



by PANAGIOTA BAILEY

Mind the GAP: Questioning the Investment Manager's Stated Benchmark

Selecting and monitoring fund investments is an arduous task, with no guarantee of success. For the majority of funds, the investment and due diligence process involves an evaluation of past performance and risk exposure relative to some benchmark. Thus, the choice of benchmark is critical in this context, and simply adopting the investment manager's stated benchmark for fund analysis may not yield accurate results. To help with this task, I developed the concept of "GAP" analysis—an easy-to-calculate measure for identifying the presence and extent of sub-optimality of the stated benchmark.

Since the 1986 publication of “Determinants of Portfolio Performance” by Gary P. Brinson, L. Randolph Hood, and Gilbert L. Beebower, there has been a growing literature about time-series and cross-sectional analysis. On the one hand, researchers have attempted to attribute the variation of the total return of the fund to the variation of the overall market return, the return from asset allocation policy, and the active return. On the other hand, researchers have attempted to attribute the variation across fund returns to these same components.¹ But before decomposing the sources of return, it is vital to select the appropriate benchmark. In the following, I develop the “GAP” concept, which builds on this literature and defines an easy-to-calculate measure aimed at helping practitioners identify the presence and extent of sub-optimality of the stated benchmark.

IMPORTANCE OF BENCHMARK SELECTION

Selecting and monitoring fund investments requires conducting in-depth due diligence in fund operations, compliance, investment management, and risk management.² Investment managers provide a host of information concerning different facets of the funds they manage, including the choice of benchmark.³ Investors, however, may lack the expertise necessary to develop a critical eye that cuts through the many sales pitches that embellish almost every fact sheet and presentation of

a fund, or they may simply lack the time and personal resources to dedicate to the task of fund due diligence. In either case, investors may be compelled to take information presented to them by the investment manager at face value. Critically determining the optimal benchmark for each fund under review is vital because it has direct implications for analyzing the relative historical performance of fund investments in terms of their return and risk characteristics.

I repeatedly come across instances in which the benchmark presented as appropriate by the investment manager (henceforth the “stated benchmark”) is in fact anything but. For example, I was recently analyzing the strategy of an investment manager whose aim was to generate positive excess returns by identifying U.K. stocks with growth potential and taking positions in these stocks when valuations were compelling. Over the life of the fund, the manager would invest around one-third of the value of the fund in large-capitalization stocks, another third in mid-cap stocks, and the final third in small-cap stocks. Hence, the appropriate investment universe for this fund was the universe of U.K. equities (except micro-caps), and the appropriate benchmark was a composite benchmark giving equal weights to indices of large-cap U.K. stocks (e.g., the FTSE 100), mid-cap U.K. stocks (e.g., the FTSE 250), and small-cap U.K. stocks (e.g., the FTSE SmallCap). In all oral and written presentations,

fund returns were being compared with the FTSE 100, which covers only the large-cap segment of the U.K. equity market, thereby providing a misleading indication of the alpha-generating capability of the fund manager.⁴ Furthermore, the exercise of comparing the fund returns with the appropriate composite benchmark was not even carried out internally (so the fund manager herself did not know her true alpha-generating capability!).

Various methods can be used to determine an appropriate benchmark, such as returns-based analysis, whereby the historical record of fund returns is analyzed against those of a set of indices; holdings-based analysis, in which the current characteristics of the securities in the portfolios are scrutinized in isolation; and factor-based models, which determine the factor exposures of the fund returns. Moreover, the forward-looking views of the managers must be considered. The end result should be to select a benchmark that (1) appropriately captures the exposure of the fund to systematic sources of risk and (2) uses a universe of securities as close as possible to the universe that the manager uses as the “opportunity set” for selecting investments.

The ideal benchmark is often not readily available, and so the best option may be to construct a fully customized benchmark.⁵ But in many cases, even with a custom benchmark, it may be necessary to compromise (because of time and cost considerations) and construct a benchmark that captures just the main systematic sources of risk of the fund manager and that reasonably approximates the universe of securities. In any case, selecting or constructing the appropriate benchmark is a non-trivial task, with significant implications for risk analysis and performance measurement, attribution, and evaluation.

