The Active Management Premium in Small-Cap U.S. Equities

Real, or a figment of universe construction bias?

Gregory C. Allen

A portfolio that earned the average return for a broad universe of institutional small-cap U.S. equity managers in each quarter over the 20 years ended June 30, 2004, would have outperformed the Russell 2000 by over 500 basis points per year. While this is striking outperformance, its consistency is perhaps even more remarkable. Over the same period, the median small-cap manager outperformed the Russell 2000 in every rolling three-year period. By contrast, a portfolio that earned the average return for a broad universe of large-cap U.S. equity managers over the last 20 years would have underperformed the Russell 1000 by 30 basis points, and the median large-cap manager would have beaten the index in only 35% of the rolling three-year periods.[1]

The numbers posted by small-cap managers have attracted significant attention from institutional investors. Plan sponsors are under increasing pressure to extract return from their portfolios, whether through beta (market exposure) or alpha (active management exposure). With the increasing acceptance of risk budgeting as a framework for allocating active management risk, plan sponsors are taking a hard look at historical performance, seeking to allocate active risk capital to the areas that have delivered the best bang for the buck.

All this begs the question as to whether the historical outperformance of small-cap managers is a legitimate phenomenon. Efficient markets adherents argue that persistent outperformance in any reasonably efficient market is impossible. Active management disciples (many of

Gregory C. Allen is the director of research and an executive vice president at Callan Associates Inc., in San Francisco, CA.

them successful small-cap managers) argue that, even in the U.S. market, perhaps the most efficient in the world, small-cap stocks are underresearched and often mispriced. This, they say, conveys a significant advantage to well-resourced and disciplined investors.

The only thing these two camps seem to agree on is that the small-cap universes purveyed by consultants and various other vendors of data are fraught with biases and inconsistencies that make them suspect in any type of rigorous analysis.

That said, I use just such a universe to attempt to shed light on whether the small-cap premium is in fact real. I examine the techniques used in construction of the small-cap universe and its historical behavior relative to the Russell 2000. I discuss the impacts of survivorship bias and instant history bias, two well-documented biases that plague the constructors of universes. Finally, I evaluate the impact of any persistent factor bets that the small-cap managers may have taken over this period relative to the broad benchmark.

My conclusion is relatively straightforward. After taking into account universe construction biases and persistent factor biases, active institutional small-cap U.S. equity managers have significantly and consistently outperformed their passive benchmarks over the last 20 years. This achievement is all the more striking when we look at the record for institutional large-cap U.S. equity managers.

UNIVERSE CONSTRUCTION AND COMPOSITION

While much has been written on the relative performance of active managers to benchmarks, we have generally done a poor job describing how the underlying universes are constructed. This may seem a trivial detail, but in fact the way a universe is constructed and maintained can have an influence on the conclusions we reach when studying its behavior.

The small-cap universe used in this analysis is called the Total Institutional Small-Cap universe (TISC). It is a subset of a much larger manager database built and maintained by Callan Associates, Inc., an institutional investment consulting firm. This database was built and has been consistently maintained following a set of rules established over 20 years ago. The rules are designed to allow the database to support a performance measurement and reporting process for large and complex institutional portfolios. In aggregate, this database represents virtually every product that has been marketed to U.S. tax-exempt institutional investors over the last 20 years.

The first step in the construction of any universe is data collection. For the TISC universe, this process begins on the first business day of every quarter. Quarterly rates of return are collected directly from managers in several ways, including direct mail and e-mail-based questionnaires, a web interface, and as a final resort, targeted telephone calls. The objective is 100% coverage on a trailing two-quarter basis. While this goal is never actually achieved, the process results in a very robust level of coverage over time.

The returns in the TISC universe are reported gross of fees, and represent the composite return for institutional separate account products or commingled funds. Mutual funds are not included, as their returns are reported net of fees; they typically maintain significant cash positions; and they are generally managed with different objectives. No attempt is made to verify the validity of the returns other than application of some simple algorithms designed to catch data entry errors. Managers are given the opportunity to review and correct the entire return history for each of their products every quarter. In general, returns since about 1990 are AIMR-compliant.

