Does the composition of the market portfolio *really* matter?

It does matter, but the "all-inclusive" market portfolio remains an elusive concept.

Keith C. Brown and Gregory D. Brown

Few topics in recent memory have generated more debate than the usefulness of the Capital Asset Pricing Model in monitoring investment performance. The heart of the controversy appears to be Roll's [1977] contention that a "correct and unambiguous" test of the CAPM is impossible because we can never observe the true market portfolio. One consequence of this simple truth is that an asset's measured risk factor (beta) will depend as much on the proxy we select to represent the universe of assets as well as on security-unique attributes.

In what must be a truly harrowing finding for portfolio managers, Roll used this fact to demonstrate that it is theoretically possible to generate exactly opposite rankings of "winners" and "losers" by using two different approximations of the market portfolio. His argument that "judicious choice of the index can produce any desired measured 'performance' " (1978, p. 1056) is the foundation for questions on the value of CAPM-based evaluation techniques.

The contrary viewpoint in this dispute perhaps is best expressed by Rosenberg [1981], who regards the criticism that the market portfolio is observed imperfectly as an unnecessary distraction. In fact, Rosenberg contends that "likely specification errors [in the approximated index] are relatively unimportant" and that the " 'true efficient portfolio' is not a useful construct" (p. 5). The source of such a divergent opinion is the belief that it is the application of the asset pricing paradigm, rather than theory itself, that is of importance. From this perspective, the central question is not whether some inefficient market proxy *could* lead to invalid inference, but, rather, whether a reasonable estimate is *likely* to generate distorted results.

This is clearly the philosophy adopted by Stambaugh [1982]. After creating four separate market indexes consisting of stocks, bonds, real estate, and consumer durables, he found that the separate proxies led to identical inferences about the CAPM. He concluded that a reversal of inference may still be possible, but the reversal is not so inevitable as previously believed.

To this point, little is known about the actual empirical sensitivity of CAPM-based performance measures to different specifications of the market portfolio. Certainly, Stambaugh's results imply that rankings of investor performance should be robust with respect to a wide range of estimates. On the other hand, French and Henderson [1985] have shown that evaluation measures may still have trouble assessing superior or inferior investment capabilities even under ideal conditions involving no identity problems. Further, Roll [1981] and Blume [1984] have already suggested alternative techniques for correcting for the observability problem.

The purpose of this study is to evaluate the historical performance of several publicly traded investment portfolios relative to six different market proxies. Specifically, we use the evaluation method first described by Jensen [1968] to analyze the annual

KEITH C. BROWN is Assistant Professor at the Graduate School of Business at the University of Texas at Austin (TX 78712). GREGORY D. BROWN is a financial analyst with Arthur Andersen & Company in Dallas (TX 75243). The authors wish to acknowledge Richard Roll's contribution to the manuscript, especially his insight into the problems of measuring portfolio management performance.

returns of thirty-two mutual funds in the period 1947–1978. We will show that the array of indexes do produce substantially similar rankings of the funds, but they lead also to dramatically different inferences about abnormal investment performance. The inclusion of real estate in the market proxy in particular can lead to highly erratic results. Our conclusion, therefore, is that the composition of the market proxy does indeed matter for the particular set of assets under scrutiny.

CONSTRUCTING THE MARKET INDEXES

To formulate each separate market proxy, we obtained the historical rates of return and the relative market values of five different classes of capital market assets. The source of these data was Ibbotson and Fall [1979], whose study covered the thirty-two year period from 1947 through 1978.

The major classes of securities are:

1. Common stocks, including NYSE, AMEX, and OTC equities.
2. Fixed-income corporate issues, including preferred stocks, intermediate- and long-term bonds, and commercial paper.
3. Real estate, including the USDA aggregate market value of farm investments and residential housing aggregate values (excluding urban land values) assembled from estimates of net rental yields and an index of capital appreciation.
4. United States government issues, including Treasury bills, notes, and bonds, as well as government agency securities.
5. Municipal bonds, including both state and local, and short- and long-term bonds.[1]

On the surface, an index consisting of all five asset categories might be considered to be the optimal market portfolio. Ibbotson and Fall were quick to point out that this is not the case, however. Some assets, such as human capital and consumer durables, are left out while others might be overstated.[2] Despite these drawbacks, the data included in this study do measure the most identifiable and liquid of the capital market securities. More important, these assets also comprise the set of opportunities for the vast majority of investors, so they are reasonable representation of the investment market.

