Toward an Optimal Domestic Large-Cap Equity Index

Choice of a benchmark is critical.

Hamish Seegopaul, Francis Gupta, and John Prestbo

Appropriate domestic large-cap equity benchmarks may play a key role in the performance of a plan's overall portfolio. For many plan sponsors, the greatest allocation to a single asset class is generally to domestic large-cap equities, so even a marginal difference in the performance of this asset class over a long period can result in significant differences in the value of the overall plan. Consequently, it is crucial that institutional investors get the right benchmark to represent this asset class.

For the domestic large-cap equity asset class, the benchmark decision becomes even more critical because it plays an important role in implementation of the allocation decision. In comparison to other asset classes, the benchmarks representing this asset class have historically proven to be notoriously hard to outperform. It is thus commonly the case that most investors tend to have a good portion of their allocation passively managed, i.e., indexed or closet indexed.[1]

The use of active managers within this asset class is for the sole purpose of generating excess returns or alpha over the agreed-upon benchmark. Given this quest for alpha, it becomes even more essential that the benchmark representing this asset class be efficient.

In other words, the passive return of this asset class's benchmark should be able to take advantage of all the return-enhancing opportunities within the asset class, and any overperformance by an active manager can be considered as an outcome of the ability or skill of the manager.

HAMISH SEEGOPAUL
is an assistant manager of Index Development at Dow Jones Indexes in Frankfurt, Germany.
hamish.seegopaul@dowjones.com

FRANCIS GUPTA
is director of Risk Management & Research at Dow Jones Indexes in Princeton.
francis.gupta@dowjones.com

JOHN PRESTBO
is editor and executive director of Dow Jones Indexes in Princeton.
john.prestbo@dowjones.com

EXHIBIT 1

Performance of Major Domestic Large-Cap Equity Indexes by Style—10 Years Ending 4Q2004

Domestic Large-Cap Equity Style Index	Ann. Return	Ann. S.D. (%)	Return/S.D.
Blend:			
S&P 500	12.07	17.77	0.68
Russell 1000	12.16	18.13	0.67
Wilshire Target 750	12.08	18.31	0.66
Value:			
Barra/S&P 500 Value	12.24	17.55	0.70
Russell 1000 Value	13.83	16.20	0.85
Wilshire Target LC Value	12.35	16.32	0.76
Growth:			
Barra/S&P 500 Growth	11.44	19.89	0.58
Russell 1000 Growth	9.60	22.56	0.43
Wilshire Target LC Growth	12.65	20.54	0.62

Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network.

EXHIBIT 2

Domestic Large-Cap Equity Institutional Product Composites with Ten Years of Performance—Ending 4Q2004

	Benchmark		
Large-Cap Style	S&P Indexes	Russell Indexes	Total
Blend	359 (95%)	20 (5%)	**379 (100%)**
Value	21 (15%)	120 (85%)	**141 (100%)**
Growth	6 (5%)	109 (95%)	**115 (100%)**
Total	**386 (61%)**	**249 (39%)**	**635 (100%)**

Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network gross of fees.

ROLE OF A BENCHMARK

Depending on the function, different benchmarks may have different characteristics. For our purposes, benchmark criteria as follows serve as the litmus test:[2]

Representative – Plan sponsors make their asset allocation policy decision at the asset class level. It follows that the benchmark chosen to represent each asset class should be a true representation of the opportunity set.

Measurable – To get a good idea of the opportunity cost and to evaluate active managers, the benchmark should be measurable.

Replicable – An investor should be able to replicate the returns of the benchmark passively.

Efficient – A benchmark should be efficient; i.e., it should be hard to outperform. For a benchmark to serve as an appropriate yardstick for active managers, it should be free of any inherent inefficiency. A benchmark whose performance can be easily enhanced by a transparent and obvious rule should incorporate that rule into its construction.

DOMESTIC LARGE-CAP EQUITY BENCHMARKS

Exhibit 1 shows the past performance of three families of domestic large-cap equity indexes over a ten-year period ending December 2004. Each family includes style indexes. We categorize performance in terms of three commonly accepted styles: blend, value, and growth.

For the ten-year period studied, the blend indexes exhibited very similar return and risk characteristics. Annualized performance ranges from 12.07% for the S&P 500 to 12.16% for the Russell 1000, a difference of 9 basis points. The S&P 500's standard deviation of 17.77% is the lowest, giving it the best return-risk trade-off.