The basic rule that guides the maintenance of the overall database is that managers can overwrite returns submitted in the past with corrected returns, but they cannot delete them. This rule, designed originally to create stability in the reporting of return distributions for periods ended in the past (calendar years, for example), is what saves much of the history in this database from the "great AIMR restatement" of 1994 and 1995.

One of the unintended consequences of requiring managers to become AIMR-compliant in the early to mid-1990s was that when they undertook their audits to become compliant, most went only as far back into history as was practical at that time (typically five to ten years). This meant much of the historical performance record from the 1980s was wiped out in many commercially available databases.

The data I use in my analysis spans the 20-year period ended June 30, 2004. The TISC universe represents a total of 783 products, all with at least a three-year performance record. The universe started the period with 126 products as of September 30, 1984. Only 57 of these original 126 products survived the entire period. A total of 657 products were created over this period. A total of 205 products stopped reporting returns (terminated).

Exhibit 1 shows the total size of the universe at the beginning of each calendar year and the number of terminations and inceptions in that year.

EXHIBIT 1

Universe Size, Inception, and Termination History—Total Small-Cap Database

Year	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01	02	03	04	Total
Count	126	158	186	223	252	279	309	330	356	392	435	477	535	585	621	648	654	663	641	601	783
Inceptions	32	29	37	29	27	32	24	33	41	46	50	66	57	48	49	39	18	0	0	0	657
Terminations	0	1	0	0	0	2	3	7	5	3	8	8	7	12	22	33	9	22	40	23	205

In compiling the TISC universe, every effort was made to identify all the small-cap composites in the database over the last 20 years. This includes not only traditional small-cap core products managed against the Russell 2000, but also products managed against style benchmarks. Products managed against the slightly larger-capitalization Russell 2500 are also included.

Exhibit 2 shows the breakdown of the universe by investment objective as defined by the stated benchmark for each product.

No database includes the entire history of all the small-cap products that were available to institutional investors over this period. All commercially available databases represent subsets or samples of the population. The TISC universe, with its breadth of coverage, stable construction rules, and long-term history, appears to be a robust sample. Much of the rest of my analysis is devoted to determining whether it is a reasonably unbiased sample compared to the actual population.

HISTORICAL RECORD

In an evaluation of the TISC universe, the historical record of active small-cap managers is striking both in the extent and the consistency of outperformance of benchmarks. Exhibit 3 compares the mean and median returns for the small-cap universe with the Russell 2000 index over various periods ended June 30, 2004.

The mean return over any period is calculated by taking the average return of all the products in the universe in each quarter. These average quarterly returns are then linked to calculate a cumulative mean return. Thus, the mean return for the 20-year period takes into account the performance history of all the 783 products in the TISC universe over the period.

The median return is relatively straightforward. It represents the middle of the return distribution for all managers with a return series that spans the entire period. Thus, for the 20-year period, the median represents the midpoint of the return distribution for the 57 products that had a full 20-year record.

EXHIBIT 2

Composite Count by Benchmark—TISC Universe

Benchmark Name	Count	Percent
Russell 2000	242	30.9%
Russell 2000 Growth	213	27.2%
Russell 2000 Value	172	22.0%
Russell 2500	55	7.0%
Russell 2500 Growth	37	4.7%
Russell 2500 Value	36	4.6%
S&P 600 Small-Cap	22	2.8%
S&P 600 Small-Cap Growth	3	0.4%
S&P 600 Small-Cap Value	3	0.4%
Total	**783**	**100.0%**

EXHIBIT 3

Returns for Periods Ended June 30, 2004—Total Institutional Small-Cap Universe

	5 Years	10 Years	15 Years	20 Years
Russell 2000 Index	6.63	10.93	10.44	11.11
TISC Mean	11.79	16.21	15.94	16.77
TISC 10th Percentile	19.03	18.76	17.87	18.32
TISC 25th Percentile	15.39	16.97	15.95	16.84
TISC Median	11.55	15.10	14.48	15.16
TISC 75th Percentile	6.60	12.95	12.75	13.73
TISC 90th Percentile	2.59	10.72	11.14	12.55
Number of Products	493	274	151	57

Exhibit 4 plots the rolling three-year excess return relative to the Russell 2000 over the last 20 years. The zero line represents the Russell 2000. Positive values indicate that the manager outperformed the Russell 2000 over the trailing three-year period. Negative values indicate underperformance. The gray area represents the range of excess returns for the TISC universe from the 90th through the 10th percentile for each period. The line in the center of the gray area represents the excess return for the mean of the universe.