To create the six market proxies, we began, as Stambaugh did, by including only the returns to common stock. We then added the other securities incrementally to form an increasingly broader portfolio. To extend the sensitivity analysis, we formed a final index with all the assets except common stock. In each case, the annual market values were used to provide the percentage weighting that a security class would receive in the portfolio that year. We then multiplied these weights by the annual returns to each asset and added them to create a value-weighted market index. The six market proxies are summarized as follows:

Index 1. Common stock,

Index 2. Index 1 plus fixed-income corporate issues,

Index 3. Index 2 plus real estate,

Index 4. Index 3 plus United States government issues,

Index 5. Index 4 plus municipal bonds, and

Index 6. Index 5 less common stock.

Table 1 gives a statistical summary of the thirty-two annual returns for each of the indexes. For purposes of comparison, the characteristics of the Treasury bill returns are included for the same time period. T-bills of course are a component of the United States government issue category, and we will use them later as the risk-free rate in the tests using Jensen's Performance Index.

Note that the index made up solely of common stock had both the highest historical risk (standard deviation) and return (mean) of any other portfolio. As the market indexes broaden, the dominating effect of common stock diminishes, and risk and return both drop off considerably. The decline in the standard deviation is most apparent when real estate is introduced in Index 3. The ultimate impact of this decrease in measured risk will be documented in subsequent sections.

CONSTRUCTING THE MUTUAL FUND RETURN SERIES

In his original research on mutual fund performance, Jensen identified 115 investment companies with a complete set of annual returns over the 1955–1964 period. This collection of portfolios became the base for the data sample we used, but several

TABLE 1

Statistical Summary of Market Indexes, 1947–1978

	Arithmetic Mean (%)	Standard Deviation (%)	Coefficient of Variation
Index 1	11.79	17.73	1.50
Index 2	9.07	13.62	1.50
Index 3	8.26	5.63	0.68
Index 4	7.25	4.71	0.65
Index 5	6.97	4.58	0.66
Index 6	6.08	2.88	0.47
Risk-Free Rate	3.53	2.07	0.59

1. Footnotes appear at the end of the article.

refinements in the original list were necessary for our purposes.

The market indexes were formed for the years from 1947 through 1978, so we consulted Weisenberger's *Investment Companies* for the funds that operated continuously in this time period. A total of forty-nine companies satisfied this constraint. In order to insure consistency, we imposed an additional set of restrictions that eliminated any fund that had altered its management strategy or changed its name.[3] We also removed several other companies that had acquired or had been acquired by other mutual funds during the data period. As a final refinement, we selected funds so as to achieve an equal representation from each of Weisenberger's four investment objective classifications: growth, income, income and growth, and balanced investment. As only eight income funds satisfied the requirements above, this became the effective constraint on the ultimate size of the sample. The final list of the thirty-two mutual funds with coded investment classifications appears in Table 2.

Having established the data base, we next determined the yearly returns for each of the sample companies. The mutual fund yields, unlike returns to the component assets of the market indexes, were not readily available. The Weisenberger manuals provided the initial observations for net asset values (NAV), income from dividends (ID), and capital gains distributions (CG). The fund j, period t, return could then be written as:

$$R_{jt} = \frac{(NAV_{jt} - NAV_{jt-1}) + ID_{jt} + CG_{jt}}{NAV_{jt-1}} \qquad (1)$$

Of course, all the variables were adjusted for stock splits and stock dividends. Consequently, Equation (1), computed on an annual basis for each individual investment company, was the source for all remaining data necessary for the performance evaluation tests. The means and standard deviations for these thirty-two return series are listed in Table 2.

TESTING FOR ABNORMAL INVESTMENT PERFORMANCE

The primary result of the CAPM states that the expected return on the j-th portfolio during period t is equal to the risk-free rate plus a percentage of the