There is a great deal more variability among the growth and value indexes, and here benchmark selection does become critical. While the blend indexes all base their component selection loosely on market capitalization, the determination of styles is more of a subjective task, and each index provider uses a different methodology. This approach is apparent in the return-risk ratios of the indexes; value indexes range from 0.70 to 0.85, and growth indexes from only 0.43 to 0.62.

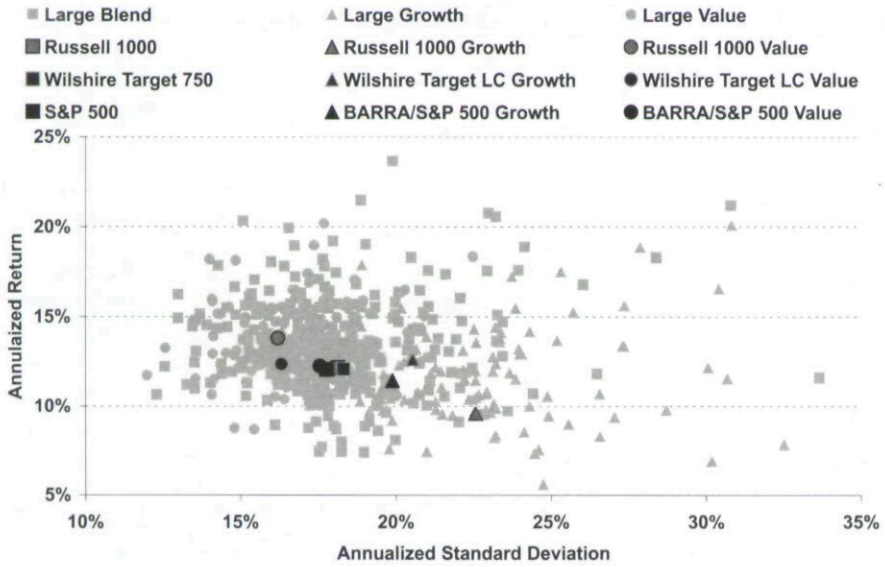
Further, similarities between providers in one style do not translate to another style. The annualized performances of the Russell and Wilshire large-cap indexes differ by only 148 basis points in the value space, but 302 basis points in the growth space. Over this ten-year period, the Russell 1000 Value and Wilshire LC Growth resulted in the best return-risk trade-offs.

DOMESTIC LARGE-CAP EQUITY INSTITUTIONAL COMPOSITES

Exhibit 2 presents the breakdown by benchmark of large-cap equity institutional composites in the Plan Sponsor Network (PSN) database. While the three blend indexes in Exhibit 1 have very similar return and risk characteristics, it is clear that in terms of benchmarking prod-

EXHIBIT 3

Outcome of Institutional Product Composites – Domestic Large-Cap Equity—10 Years Ending 4Q2004



Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network gross of fees.

ucts, the S&P dominates, with 95% of the 379 large-cap equity institutional composites benchmarked to the S&P 500. The remaining 5% of the composites are benchmarked to the Russell 1000. The number of composites benchmarked to the Wilshire is negligible.

The Russell style indexes dominate the growth and value space; 85% of the 141 large-cap value composites are benchmarked to the Russell 1000 value index, and 95% of the 115 growth composites are benchmarked to the Russell 1000 growth index. The remaining value and growth composites are benchmarked against the Barra/S&P 500 value and growth indexes, respectively. Again, the number of composites benchmarked to the Wilshire style indexes is negligible.[3]

Overall, 61% of the 635 domestic large-cap equity institutional composites with a ten-year history available in the Plan Sponsor Network database are benchmarked against S&P 500 indexes; the remaining are benchmarked against the Russell indexes.

ACTIVE MANAGER PERFORMANCE

Exhibit 3 displays the performance of the 635 composites in return-risk space. It also presents the performance of the various blend, value, and growth indexes shown in Exhibit 1.

The most striking observation to emerge from this exhibit is that while all the benchmarks lie near one another—more so for the blend indexes than the growth and value indexes—active managers span a much wider spectrum of return and risk outcomes. Another striking observation is that none of the managers produces negative returns, although this finding may be an outcome of the survivorship bias that is inherent in the data.[4]

Exhibit 4 takes the same institutional manager composites and presents their performance in the alpha tracking error space relative to their benchmarks. These graphs make it easier to compare the performance of active managers across benchmarks and across styles.