As Exhibit 4 illustrates, the excess return for the mean of this universe has been positive in every rolling three-year period during this 20-year span. The worst relative periods were the three years ended in December of 1993 and 1994 when the mean excess return dropped to

EXHIBIT 4

Rolling Three-Year Excess Return Relative to Russell 2000 Index

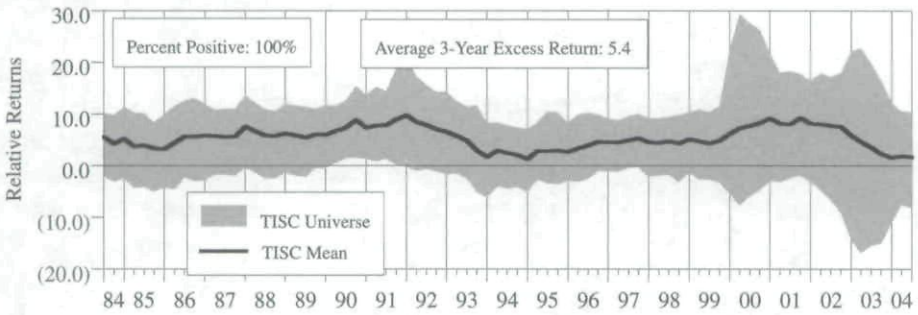


EXHIBIT 5

Rolling Three-Year Excess Return Relative to Russell 1000 Index

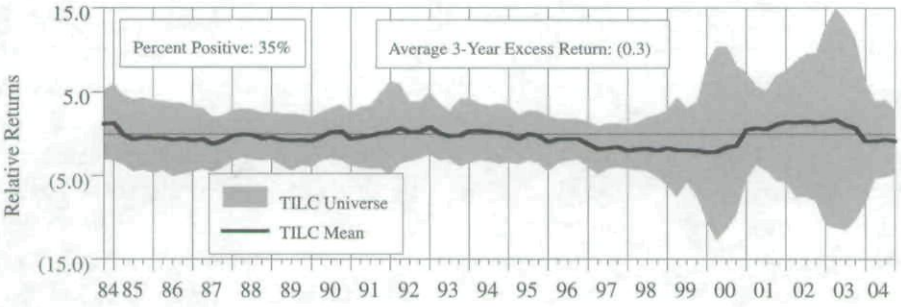
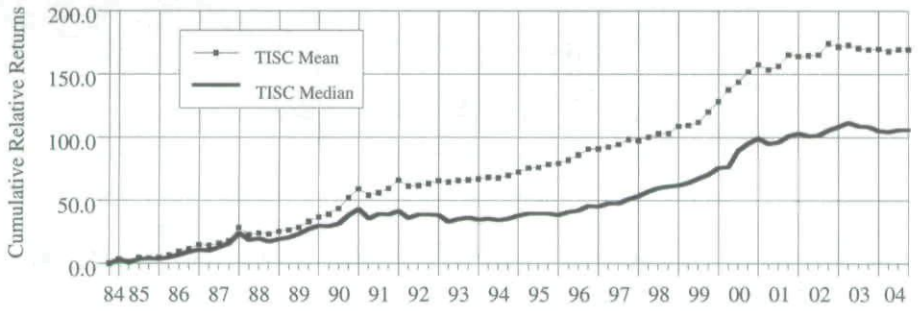


EXHIBIT 6

Cumulative Excess Return Relative to Russell 2000 Index



roughly 1.5% annualized. The average three-year excess return for the mean of the TISC universe over this period was 5.4% annualized.

