning of the first quarter in the analysis, the dataset consisted of 814 products. A total of 4,178 products were created over the analysis period, and a total of 1,498 products stopped reporting data (approximately one-third of the dataset). Only 488 of the original 814 products survived the entire period.

Exhibit 1 shows some of the key demographics of the dataset broken down by asset class and style. U.S. equity products are separated into capitalization ranges and styles to account for the significant differences in return across market segments over the analysis period.

In compiling this broad dataset, I made every effort to identify all the composites in the Callan database over the last 20 years that have the minimum required level of data. A considerable amount of additional effort went into cleaning up the data to ensure that there are no gaps in a product's record. As a final check, all the AUM and return data for each product were evaluated cross-sectionally

EXHIBIT 1

Number of Products and Assets Under Management (in Millions)

Asset Class	Total Products	Products as of 6/30/86	Products as of 12/31/05	Inceptions	Non Survivors	Combined Assets as of 12/31/05
Large Cap US Equities	1,106	232	743	874	363	$ 2,353,712
Large Cap Core	390	58	261	332	129	$ 997,098
Large Cap Growth	360	81	231	279	129	$ 562,597
Large Cap Value	356	93	251	263	105	$ 794,016
Mid Cap US Equities	818	122	637	696	181	$ 891,841
Mid Cap Core	235	28	185	207	50	$ 249,987
Mid Cap Growth	319	55	246	264	73	$ 256,413
Mid Cap Value	264	39	206	225	58	$ 385,441
Small Cap US Equities	706	45	545	661	161	$ 550,857
Small Cap Core	177	7	127	170	50	$ 141,358
Small Cap Growth	268	25	197	243	71	$ 160,226
Small Cap Value	261	13	221	248	40	$ 249,272
Core Fixed Income	852	182	397	670	455	$ 954,461
Core Plus Fixed Income	113	16	85	97	28	$ 357,685
High Yield Fixed Income	236	50	181	186	55	$ 321,754
Developed Non US Equities	946	103	651	843	295	$ 2,072,224
Emerging Non US Equities	166	3	115	163	51	$ 298,070
Total	**4,943**	**753**	**3,354**	**4,190**	**1,589**	**$ 7,800,603**

(compared against other similar products) and through time to ensure reasonableness. Products with questionable or missing data were eliminated from the dataset.

EXCESS RETURN BY SIZE QUARTILE

The test to measure the impact of assets under management on performance is designed to be relatively straightforward and intuitive. That is, I want to compare the relative performance of the smallest managers in each asset class with the performance of their largest counterparts over the 20-year analysis period. To construct this test, I began by calculating a series of quarterly excess returns for every product in the dataset over its entire history.

EXHIBIT 2
Indexes Used in Excess Return Calculations

Asset Class	Index
Large Cap Core US Equities	Russell 1000
Large Cap Growth US Equities	Russell 1000 Growth
Large Cap Value US Equities	Russell 1000 Value
Mid Cap Core US Equities	Russell Midcap
Mid Cap Growth US Equities	Russell Midcap Growth
Mid Cap Value US Equities	Russell Midcap Value
Small Cap Core US Equities	Russell 2000
Small Cap Growth US Equities	Russell 2000 Growth
Small Cap Value US Equities	Russell 2000 Value
Core Fixed Income	Lehman Aggregate
Core Plus Fixed Income	Lehman Aggregate
High Yield Fixed Income	Lehman High Yield
Developed Non US Equities	MSCI EAFE (Net Dividends Reinvested)
Emerging Non US Equities	MSCI Emerging Markets (Free)

EXHIBIT 3
Assets Under Management Limits by Asset Class and Size Quartile

	Upper Limits (in Millions) Year-End 2005			
Asset Class	Smallest	Second	Third	Largest
US Large Cap Core	$ 238	$ 1,044	$ 3,388	$ 113,223
US Large Cap Growth	$ 127	$ 602	$ 2,165	$ 41,268
US Large Cap Value	$ 162	$ 719	$ 2,660	$ 78,614
US Mid Cap Core	$ 102	$ 268	$ 1,049	$ 20,126
US Mid Cap Growth	$ 87	$ 324	$ 1,056	$ 23,669
US Mid Cap Value	$ 146	$ 547	$ 1,845	$ 29,257
US Small Cap Core	$ 153	$ 628	$ 1,348	$ 17,448
US Small Cap Growth	$ 91	$ 461	$ 1,100	$ 7,524
US Small Cap Value	$ 186	$ 575	$ 1,526	$ 13,382
Core Fixed Income	$ 157	$ 496	$ 1,886	$ 77,685
Core Plus Fixed Income	$ 327	$ 1,566	$ 4,114	$ 63,876
High Yield Fixed Income	$ 198	$ 599	$ 1,699	$ 18,165
Developed Non US Equities	$ 147	$ 653	$ 2,675	$ 70,810
Emerging Non US Equities	$ 363	$ 820	$ 2,955	$ 22,702

Exhibit 2 details the index used for each asset class and style in calculating these excess returns. I use excess returns (rather than total returns) because that allows me to compare the effectiveness of active management across many different asset classes and styles.