A summary of performance statistics is presented in the appendix.

A point worth noting in Exhibit 4 is the tracking error scales across the different styles. The tracking error scales go from 25% to 15% to 20% for blend, value, and growth managers, respectively. The value managers tend to cluster together, while growth managers tend to be more varied.

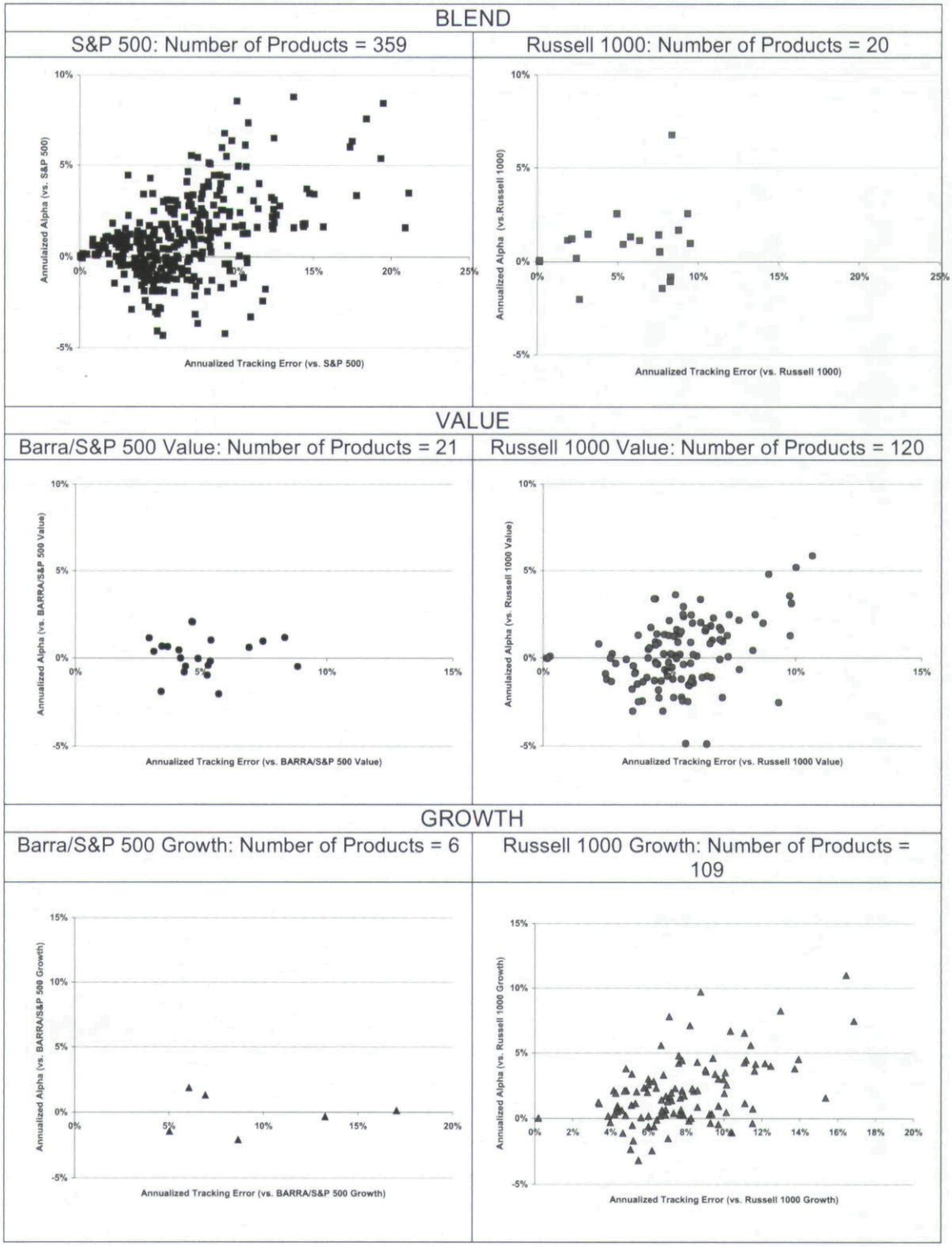
Exhibit 5 presents the percentage of winners versus the self-reported benchmark. The winners are determined using the average alpha (difference between product composite and benchmark quarterly returns) over the ten-year period.