TABLE 2

Sample of Mutual Funds

Fund Number	Code	Name	Annual Returns (%) Mean	Annual Returns (%) Standard Deviation
1	GI	Affiliated Fund, Inc.	10.65	14.74
2	I	Axe-Houghton Income Fund, Inc.	9.52	15.90
3	G	Axe-Houghton Stock Fund, Inc.	9.94	19.74
4	I	Century Shares Trust	10.37	20.11
5	G	Chemical Fund, Inc.	10.13	20.54
6	GI	The Colonial Fund, Inc.	9.77	16.25
7	B	Composite Bond & Stock Fund, Inc.	7.72	11.52
8	GI	Delaware Fund, Inc.	7.98	17.53
9	B	Dodge & Cox Balanced Fund	8.04	12.16
10	B	Eaton & Howard Balanced Fund	7.53	10.74
11	GI	Fidelity Fund, Inc.	11.46	17.21
12	G	Growth Industry Shares, Inc.	10.68	18.98
13	GI	The Investment Company of America	12.05	17.05
14	GI	Investment Trust of America	10.69	18.05
15	B	Investors Mutual, Inc.	7.30	11.63
16	I	Investors Selective Fund, Inc.	5.25	5.78
17	G	Johnston Mutual Fund, Inc.	9.68	15.22
18	B	Loomis-Sayles Mutual Fund, Inc.	7.37	11.12
19	G	Massachusetts Investors Growth Stock Fund	11.21	19.08
20	I	Mutual Investing Foundation–MIF Fund	9.72	15.82
21	G	National Investors Corporation	11.92	18.58
22	G	National Growth Fund	10.46	20.74
23	B	Nation-Wide Securities Company, Inc.	8.07	10.94
24	I	Puritan Fund, Inc.	11.19	15.79
25	B	The George Putnam Fund of Boston	9.07	13.42
26	G	Scudder Common Stock Fund, Inc.	9.72	17.85
27	I	Scudder Income Fund, Inc.	6.76	12.08
28	GI	Selected American Shares, Inc.	8.89	16.31
29	GI	State Street Investment Corporation	10.80	15.73
30	I	United Income Fund	10.17	16.50
31	B	Wellington Fund	7.13	11.18
32	I	Wisconsin Income Fund, Inc.	8.36	15.04

Weisenberger Classifications: G Growth; GI Growth and Income; I Income; and B Balanced.

expected risk premium on the market portfolio:

$$E(R_{it}) = RF_t + \beta_i[E(R_{mt}) - RF_t]. \quad (2)$$

Jensen's performance measure, α_i, is created by adapting Equation (2) into a regression of ex post, excess returns:

$$(R_{it} - RF_t) = \alpha_i + \beta_i(R_{mt} - RF_t) + \mu_{it}. \quad (3)$$

The estimate for α_i represents the constant periodic return that the portfolio generated above or below its expected risk premium. One advantage of this regression-based indicator is that it can be interpreted in a statistically meaningful way.

We initiated the testing procedure by regressing the historical returns for all thirty-two mutual funds against the returns to each of the six different market proxies. The average results for each index are listed in Table 3. The display gives the cross-sectional mean value for α_i, β_i, and R^2, as well as the average significance levels for the estimated parameters.

TABLE 3

Cross-Sectional Average Regression Results for Each Market Index

	Alpha	t-statistic	Beta	t-statistic	R-Squared
Index 1	−0.72	−0.84	0.79	16.85	0.86
Index 2	0.20	0.05	1.02	15.66	0.86
Index 3	−4.40	−2.67	2.18	10.21	0.76
Index 4	−3.96	−2.53	2.63	10.60	0.77
Index 5	−3.33	−2.16	2.65	10.30	0.76
Index 6	2.24	0.42	1.41	1.22	0.06

Two facts stand out. First, as measured by the coefficient of determination, only the first five indexes provide any significant explanation for movements in the mutual fund yields. This is interesting, because R^2 by definition can be interpreted as the degree of diversification in the underlying investment portfolio. Recall that Index 6 consists of all assets *except* common stock. Thus, we find that a market portfolio without stock is worthless for measuring the performance of funds investing primarily in corporate securities. Therefore, we will ignore any further results generated by Index 6.

The second fact of note from Table 3 is that the indexes including real estate, Indexes 3, 4, and 5, produce different inferences from those without real estate — Indexes 1 and 2. In particular, the average level of Jensen's alpha is not significant when real estate is left out, but the alpha is significantly negative otherwise. This dichotomy is in part the consequence of the dramatic decrease in risk when real estate is added to the market portfolio. In addition, the average level of beta is far different for the two index groups. As beta measures risk relative to the approximated market portfolio, the proxy with the average coefficient closest to 1.00 (Index 2) is also the one whose composition most closely reflects that of the funds themselves.