METHOD FOR DETERMINING BENCHMARK SUB-OPTIMALITY

Methodology and data To determine whether the stated benchmark is *not* optimal, it is sufficient to show that another benchmark is better. I do so by using simple regression analysis to uncover whether the stated benchmark or some other index has a better fit with the fund’s historical return record. To judge the appropriateness of each index as a benchmark for a given fund, I use the R^2 measure⁶, which captures the variation of the dependent variable that is attributable to the variation of the independent variable. I use weekly returns over periods of one, three, and five years. Specifically, I included open-end funds that had at least a one-year historical record as of the end of June 2013, that presented themselves

as following an active strategy within emerging market equities, and that had the MSCI Emerging Markets Index (MSCI EM Index) as their stated benchmark.⁷ Using the Bloomberg database, I found 187 such funds corresponding to different strategies (i.e., not including different share classes for the same fund strategy). It is interesting to note that of these 187 funds, just 112, or 60%, were presented as having a multi-cap approach; a further 20% as not having a specific focus in terms of market capitalization; a further 19% as having a large-cap approach; and the remaining 1% as having a small-cap approach. Because the MSCI EM Index covers large-, mid-, and small-cap securities, there are grounds to be suspicious about the 40% of funds that do not classify themselves as having a multi-cap approach yet present a multi-cap benchmark as the appropriate benchmark for investors to measure their returns against.

All the funds in the sample use the MSCI EM Index⁸ as their stated benchmark. To then check the appropriateness of the MSCI EM Index, I selected four sets of MSCI indices⁹ that capture possible misalignments of fund returns with the stated benchmark. The first set consists of selected regional indices (world, developed, and frontier markets) that cover companies across market capitalizations. The second, third, and fourth sets consist of indices that divide the universe of emerging market stocks according to market capitalization (small-, mid-, and large-cap), style (value and growth), and sector (consumer discretionary, consumer staples, energy, financials, health care, industrials, information technology, materials, telecommunications services, and utilities).

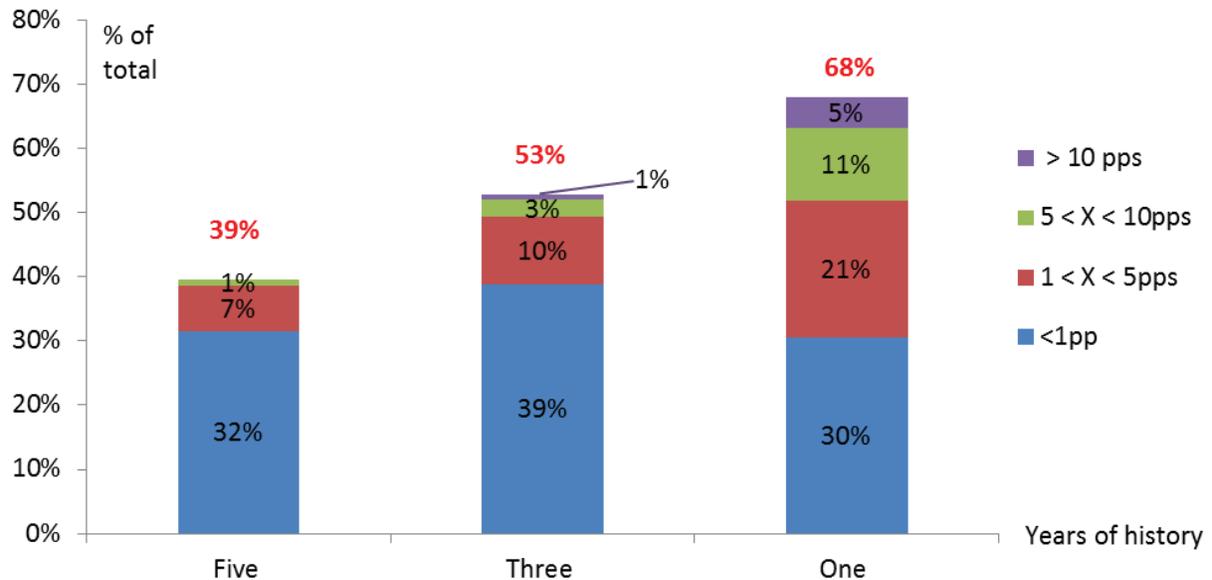
Results: Many managers’ benchmark choices are misleading The empirical results suggest strongly that relying on the investment manager for the choice of benchmark may at times be a very bad move.