These results reinforce those presented in Exhibit 1. There is little difference in selecting large-cap blend benchmarks, while benchmark selection for benchmarking growth and value managers is critical. Once again, there is also no continuity in providers across styles—the Russell Large Value index was by far the toughest to outperform, but the Russell Large Growth index was by far the easiest to outperform.

Other than the Russell 1000 value and the Wilshire growth indexes, product composites were able to outperform the benchmark indexes more than half the time. In fact, value composites were able to outperform two of the benchmark indexes at least 75% of the time. Product composites outperformed the Wilshire growth index by a more reasonable 34% of the time.

EXHIBIT 4

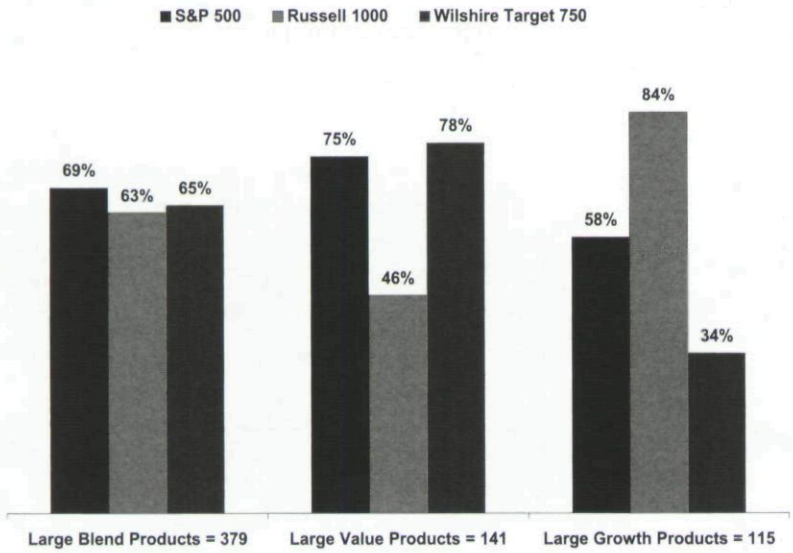
Product Composites Benchmarked Against Domestic Large-Cap Style Indexes—10 Years Ending 4Q2004



Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network gross of fees.

EXHIBIT 5

Product Composite Outperformance by Style—10 Years Ending 4Q2004



Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network gross of fees.

OPTIMAL DOMESTIC LARGE-CAP INDEX

These findings raise a question as to whether performance of a product can be attributed to the skill of the managers, or if the benchmarks may have some inefficiencies that active managers can exploit in a systematic way. The answers to both these questions are difficult, but analysis of the data seems to suggest that a combination of the best benchmarks by style may lead to an efficient domestic large-cap equity index.

For purposes of illustration, we create in Exhibit 6 an equally weighted large-cap index using the S&P 500, Russell 1000 value, and Wilshire LC growth benchmarks. We call this index the *optimal domestic large-cap benchmark* (ODLC).[5]

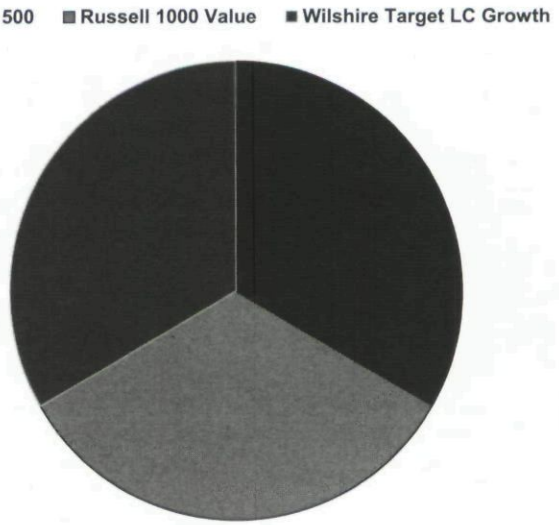
Exhibit 7 summarizes the performance of the 635 products with respect to their own performance benchmark, the S&P 500, and the ODLC. The results are encouraging. Over the ten-year period, 68% of the managers outperformed their own benchmark, and 73% outperformed the S&P 500, but only 52% outperformed the ODLC (a number one would expect if alpha were normally distributed around zero).