The record for institutional large-cap managers will help put the results for the TISC universe into perspective. Exhibit 5 is an analogue to Exhibit 4 that compares the performance of the Total Institutional Large-Cap (TILC) universe to the Russell 1000 over the same 20-year period.[2]

As Exhibit 5 illustrates, the excess return for the mean of the TILC universe is positive during only 35% of the rolling three-year periods. Furthermore, the average three-year excess return for the mean of the TILC universe over this period was negative at –0.3% annualized.

These results indicate a significant difference between the relative performance results for institutional small-cap and large-cap managers, especially given that the two universes were constructed by the same organization, using the same process, and the same basic construction and maintenance rules. So what is the source of this difference?

To help answer this question, I examine some of the commonly cited criticisms of manager databases in connection with the TISC universe.

SURVIVORSHIP BIAS

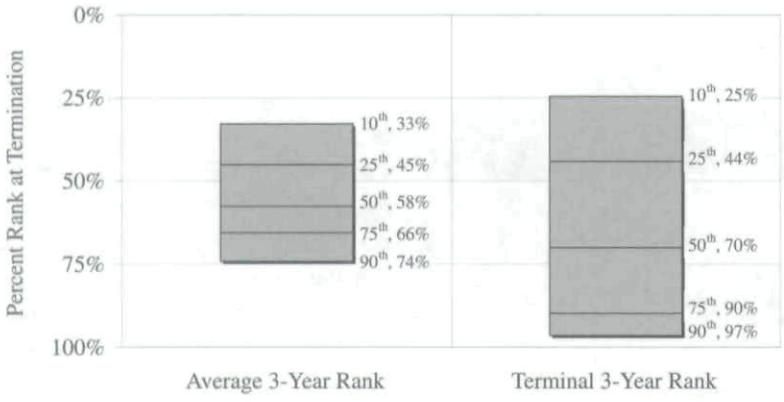
A bias in the construction of a manager universe is any flaw in the process that systematically causes the results for the sample (the universe) to differ from those of the population (the true opportunity set). Survivorship bias is the most commonly cited bias ascribed to universe construction.

Survivorship bias, it is argued, stems from the fact that poorly performing managers are much more likely to stop reporting returns to manager databases than top-performing managers. Thus, over time, and particularly for longer periods, better-performing managers tend to dominate the results for manager universes. This biases the performance results for the sample upward.

One way to examine the impact of survivorship bias is to compare the cumulative return for the mean of the universe to the cumulative return of the median manager over longer and longer periods. Since the median can include only the managers that survived over the entire measurement period, and the mean includes all managers in the universe each quarter, you would expect these to diverge in a biased sample. In particular, you would expect the median to significantly outperform the

EXHIBIT 7

Range of Ranking in TISC Universe of Non-Surviving Products



mean in a universe plagued by survivorship bias.

Exhibit 6 examines this relationship for the TISC universe. Clearly the mean and the median for the TISC universe diverge steadily and significantly over time. The fact that the mean outperforms the median, however, does not support the contention that survivorship bias is the dominant bias affecting this universe. In fact, the evidence supports a conclusion that the longer a small-cap manager survives, the more poorly it tends to perform compared to its newly introduced peers.

A second way to examine the impact of survivorship bias is to simply compare the performance of the managers that didn't survive with those that did. Over this 20-year history, 205 of the managers in the TISC universe were non-survivors. Exhibit 7 compares the relative performance for the subuniverse of non-survivors to the overall TISC universe using two common metrics of relative success.

The first bar shows the range of values for the average three-year ranking of the manager over its entire history until termination. The median value for the non-surviving group for this metric was the 58th percentile (versus the 50th percentile for the group as a whole). This indicates that, in general, the non-surviving funds were slight underperformers of their peers. While this is an indication of survivorship bias, it is interesting to note that over 40% of the non-survivors had generated above-median performance over the course of their histories leading up to termination.

The second bar shows the range of values for the trailing three-year ranking of the manager, as of the quarter of termination. The median value for the non-surviving group for this metric is the 70th percentile, and fully 25% of the non-survivors were in the bottom decile of the TISC universe when they stopped reporting returns.

This indicates that a high percentage of the non-survivors were in a particularly difficult period in their performance history when they elected to stop reporting. It is interesting to note again, however, that 30% of the non-survivors had above-median track records for the three-year period immediately preceding their termination.