Table 1 presents some key statistics for R^2 measured against the MSCI EM Index (the stated benchmark for the samples under study).¹⁰ For many managers, the investment manager’s choice of benchmark is optimal, in that the return history of that benchmark is able to account for the preponderance of the (in-sample) variation in the return history of the fund. But in some instances, the manager’s choice of benchmark fails miserably on this score.¹¹ From Table 1, we can already observe that the generated time series of R^2 is in all instances negatively skewed, showing that most fund managers have an R^2 that lies to the right of the mean (also shown in Table 1). This finding provides little comfort, as Figure 1 shows.

Table 1. Key Statistics for R^2 Measured against the Stated Benchmark

Years of History	Average	Median	Min	Max	Standard Deviation	Skew	Kurtosis
5	87%	90%	58%	99%	9%	-1.1	0.9
3	84	89	3	99	14	-2.4	9.2
1	79	85	0	97	16	-1.9	5.5

Figure 1. Proportions of funds falling in GAPs over one-, three-, and five-year periods



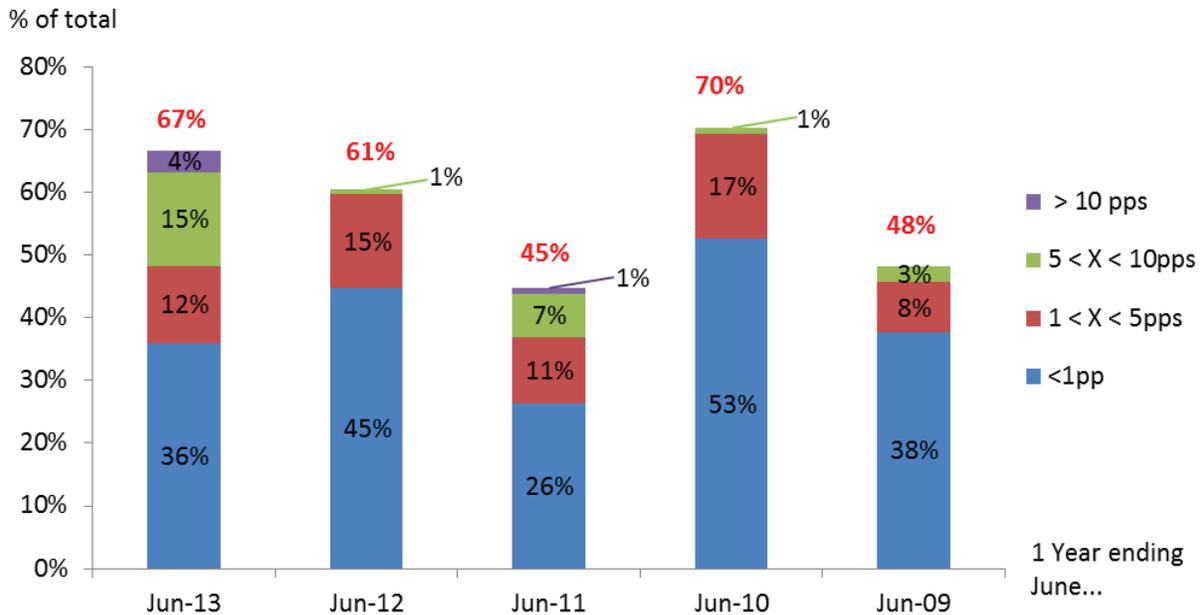
This figure groups the funds according to the difference (which I will henceforth call the “GAP”) between their R^2 metric when it is measured against the stated benchmark and the highest R^2 for the wider set of benchmarks and shows the proportion of funds falling within each GAP. A number of interesting observations arise from this analysis. Across all periods, and notably even in the longer term, the stated benchmark for many funds is not optimal from the point of view of accounting for the variation of fund returns, as shown by the high proportion of funds with a non-zero GAP.¹² Moreover, the GAP widens notably for shorter return histories.¹³ Even though active managers purposely deviate from their stated benchmarks on a tactical basis, the proportion of funds falling into wide GAPs is surprisingly high, indicating that the stated benchmark may not be optimal for analyzing relative fund returns.