The implication of these results is that, at least on average, the specification of the market portfolio does make a difference. This finding coincides with Stambaugh's conclusion that the reasonableness of the market proxy depends to a great extent on the nature of the assets being tested. On the other hand, can we make a similar statement after examining investment performance on an individual basis?

For each of the five valid indexes, we ranked the mutual funds according to a three-step sorting routine. First, we separated the funds into three groups according to the statistical significance of their alpha estimates in order to make a more accurate distinction between expected and abnormal performance. We used the 95% confidence level to identify companies with significantly positive, insignificant, and significantly negative coefficient values, respectively. The second step in the process then consisted of sorting the funds by alpha level in descending order of magnitude within each of the three significance groups. Finally, we determined an overall ranking for the sample of funds by selecting the ordering from the Significantly Positive group first, the ordering from the Insignificant group second, and the ordering from the Significantly Negative group last. The advantage of sorting by significance class before looking at the numerical value of the alpha coefficient is that any spurious estimate of excess return is less likely to be interpreted as representing abnormal performance.

The rankings generated by each index, along with an average of the five separate results, are listed in Table 4. For purposes of comparison, the last column of Table 4 shows a performance ranking based solely on the mean return figures reported earlier. Thus, Table 4 gives five performance rankings that adjust for systematic risk as well as one ranking that ignores risk entirely.

To establish the degree of association between the orderings, Table 5 reports the Spearman rank correlation coefficients for each pairwise comparison. T-statistics appear in parentheses directly beneath these parameters.

With thirty-two mutual funds being ranked, the 0.05 critical value for each coefficient is 2.04. Therefore, we can be more than 95% certain that all estimated correlations for the risk-adjusted rankings are significantly positive. This suggests that the five indexes generate rankings that are substantially similar to one another.

TABLE 4

Summary of Mutual Fund Performance Rankings

Fund Number*	Index 1	Index 2	Index 3	Index 4	Index 5	Index Average	Mean Return
1	3	4	6	10	3	5.2	9
2	8	9	9	4	6	7.2	19
3	11	13	5	6	10	9.0	14
4	7	8	4	8	12	7.8	11
5	12	14	3	5	7	8.2	13
6	16	15	10	12	8	12.2	15
7	15	20	11	16	19	16.2	26
8	32	30	32	32	32	31.6	25
9	25	23	19	21	22	22.0	24
10	21	22	13	15	20	18.2	27
11	5	5	14	11	5	8.0	3
12	20	19	28	26	16	21.8	8
13	1	1	7	2	2	2.6	1
14	18	11	23	13	11	15.2	7
15	27	31	22	25	26	26.2	29
16	14	21	1	9	13	11.6	32
17	10	10	16	19	14	13.8	18
18	22	24	17	20	25	21.6	28
19	13	7	25	7	9	12.2	4
20	17	16	21	22	15	18.2	17
21	4	3	8	3	4	4.4	2
22	23	25	31	31	31	28.2	10
23	9	12	12	14	17	12.8	23
24	2	2	2	1	1	1.6	5
25	19	17	15	18	21	18.0	20
26	28	26	30	30	30	28.8	16
27	31	32	26	27	29	29.0	31
28	30	28	29	29	28	28.8	21
29	6	6	18	17	18	13.0	6
30	24	18	24	24	23	22.6	12
31	26	27	20	23	24	24.0	30
32	29	29	27	28	27	28.0	22

* Fund numbers as shown in Table 2.

TABLE 5

Spearman Rank Correlations Between Market Indexes and Mean Return

	Index 1	Index 2	Index 3	Index 4	Index 5	Mean Return
Index 1	1.0					
Index 2	0.9512 (9.93)	1.0				
Index 3	0.7749 (5.56)	0.6298 (3.99)	1.0			
Index 4	0.8669 (7.11)	0.8288 (6.38)	0.8710 (7.20)	1.0		
Index 5	0.8981 (7.87)	0.8992 (7.91)	0.7676 (5.46)	0.9216 (8.61)	1.0	
Mean Return	0.6367 (4.05)	0.7823 (5.66)	0.1580 (0.86)	0.4663 (2.72)	0.6221 (3.96)	1.0

T-statistics in parentheses beneath estimated coefficients.

Nevertheless, a closer inspection of the display reveals that a change from Index 2 to Index 3 produces a sharp reduction in the measured level of correlation. A further indication of this phenomenon is that the only two rankings that are not significantly correlated with one another are those generated by Index 3 and the mean returns. Once again, these findings correspond to the introduction of real estate into the market portfolio. Average mutual fund performance drops off precipitously when measured against Index 3, and there is a greater tendency for the individual rankings to be reversed as well.