Overall these results indicate that the non-survivors in the TISC universe generally underperformed the survivors. This leads to a conclusion that the universe does suffer from some level of survivorship bias. Its impact, however, is tempered significantly by the fact that successful small-cap managers (up to 40% of the non-survivors) often stop reporting returns to consultant databases when they reach capacity and close their products to new assets. This phenomenon is not observed in the more liquid asset classes including large-cap U.S. equity and U.S. fixed-income.

Ultimately, that the mean return for the TISC universe has consistently and significantly exceeded the median manager return in the universe is an indication that survivorship bias, while present, is not the dominant bias affecting this universe.

INSTANT HISTORY BIAS

Instant history bias occurs because managers typically start reporting to consultant databases only after they have generated a solid three-year performance record. When they do start reporting, they submit their entire performance history on the first day. Thus, while institutional investors are only beginning to recognize them as part of the real opportunity set, they have already contributed three years of history to the sample universe.

In small-cap universes, this bias is compounded by the fact that the first three years of any manager's record is typically built on a very small asset base. Given the illiquid nature of the asset class, this conveys a significant advantage to new managers over the rest of the investable population.

One way to measure the impact of instant history bias is to simply truncate the return history for every manager in the universe (eliminate their first three years), and then compare the performance of the truncated universe with that of the universe as is. Exhibit 8 illustrates the cumula-

tive performance of the truncated TISC universe and the TISC universe as is over the 20-year analysis period.

As the plot illustrates, the truncated universe significantly and consistently underperformed the intact universe. Over the entire 20-year period, the truncated universe underperformed by 94 basis points per year with almost no subperiods of outperformance.

This result supports the intuitive conclusion that instant history bias significantly impacts the performance results of small-cap universes relative to the actual investable universe of small-cap managers. It also suggests a reasonable approach to adjust for this bias.

Applying a screen that requires either three years of performance or a reasonable level of assets under management ($50 million, for example), before a manager can be added to the universe, would effectively reduce the impact of the bias. This adjustment would result in a sample that more accurately represents the true opportunity set for institutional investors.

PERSISTENT FACTOR BIASES AND BENCHMARK MISSPECIFICATION

Survivorship bias and instant history bias result in samples that systematically misrepresent their underlying populations. This undermines the credibility of any analysis that uses these samples to make claims about the behavior of the population as a whole. Factor biases are somewhat different in terms of their impact.

A factor bias in a manager universe is defined as a persistent over- or underweighting of some recognizable factor that explains performance over time. The under- or overweighting is calculated relative to the passive benchmark that is used to evaluate the relative performance of the universe. Thus, a factor bias does not imply that the sample universe misrepresents the population. Rather, it implies that the sample universe (and by extension the population) may have some kind of systematic advantage (or disadvantage) relative to the benchmark because of a persistent mismatch between the two.

There are two ways to correct for this type of bias when we compare the performance of a universe to that of a benchmark. The universe can be adjusted to look more like the benchmark, or the benchmark can be adjusted to look more like the universe.

Adjusting the universe to match the benchmark has a number of methodological problems, not the least of which is the potential of a systematic bias in the sample relative to the underlying population. Adjusting the benchmark to reflect the composition of the universe is more straightforward, and is generally useful in helping us better understand the nature of the bias.

Exhibit 2 showed the breakdown of the TISC universe by investment objective as defined by the benchmark for each product. 15% of this universe is benchmarked to some derivative of the Russell 2500, a somewhat larger-capitalization benchmark than the Russell 2000. Managers in this universe benchmarked to growth indexes outnumber those benchmarked to value indexes, 253 to 211. This information would suggest that the universe has a slight mid-cap growth bias relative to the Russell 2000.