In the next step of the analysis I separate all the products in each asset class or style into quartiles on the basis of AUM in each quarter. Thus, in a quarter where a particular asset class or style has 100 products, I put the 25 smallest products into the smallest quartile, the next 25 products into the second quartile, and so on. I go through this process for every asset class and every style for each quarter in the 20-year analysis period.

To provide some perspective on the makeup of products in the size quartiles, Exhibit 3 shows the upper AUM limits (in millions of dollars in AUM) for each size quartile for each asset class and style for the quarter ended December 31, 2005.

In the last step of the analysis, I calculate the mean excess return across the products within each size quartile for each asset class and style. For ease of interpretation, I then annualize these quarterly mean excess returns to yield the average annual mean excess return for each size quartile for each asset class and style. Exhibit 4 summarizes these annualized mean excess returns.

Note that this technique of using the mean excess return across all products in each quarter ensures that non-surviving products receive an appropriate weight in each quarter they were in existence. Combined with the fact that the underlying dataset includes substantially all the non-surviving products that reported performance to Callan during this period, this technique eliminates (or at least significantly mitigates) the impact of survivorship bias on the results.[1]

The last three columns in Exhibit 4 represent the results of a difference in means test that compares the mean excess return of the smallest quartile and the largest quartile for each asset class and style.[2] The first column is the simple difference between the two annualized mean excess returns. The second column is the t-statistic from the difference in means test, which compares the two distributions of excess returns.

The final column represents the confidence level associated with the hypothesis that the mean annualized excess returns for the smallest and largest quartiles are in fact different. It is generally accepted by statisticians

EXHIBIT 4

Annualized Mean Excess Return by Asset Class and Size Quartile—Full Universe

Asset Class	Annualized Mean Excess Return: Smallest Quartile	Second Quartile	Third Quartile	Largest Quartile	Difference in Means	T-Stat.	Confidence Level
US Large Cap Core	0.96%	0.79%	0.76%	0.44%	0.52%	2.91	99.91%
US Large Cap Growth	2.75%	2.19%	1.92%	1.76%	0.99%	2.72	99.79%
US Large Cap Value	0.83%	0.27%	0.35%	0.03%	0.79%	3.48	99.96%
US Mid Cap Core	0.42%	1.09%	1.10%	-0.01%	0.43%	0.73	76.11%
US Mid Cap Growth	3.73%	2.81%	2.48%	1.42%	2.31%	3.49	99.97%
US Mid Cap Value	0.87%	0.21%	0.47%	-0.13%	1.00%	2.24	98.74%
US Small Cap Core	6.37%	4.61%	3.83%	2.91%	3.45%	5.31	100.00%
US Small Cap Growth	11.20%	7.57%	5.45%	4.54%	6.65%	9.69	100.00%
US Small Cap Value	3.52%	2.21%	1.27%	0.90%	2.62%	5.80	100.00%
Core Fixed Income	0.02%	0.12%	0.16%	0.20%	-0.17%	-4.51	100.00%
Core Plus Fixed Income	0.68%	0.83%	0.70%	0.64%	0.04%	0.20	58.71%
High Yield Fixed Income	1.69%	1.19%	0.70%	0.45%	1.24%	4.22	100.00%
Developed Non US Equities	3.75%	3.25%	3.06%	2.80%	0.95%	3.11	99.93%
Emerging Non US Equities	4.12%	2.89%	2.87%	2.73%	1.39%	1.95	97.44%

that a t-statistic higher than 1.96, or a confidence level higher than 97.5%, in this type of test represents a statistically significant result.

The results reveal that the difference in annualized excess return is positive (that is, the smallest products have outperformed the largest products) across every asset class and style except for core fixed-income. There are generally greater differences in excess return across the more illiquid asset classes (U.S. small-cap equity and high-yield fixed-income) and smaller differences across the more liquid asset classes (U.S. large-cap equity and U.S. fixed-income). Finally, with the exception of U.S. mid-cap core equity and U.S. core fixed-income, all the difference in means test results appears to be statistically significant.

Given the signs and the significance of these results, it would seem reasonable to conclude that the use of minimum AUM screens has generally had a negative impact on institutional investor returns.