Even though the Spearman coefficients reported above suggest a broad concurrence among the different market indexes, there is reason to believe that the agreement is weaker than it appears to be. Fuller [1981, pp. 39-41], citing the findings of Roll, notes that the measured correlation between two risk-adjusted performance rankings often will be overstated. He based this claim on the observation that the correlation coefficient between a risk-adjusted ranking and a ranking that ignores risk and uses just the mean returns usually will still be significant. The last row of Table 5 confirms this fact for all market proxies except Index 3.

One method of correcting for the upward bias this may impose is the partial rank correlation measure described in Kendall [1975, pp. 117-122]. This statistic estimates the degree of association between the rankings of two indexes (i and j) above and beyond their individual relationships with the ordering of a third index (r). Mathematically, the Kendall Partial Tau coefficient can be expressed:

$$\tau_{ij.r} = [\tau_{ij} - (\tau_{ir})(\tau_{jr})] \div [(1 - \tau^2_{ir})(1 - \tau^2_{jr})]^{0.5};\ i,j = 1, \ldots, 5, \qquad (4)$$

where τ_{ij} represents the Kendall rank correlation coefficient between rankings i and j. In this context, we let the index r denote the mean return rankings.

There are two interesting facts worth mentioning about the statistic summarized by Equation (4). First, as is true for any correlation measure, $\tau_{ij.r}$ is restricted to values between −1 and +1 and therefore can be interpreted in the usual manner. Second, as Kendall pointed out, "no tests of significance are yet known for partial τ" (p. 122). Thus, as a comparative technique, its usefulness is restricted.

Nonetheless, Table 6 contains the pairwise par-

TABLE 6

Kendall Partial Tau Correlations Between Market Indexes (Mean Return Ranking Held Constant)

	Index 1	Index 2	Index 3	Index 4	Index 5
Index 1	1.0				
Index 2	0.7938	1.0			
Index 3	0.5245	0.4731	1.0		
Index 4	0.6562	0.5804	0.7536	1.0	
Index 5	0.6834	0.6447	0.6124	0.7700	1.0

tial correlation coefficients for the risk-adjusted measures holding the mean return rankings constant. In general, we can conclude that the relationships between the market index orderings are indeed weaker than the Spearman coefficients imply. As before, Index 3 appears to be primarily responsible for the overall reduction in the association exhibited by the rank correlation structure.

Finally, we point to two additional pieces of evidence that the composition of the market proxy significantly affects the evaluation of individual portfolio performance. Both examples emphasize the issue of statistical inference about the performance measures, rather than using the simple numerical rankings considered above.

First, Table 7 gives a frequency distribution of the estimated alphas for each of the five market portfolios. For a particular index, all thirty-two funds are classified according to whether their performance measures are significantly different from zero. For each of the tests, we adopted a 95% confidence level.[4] Easily the most dramatic finding displayed in Table 7 is the change in inference that occurs between Index 2 and Index 3. Indeed, the results for Index 2 are almost identical to those described in Jensen's original study. Consequently, while the majority of funds perform as expected relative to a corporate security proxy, the same portfolios do quite poorly compared to an index dominated by real estate.

TABLE 7

Frequency Distribution of Estimated Alphas by Market Index

	Index 1	Index 2	Index 3	Index 4	Index 5
Significantly positive	0	1	0	0	0
Insignificant	25	30	10	13	16
Significantly negative	7	1	22	19	16
Total	32	32	32	32	32

All tests conducted at the 0.05 significance level

Table 8 offers a second confirmation of this phenomenon, showing the number of individual funds that have a reversal in the inference about the alpha coefficient when the market proxy is changed.[5] For instance, when switching from Index 2 and Index 3, we found that twenty-two out of the thirty-two mutual funds had different performance evaluations.

The inclusion of real estate in the market portfolio appears to be the driving force behind this result. The problem may be even more pervasive, however, because a shift from an all-stock index (Index 1) to one with stocks and fixed-income instruments (Index 2) still causes a reevaluation in 22% of the funds. Clearly, then, the specification of the market index has an obvious impact on performance measurement.