Return-based regression analysis confirms this conclusion. Using an algorithm that searches for the best historical fit between the mean returns for the universe, and some combination of the Russell indexes, reveals that a custom benchmark of 85% Russell 2000 and 15% Russell 2500 growth has been a better benchmark for the TISC universe than the Russell 2000. An evaluation of this custom benchmark in the context of the portfolio characteristics of the TISC universe further supports this conclusion.[3]

Exhibit 9 illustrates the market capitalization rank-

EXHIBIT 8

Cumulative Excess Return Relative to TISC Universe

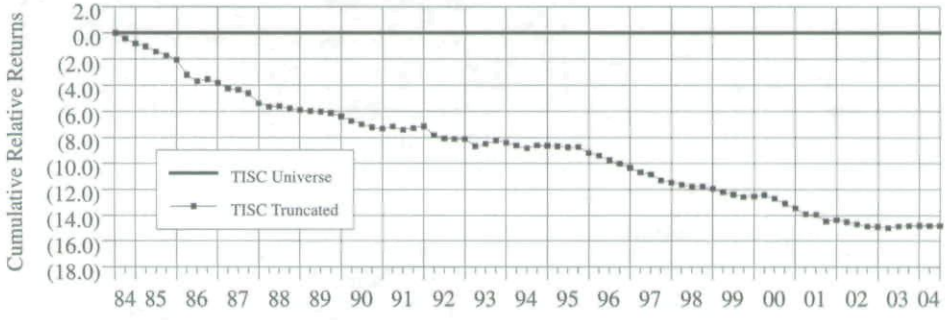


EXHIBIT 9

Weighted Median Market Cap Rankings—TISC Universe

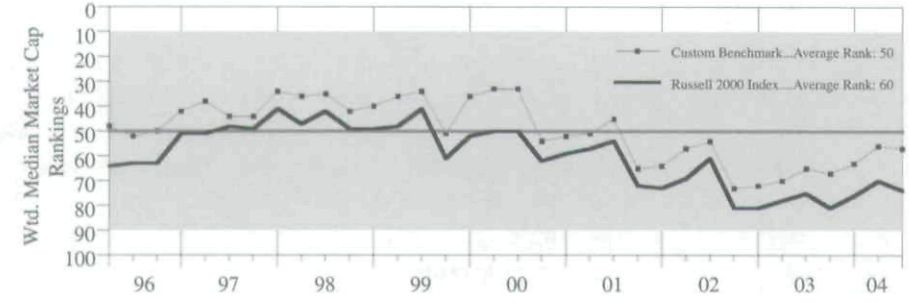
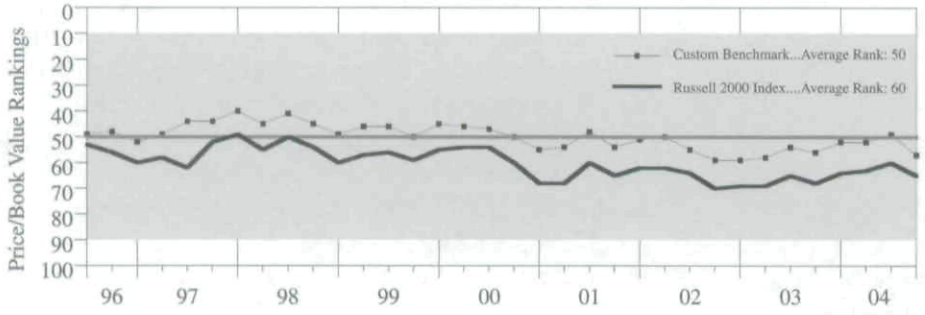


EXHIBIT 10

Price/Book Value Rankings—TISC Universe



ing of the Russell 2000 and the custom benchmark relative to the TISC universe since the first quarter of 1996. While neither of the two benchmarks ranks at the median for the universe in every quarter, the custom benchmark is clearly the better fit over time, with an average ranking of the 50th percentile versus an average ranking of the 60th percentile (lower than median) for the Russell 2000.[4]

Exhibit 10 is the analogue to Exhibit 9 on price-to-book ratio, a measure of growth versus value. Again, the custom benchmark is the better fit, with an average rank of the 50th percentile versus the 60th percentile (lower price/book ratio, thus fewer growth characteristics) for the Russell 2000.

These results support a conclusion of a persistent and identifiable mismatch between the TISC universe and the Russell 2000 over this period. The question remains as to whether this bias conferred any kind of systematic performance advantage.