IMPACT OF BACK-FILL BIAS

The use of the mean quarterly excess returns, combined with a database that includes non-survivors, largely corrects for the impact of survivorship bias on the results of this analysis. Survivorship bias, however, is only the first of two pervasive biases that tend to inflate the rates of return in self-reported manager databases.

The second bias, known as *back-fill or instant history* bias, stems from the fact that managers typically start reporting to consultant databases only after they have generated a solid three-year performance record. At that point they will usually submit the entire return history for the product. Since managers will typically market only products with superior relative performance, this results in a self-selection bias, in that many of the three-year records for the more poorly performing products never get reported. The exclusion of these records biases the sample observed in consultant databases upward relative to the returns generated by the population as a whole.

Unfortunately, the high correlation between early performance records and small assets under management makes this analysis particularly sensitive to the impact of back-fill bias. While it is difficult to quantify, it is safe to assume a significantly greater impact of back-fill bias on the results for the smallest AUM quartile than on the results for the largest AUM quartile. This implies that the difference in means test probably overstates the impact of AUM on performance.

To adjust for the effect of back-fill bias in this analysis, I use an aggressive approach: I simply eliminate the first three years of every product's performance record from the analysis. I then regroup the products into size quartiles using the same method used for the original dataset. Finally, I recalculate the annualized mean excess returns for each size quartile for the same 20-year period. Exhibit 5 is the analogue to Exhibit 4 for this truncated dataset.

The results shown in Exhibit 5 are not surprising. Eliminating the first three years from each product's performance history has a relatively small impact on the observed mean excess returns for the largest AUM quartiles. As expected, the smallest AUM quartiles are more significantly impacted by this adjustment.

The end result is that the statistical significance of the results of the difference in means tests drops for every asset class and style. The differences, however, remain positive for virtually all asset classes, and both positive and statistically significant for high-yield fixed-income and all three of the U.S. small-cap equity groups. This should not be surprising, as these asset classes are widely acknowledged to be the most capacity-constrained (and therefore the most significantly impacted by asset size).

The other interesting result is that for core fixed-income the relation between size and excess return is significant and negative. This suggests that size (and presumably the associated resources and trading expertise) may be an advantage in these deep liquid markets.

EXHIBIT 5

Annualized Mean Excess Return by Asset Class and Size Quartile—Truncated Universe

	Annualized Mean Excess Return						
Asset Class	Smallest Quartile	Second Quartile	Third Quartile	Largest Quartile	Difference in Means	T-Stat.	Confidence Level
US Large Cap Core	0.72%	0.68%	0.55%	0.43%	0.29%	1.48	93.05%
US Large Cap Growth	2.53%	2.26%	1.89%	1.78%	0.75%	1.87	96.92%
US Large Cap Value	0.40%	-0.10%	0.20%	0.00%	0.41%	1.66	95.15%
US Mid Cap Core	-0.40%	0.25%	1.03%	-0.47%	0.06%	0.09	53.59%
US Mid Cap Growth	2.29%	2.19%	1.66%	1.18%	1.11%	1.55	93.94%
US Mid Cap Value	0.07%	0.33%	0.28%	-0.47%	0.54%	1.11	86.65%
US Small Cap Core	4.45%	3.65%	3.32%	3.00%	1.45%	2.01	97.78%
US Small Cap Growth	8.47%	6.06%	4.16%	4.36%	4.10%	5.65	100.00%
US Small Cap Value	2.25%	1.68%	1.01%	0.95%	1.31%	2.49	99.36%
Core Fixed Income	0.02%	0.08%	0.15%	0.19%	-0.17%	-4.29	100.00%
Core Plus Fixed Income	0.83%	0.71%	0.67%	0.73%	0.10%	0.44	67.00%
High Yield Fixed Income	1.41%	0.99%	0.37%	0.50%	0.92%	2.59	99.52%
Developed Non US Equities	2.93%	2.62%	2.42%	2.63%	0.30%	0.92	82.12%
Emerging Non US Equities	2.14%	2.43%	2.81%	2.58%	-0.44%	-0.56	28.78%