Across all large-cap domestic equity products, about the same number of products outperformed their respective benchmark and the S&P 500. Their alphas are also very similar (1.03% versus 1.10%). The average alpha with respect to the ODLC was only 0.31% (gross of fees). But, as one would expect, manager tracking errors were the lowest relative to their own benchmarks (6.58%) and climbed to 7.67% and 7.59% relative to the S&P 500 and the ODLC, respectively. Consequently, while the average information ratio with respect to the S&P 500 and the product's respective benchmark were similar (0.13 and 0.13), the average information ratio with respect to the ODLC was no different from zero (–0.02).

The finding that the average information ratio relative to an optimal benchmark is around zero verifies the optimality or efficiency of the benchmark. Even though the average alpha is positive (31 basis points gross of fees), the relatively high tracking error of the managers pulls the information ratio toward zero. And if we were to assume an average active management fee of 31 basis points, the average alpha would also be zero. This would make the time series of manager returns consistent with a belief that over the long run, relative to an efficient index, the average manager and therefore the average investor experiences an excess return of zero. Investors with pos-

EXHIBIT 6

Optimal Domestic Large-Cap (ODLC) Equity Index



Source: Dow Jones Indexes Research.

EXHIBIT 7

Performance Comparison by Index—10 Years Ending 4Q2004

	Product Composite Benchmark	S&P 500	ODLC
Total Number of Product Composites	**635**	**635**	**635**
Number of Product Composites Outperforming	429	465	328
% of Product Composites Outperforming	68%	73%	52%
Average Annual Alpha	1.03%	1.10%	0.31%
Average Annual Tracking Error	6.58%	7.67%	7.59%
Average Information Ratio	0.13	0.13	-0.02

Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network gross of fees.

itive alpha were either lucky or clever to have selected from the managers who generated positive alpha, and vice-versa.

This means that the quest for alpha within this asset class begins with the ability to identify alpha-generating managers. Without this ability, investors may be better off indexing to an optimal benchmark.

Exhibit 8 presents the excess quarterly returns of the ODLC over the S&P 500. On average, the ODLC outperformed the S&P 500 by 20 basis points a quarter. To understand the significance of 20 bp in excess return every quarter, consider this. A quarterly return of 0.20% would cause $100 to grow to $108 over 40 quarters. The cumulative difference of 8% is significant.

Finally, note that in quarters the ODLC outperformed the S&P 500, it did so with an average positive quarterly return of 47 basis points. In the months it underperformed the S&P 500, it did so with an average negative return of only 22 basis points.

CONCLUSION

We have used the performance of 635 large-cap domestic equity products over a period of ten years to compare the efficiency of three families of style indexes (see Enderle, Pope, and Siegel [2003] for another comparison of broad-capitalization indexes of the U.S. equity market).

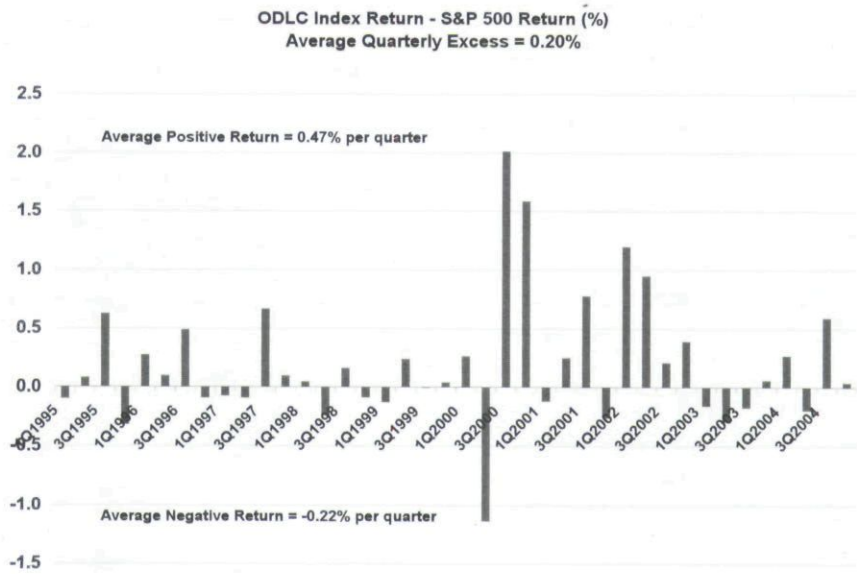
The result of the analysis prompts several conclusions:

First, one size or style or family does not fit all. Each of the three large-cap domestic styles analyzed has a return-risk profile that is similar, but unique. One single family is not efficient in all the styles.