Exhibit 11 compares the performance of the mean of the TISC universe and the truncated TISC universe to various small-cap benchmarks, including the custom benchmark, over a number of periods ended June 30, 2004. The performance of the custom benchmark is remarkably similar to that of the Russell 2000 over all the subperiods shown. This fact does not support a conclusion that the observed mid-cap growth bias in the TISC universe conferred a meaningful performance advantage over any of these subperiods.

A further evaluation of the consistency of this outperformance in Exhibit 12 indicates that the mean of the TISC universe outperformed the custom benchmark in all the rolling three-year periods during the last 20 years. These results are consistent, both in terms of magnitude and consistency, with the results relative to the Russell 2000.

Thus, a careful examination of the composition, the performance history, and the portfolio characteristics of the TISC universe clearly supports a persistent mid-cap growth bias relative to the Russell 2000. The mid-cap component of this bias did confer a performance advantage over this period (as evidenced by the outperformance of the Russell 2500 relative to the Russell 2000). The growth component of the bias, however, more than mitigated this advantage.

Ultimately, the mean for the TISC universe significantly outperformed all the Russell 2000 and Russell 2500 derivative indexes over the last 20 years. The extent of this outperformance does not support the contention that it is the result of a consistent factor bias.

CONCLUSION

The performance history of the TISC universe strongly supports the conclusion that institutional small-cap managers as a group have consistently and significantly outperformed their benchmarks over the last 20 years. A careful examination of the evidence indicates that approximately 20% of this outperformance occurs because the first three years of a typical small-cap manager's performance record are usually their best years. For a number of structural reasons, these first three years are effectively not accessible to the typical institutional investor.

There seems to be little evidence to support the assertion that survivorship bias or persistent factor biases

EXHIBIT 11

Returns for Periods Ended June 30, 2004

	5 Years	10 Years	15 Years	20 Years
TISC Universe	11.87	16.30	16.01	16.83
TISC Truncated	11.24	15.51	15.25	15.89
Custom Benchmark*	6.07	10.83	10.38	11.07
Russell 2000 Index	6.63	10.93	10.44	11.11
Russell 2000 Growth	1.98	9.69	9.60	10.49
Russell 2000 Value	12.82	13.91	12.92	13.64
Russell 2500 Index	8.47	13.08	12.20	13.25
Russell 2500 Growth	1.98	9.69	9.60	10.49
Russell 2500 Value	11.93	14.67	13.42	14.32

**Custom benchmark equals 85% Russell 2000, 15% Russell 2500 Growth, rebalanced quarterly.*

can explain away the balance of the outperformance of this universe. Survivorship bias is substantially mitigated by the fact that many of the small-cap managers that stopped reporting their data were actually top performers who simply closed their funds. While this universe did exhibit a mild mid-cap growth bias compared to the Russell 2000, this persistent factor bet did not appear to confer any performance advantage over the last 20 years.

These findings lend strong empirical support to the conclusion that the small-cap U.S. equity market is fertile ground for institutional investors seeking to allocate their active management risk. This conclusion is further strengthened when we evaluate these results in the context of the record for the large-cap U.S. equity managers that currently manage the vast majority of institutional active risk capital.

ENDNOTES

[1]Our comparisons use the Callan Associates' broad database of institutional U.S. equity (large-cap and small-cap) separate account and commingled composites.

[2] The Total Institutional Large-Cap universe is composed of separate account and commingled fund composites managed against large-cap benchmarks. It is similar to the TISC universe in all aspects of its construction and maintenance.

[3]The Russell 2500 growth and value indexes began in the first quarter of 1986. The 20-year period used throughout this analysis begins in the third quarter of 1984. To allow the analysis to consistently span this 20-year period, returns for the Russell 2000 growth and Russell 2000 value indexes were used to back-fill the missing six quarters of data for the two Russell 2500 derivative indexes.

[4]Data limitations on portfolio characteristics for managers in the TISC universe preclude any reasonable analysis prior to January 1996.

To order reprints of this article, please contact Ajani Malik at amalik@iijournals.com or 212-224-3205.

EXHIBIT 12

Rolling Three-Year Excess Return Relative to Custom Benchmark