To determine whether the stated benchmark is the best choice for the job of relative fund return analysis, further investigation is needed. **Table 2** shows some key statistics for R^2 measured against the MSCI EM Index—the stated benchmark for all the funds. **Figure 2**, similarly to Figure 1, again groups the funds according to the proportion falling into GAPs of varying size. Here the focus is restricted to the 114 funds with a five-year history of returns (as of the end of June 2013), and the results are shown for each of five 12-month periods. If the GAP persists across these periods, then it is not the result of active management. Rather, it is evidence that the benchmark presented by the investment manager is suboptimal and thus confounding if used for evaluating the relative performance of funds. This hypothesis is confirmed by Figure 2, which indicates that in addition to large portions of funds falling into a GAP of any size, significant portions fall in wide GAPs within these five one-year periods.

Table 2. Key Statistics for R^2 Measured against the Stated Benchmark

One Year Ending	Average	Median	Min	Max	Standard Deviation	Skew	Kurtosis
June 2013	80%	86%	37%	97%	14%	-1.1	0.7
June 2012	90	92	43	100	9	-2.5	8.5
June 2011	82	90	34	99	17	-1.1	-0.1
June 2010	88	91	57	99	9	-1.4	1.9
June 2009	89	92	53	100	10	-1.7	2.8

Figure 2. Proportions of funds falling in GAPS over five one-year periods

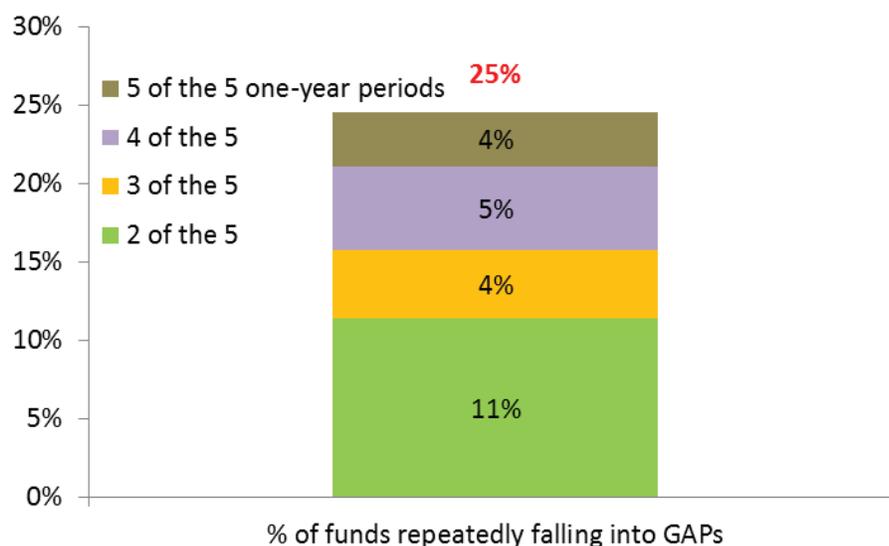


It is clear from Figure 2 that of the funds that fall into a GAP during one of these five one-year periods, many of them fall into a GAP again in one or more of the five-year periods. I then checked to see whether there is substantial rotation in the funds that fall into a wide GAP (> 1 pp) or if the funds that fall in a wide GAP once do so repeatedly. This information is summarized in **Figure 3**, which shows that 25% of the 114 funds with a five-year return history repeatedly fell in wide GAPS. Just less than half of these fell in wide GAPS in two out of the five one-year periods, whereas around 17% fell in wide GAPS in each of three, four, and five out of the five one-year periods ending June 2013.

CONCLUSION

It is important to question the investment manager's stated benchmark because there are good reasons for being skeptical. Often the explanatory power of a stated benchmark falls short of alternative benchmarks. Moreover, this discrepancy, or GAP, can persist across time periods, with some funds being repeatedly off course. Given the importance of benchmark selection for relative analysis of fund returns, investors at large, professional consultants advising on fund investments, and fund-of-funds managers should mind the GAP lest they fall in it themselves.