For example, while the Russell 1000 value index seemed like a good benchmark to serve as a bogey for a large-cap value manager, its growth benchmark was not as efficient. This means that every style methodology has a weakness and may not lend itself well to being a benchmark for all the styles. Thus, choosing a single family of

EXHIBIT 8

Quarterly Excess Returns of ODLC versus S&P 500—10 Years Ending 4Q2004



Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network gross of fees.

indexes as a set of benchmarks may not always be the best solution for a plan sponsor.

Second, a combination of indexes across families may result in a better index. Our historical analysis has shown that over the last ten years there was clearly an opportunity to create a more efficient benchmark by combining the style indexes of the three major domestic large-cap equity index providers. This combined index would have been a good representation of the asset class and serve as a better bogey for performance evaluation.

A good benchmark would be information-free, resulting in a zero average alpha for a large population of managers over a long period. Managers would still outperform and underperform the benchmark due to chance or the use of information outside the benchmark that was proprietary to the manager. If the proprietary information were valuable, the manager would probably outperform, while if it were of no value, the manager would probably underperform.

Third, even if one were to build an efficient domestic large-cap equity benchmark, there would still be a role for active managers. The key to manager selection would then be to really understand an active manager's management philosophy and select the manager on the basis of his or her value-added proprietary information.

Fourth, and perhaps the most important, is the fact that historically manager performance is consistent with theory. Relative to an efficient benchmark, average manager alpha and information ratio are zero. Therefore, the search for alpha begins with the ability to identify winning managers. In the absence of such ability, investors may be better off indexing their assets to an efficient benchmark.

Finally, critics may argue that analysis such as ours is easy to perform on an ex post basis and that without the benefit of 20/20 hindsight, it would be very difficult if not impossible to build an index such as the ODLC. We agree.

Our goal, however, is to illustrate that over the past ten years the choice of a benchmark would have clearly affected not only the performance of a plan but also the methods used to implement the asset allocation decision. In addition, over the same period, the choice of an efficient benchmark would have resulted in a handsome excess return over an inefficient benchmark.

We want a better benchmark selection by sponsors now to save looking back ten years from now at the returns they left lying on the table.

APPENDIX

Performance of Large-Cap Equity Domestic Benchmarked Products—10 Years Ending 4Q2004

Product Benchmarks (number of products)	Annualized Alpha (%)			Annualized Tracking Error (%)			Annualized Information Ratio		
	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Blend:									
S&P 500 (359)	1.03	-1.41	3.86	6.64	2.10	11.76	0.15	-0.27	0.56
Russell 1000 (20)	0.94	-1.05	2.55	5.52	1.88	8.78	0.19	-0.13	0.57
Value:									
Barra/S&P 500 Value (21)	0.20	-0.97	1.18	5.06	3.42	7.46	0.04	-0.18	0.39
Russell 1000 Value (120)	0.27	-2.23	2.48	5.37	3.52	7.37	0.02	-0.43	0.41
Growth:									
Barra/S&P 500 Growth (6)	0.09	-2.09	1.88	9.50	5.04	17.04	-0.01	-0.29	0.31
Russell 1000 Growth (109)	2.10	-0.44	5.57	8.04	4.37	11.61	0.24	-0.07	0.59

Source: Dow Jones Indexes Research. Quarterly returns data from Plan Sponsor Network gross of fees. Min and Max ignore outliers and refer to the 10th and 90th percentiles of the respective universes.

ENDNOTES

[1]See Calkins [2004] for an explanation of why plan sponsors use index investing.

[2]Also see Modigliani [1999].

[3]Another way to compare the popularity of the S&P and Russell benchmarks would be to use estimates of the assets managed against these benchmarks.

[4]Survivorship bias does not pose a problem for our analysis, because the goal is to compare the performance of managers to different benchmarks, and one can assume that survivorship bias influences the results similarly with respect to each benchmark.

[5]These weights would be optimal if an investor has no long-term views on the relative performance of the three styles.

REFERENCES

Calkins, Nancy. "Plan Sponsors Index Investing." *Journal of Indexes*, August/September 2004.

Enderle, Francis J., Brad Pope, and Laurence B. Siegel. "Broad-Capitalization Indexes of the U.S. Equity Market." *The Journal of Investing*, Spring 2003.

Modigliani, Leah. "Selecting and Rejecting Benchmarks." U.S. Investment Strategy, Morgan Stanley Dean Witter, December 21, 1999.