TABLE 8

Number of Inference Reversals Between Market Indexes

	Index 1	Index 2	Index 3	Index 4	Index 5
Index 1	—				
Index 2	7	—			
Index 3	15	22	—		
Index 4	12	20	3	—	
Index 5	9	16	6	3	—

All tests conducted at the 0.05 significance level

CONCLUDING OBSERVATIONS

The purpose of this paper has been to examine the historical performance of a sample of mutual funds with respect to several different definitions of the market portfolio. We have shown that, by creating a successively broader series of indexes, we can obtain a wide variety of inferences for the same set of individual funds. Our immediate finding is that the composition of the market proxy does matter for evaluating the returns to any given collection of assets.

A second important question might therefore be why this is the case. Roll [1980] has observed that true portfolio management performance cannot be gauged accurately if the "benchmark" index is a source of error in the measurement process. Even though all the indexes employed in this study were formed as value-weighted portfolios of the relevant asset classes, we can provide no assurance that they are "optimal." Benchmark errors may exist, but they are likely to be relatively unimportant, as we suggested in the introduction.

A more plausible contention is that the composition of the sample of securities to be tested is the element that actually matters and that the "market index" should be selected to reflect this universe. Of course, this is just what both Rosenberg and Stambaugh have said from a theoretical point of view.

The implication of these results is that the quest for the "all-inclusive" market portfolio is, from a practical standpoint, a futile one. A more enlightening endeavor would be to better define the market in terms of its relevant components.

REFERENCES

Blume, M. E. "The Use of 'Alphas' to Improve Performance." *Journal of Portfolio Management*, Fall 1984, pp. 86-92.

French, D. W., and G. V. Henderson, Jr. "How Well Does Performance Evaluation Perform?" *Journal of Portfolio Management*, Winter 1985, pp. 15-18.

Fuller, R. J. *Capital Asset Pricing Theories—Evolution and New Frontiers*. Charlottesville: Financial Analysts Research Foundation, 1981.

Ibbotson, R. C., and C. L. Fall. "The United States Market Wealth Portfolio." *Journal of Portfolio Management*, Fall 1979, pp. 82-92.

Jensen, M. C. "The Performance of Mutual Funds in the Period 1945-1964." *Journal of Finance*, May 1968, pp. 389-416.

Kendall, M. G. *Rank Correlation Methods*. London: Charles Griffin & Co., 1975.

Kon, S. J., and F. C. Jen. "Estimation of Time-Varying Systematic Risk and Performance of Mutual Fund Portfolios: An Application of Switching Regression." *Journal of Finance*, May 1978, pp. 457-475.

Roll, R.C. "Ambiguity When Performance is Measured by the Securities Market Line." *Journal of Finance*, September 1978, pp. 1051-1069.

——"A Critique of the Asset Pricing Theory's Tests." *Journal of Financial Economics*, March 1977, pp. 129-176.

——"Performance Evaluation and Benchmark Errors (I)." *Journal of Portfolio Management*, Summer 1980, pp. 5-11.

——"Performance Evaluation and Benchmark Errors (II)." *Journal of Portfolio Management*, Winter 1981, pp. 17-22.

Rosenberg, B. "The Capital Asset Pricing Model and the Market Model." *Journal of Portfolio Management*, Winter 1981, pp. 5-16.

Stambaugh, R. F. "On the Exclusion of Assets From Tests of the Two-Parameter Model." *Journal of Financial Economics*, November 1982, pp. 237-268.

[1] A more detailed description, as well as the primary sources, of all these asset classes can be found in Ibbotson and Fall [1979, pp. 85-90]. In particular, betas are reported for each of the components relative to the portfolio as a whole. Interestingly, the beta for total real estate holdings is indistinguishable from zero over the thirty-two year sample period.

[2] There is a double-counting problem where common stock is held by corporate investors, because there will be two claims on the same financial asset. Real estate and real estate mortgages present similar problems.

[3] Kon and Jen [1978] have cautioned against using standard performance measures when portfolio managers alter a fund's beta in an attempt to time general market movements. Thus, we invoked this restriction to protect against the obfuscating effects of a change in management policies.

[4] Recall that we ranked the mutual funds first by significance and then by magnitude of the alpha coefficients. Thus, it is easy to establish which specific companies had abnormal performance during the sample period. For example, the seven funds with significantly negative alphas relative to Index 1 would be those ranked in the last seven positions (i.e., 26 through 32) in Table 4.

[5] An inference reversal is defined as a change from significantly positive to insignificant, insignificant to significantly negative, and so on.