Figure 3. Proportion of funds repeatedly falling in wide GAPS



NOTES

- For a recent contribution to the literature on cross-sectional fund returns, see Xiong, Ibbotson, Idzorek, and Peng (2010).
- A concise reference on the many facets of fund analysis is the *Financial Times Guide to Investing in Funds: How to Select Investments, Assess Managers and Protect Your Wealth* by Jerome De Lavernere Lussan.
- For a discussion of the “tricks” that investment managers engage in to present themselves in a favorable light, see “Murder on the Orient Express: The Mystery of Underperformance” by Charles D. Ellis.
- In this example, style returns (which may be earned by the investor using exchange-traded or index funds rather than engaging an active manager and paying additional fees) and alpha returns are not properly distinguished. For instance, suppose you are investigating a manager of small-cap U.S. stocks who returned 10% for 20XX and that small-cap U.S. stocks returned 15% and the broad market returned 5% for the same period. The style is $S = 15\% - 5\% = +10\%$, and the active return is $A = 10\% - 15\% = -5\%$. If the manager erroneously compares her returns with the broad market, she may present herself as having achieved an active return of $A^* = 10\% - 5\% = +5\%$.
- For a review of the many intricacies of constructing and managing a custom benchmark, see “Creating and Managing Custom Benchmarks—A Practitioner’s Guide” by Stephen Campisi.
- The R^2 measure I used is defined as the square of the Pearson product moment correlation coefficient.
- Bloomberg makes available a field called “fund_benchmark_prim,” which provides the fund’s primary benchmark according to fund documents, such as the prospectus. If a benchmark is not available from such sources, then Bloomberg contacts the investment manager, who suggests to Bloomberg an appropriate benchmark and Bloomberg presents this in the fund_benchmark_prim field. The total returns used have a weekly periodicity.
- From the MSCI website (<http://www.msci.com/products/indices/tools/index.html#EM>): “The MSCI Emerging Markets Index is a free float-adjusted market capitalization index that is designed to measure equity market performance of emerging markets. The MSCI Emerging Markets Index consists of the following 21 emerging market indices: Brazil, Chile, China, Colombia, Czech Republic, Egypt, Hungary, India, Indonesia, [South] Korea, Malaysia, Mexico, Morocco, Peru, Philippines, Poland, Russia, South Africa, Taiwan, Thailand, and Turkey.”
- In all cases, MSCI indices are used that are expressed in U.S. dollars and that are gross indices—in that they provide the equity returns including dividends (the total return).
- The samples under study are open-end funds that present themselves as following an active strategy within emerging market equities with the MSCI Emerging Markets Index as their stated benchmark and have a (1) one-year history of weekly returns (187 funds), (2) three-year history of weekly returns (144 funds), and (3) five-year history of weekly returns (114 funds).
- If I remove the outliers from the one- and three-year return samples (consisting of two funds with 0% and one with a 3% R^2 , respectively), the funds with the next smallest R^2 still have a very low R^2 of 35% and 30%, respectively.
- The proportion of funds for which the stated benchmark did not yield the highest R^2 is very high in all the samples: 39%, 53%, and 68% for the five-, three- and one-year histories, respectively.
- The proportion of funds whose R^2 measured against the stated benchmark is less than the highest R^2 from across the wider set of benchmarks, by 5 to 10 pps or by more than 10 pps, jumps to 11% and 5%, respectively, from 0%, when considering the one-year histories rather than the five-year history.

REFERENCES

- Brinson, Gary P., L. Randolph Hood, and Gilbert L. Beebower. 1986. "Determinants of Portfolio Performance." *Financial Analysts Journal*, vol. 42, no. 4 (July/August):39–44.
- Campisi, Stephen. 2002. "Creating and Managing Custom Benchmarks—A Practitioner's Guide." *Journal of Performance Measurement*, vol. 6, no. 4 (Summer).
- De Lavenere Lussan, Jerome. 2012. *Financial Times Guide to Investing in Funds: How to Select Investments, Assess Managers and Protect Your Wealth*. London: FT Press.
- Ellis, Charles. 2012. "Murder on the Orient Express: The Mystery of Underperformance." *Financial Analysts Journal*, vol. 68, no. 4 (July/August):13–19.
- Xiong, James X., Roger G. Ibbotson, Thomas M. Idzorek, and Chen Peng. 2010. "The Equal Importance of Asset Allocation and Active Management." *Financial Analysts Journal*, vol. 66, no. 3 (March/April):22–30.
-
- Panagiota Bailey is an investment analyst at Peritus Investment Consultancy.*