EXHIBIT 6

Performance Comparison—Small versus Large AUM Portfolios—Truncated Universe

	Small AUM Portfolio			Large AUM Portfolio			
Asset Class	Excess Return	Tracking Error	XR/TE	Excess Return	Tracking Error	XR/TE	% Periods Small Beat Large
US Large Cap Core	0.59%	1.04%	0.56	0.55%	0.60%	0.92	31.88%
US Large Cap Growth	2.14%	1.95%	1.10	1.64%	1.56%	1.05	57.97%
US Large Cap Value	0.28%	1.30%	0.22	0.03%	1.24%	0.03	60.87%
US Mid Cap Core	-0.79%	1.99%	-0.40	-0.34%	1.66%	-0.20	43.48%
US Mid Cap Growth	2.82%	3.04%	0.93	1.22%	2.35%	0.52	81.16%
US Mid Cap Value	0.27%	1.63%	0.17	-0.11%	2.11%	-0.05	69.57%
US Small Cap Core	5.54%	2.67%	2.08	3.01%	1.40%	2.15	79.71%
US Small Cap Growth	9.29%	2.51%	3.69	5.41%	2.55%	2.12	97.10%
US Small Cap Value	3.16%	2.49%	1.27	1.38%	2.43%	0.57	84.06%
Core Fixed Income	0.05%	0.23%	0.20	0.19%	0.19%	1.02	14.49%
Core Plus Fixed Income	0.82%	0.70%	1.17	0.76%	0.67%	1.12	62.32%
High Yield Fixed Income	1.43%	1.29%	1.11	0.42%	1.04%	0.41	73.91%
Developed Non US Equities	2.82%	4.56%	0.62	2.55%	3.70%	0.69	72.46%
Emerging Non US Equities	0.83%	4.33%	0.19	2.42%	3.18%	0.76	17.46%

FROM STATISTICAL TO INTUITIVE

In order to provide an intuitive understanding of the significance of these statistical results, I conduct a final simple test using the truncated dataset. The test is designed to measure the difference in results that two investors with different philosophies about the potential impact of size might have experienced over the 20-year analysis period.

The first investor constructs a set of portfolios (one for each asset class and style) using only managers from the smallest assets under management quartiles. To simulate this experience, I assume this investor achieves the average return of the smallest products in each asset class and style for every quarter in the 20-year period. The second investor constructs a portfolio using only managers in the largest AUM quartiles. Consequently this portfolio achieves the average return of the largest products in each asset class and style for every quarter in the period.

Exhibit 6 shows some different perspectives on the results that these two investors would have achieved over the period.[3]

For each portfolio I show the annualized excess return, the annualized tracking error (the standard deviation of excess returns), and the ratio of these two (often referred to as the information ratio). As in the earlier results, the smaller AUM portfolios outperform the larger AUM portfolios in every asset class and style except core fixed-income and emerging markets equities.

Once these results are adjusted for risk, however, it becomes clear that most of the performance advantage of the smaller AUM portfolios comes from their systematically higher levels of risk than the larger AUM portfolios. This can be seen by comparing the information ratios; note that in many cases they are actually higher for the large AUM portfolios, despite their lower excess returns.

This suggests that the higher returns of the smaller portfolios (most notably in core plus fixed-income) represent a behavioral phenomenon, rather than a systematic trading or information advantage unique to small portfolios.

The notable exceptions to this are in small-cap equities and high-yield fixed-income, where the information ratios for the small AUM portfolios are generally equal to or higher than their larger counterparts. This suggests that in these less liquid asset categories, smaller products may enjoy some sort of systematic advantage.

CONCLUSIONS

Institutional investors can draw a number of practical lessons from this analysis. First, studies that use self-reported manager databases are inherently subject to an upward bias, particularly in the early histories of returns. Any analysis that attempts to measure the benefits of employing newer, or emerging, managers using this type of dataset is likely to be influenced by this bias.

Second, in capacity-constrained asset classes, all things held equal, there is a statistically significant negative relation between assets under management and performance. In my experience, this makes size unique among quantitative variables in its ability to explain relative performance across a broad crosssection of managers and products.

Finally, the results for core fixed-income, and to a lesser extent emerging market equities, suggest that there are asset classes and strategies where the advantages that come with additional resources can outweigh the negative impact of portfolio size. This last point underscores the inherent difficulty of the manager selection process.

While size may be the single most powerful quantitative variable in forecasting relative performance, it is only one of countless variables that ultimately influence the relative success of investment management products.

ENDNOTES

[1]Survivorship bias is a well-known bias that plagues the construction of manager databases. It occurs because in most asset classes, non-surviving products tend to predominantly be the poorer performers. To the extent that they are eliminated from the analysis, this tends to bias the observed performance of the average active manager upward. The technique I used to construct the database in this analysis allows for the inclusion of non-surviving products, thereby mitigating the impact of survivorship bias.

[2]The difference in means test is a variant of the Student-t-test known as a heteroscedastic t-test. The distinguishing assumption is that the populations from which the two samples are drawn have different variances. As a practical matter, the results of the heteroscedastic t-test are indistinguishable from those of the simpler two-sample t-test which assumes the underlying population variances are the same.

[3]Constructing a portfolio that consistently achieves the average return for each size quartile would be complicated in practice by the need to reorder the mix of managers to reflect their changes in AUM over time. This would inevitably create transaction costs that would reduce the excess returns below the results shown in Exhibit 6. I assume these costs would impact both the large AUM and small AUM portfolios equally, making the results comparable.