THE JOURNAL OF FINANCE • VOL. LI, NO. 1 • MARCH 1996

Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry

KEITH C. BROWN, W. V. HARLOW, and LAURA T. STARKS*

ABSTRACT

We test the hypothesis that when their compensation is linked to relative performance, managers of investment portfolios likely to end up as "losers" will manipulate fund risk differently than those managing portfolios likely to be "winners." An empirical investigation of the performance of 334 growth-oriented mutual funds during 1976 to 1991 demonstrates that mid-year losers tend to increase fund volatility in the latter part of an annual assessment period to a greater extent than mid-year winners. Furthermore, we show that this effect became stronger as industry growth and investor awareness of fund performance increased over time.

A TOPIC THAT HAS been of considerable recent interest within both the academic and professional communities is how portfolio managers adapt their investment behavior to the economic incentives they are provided. Most of the studies that address this issue (for example, Cohen and Starks (1988), Golec (1992), Grinblatt and Titman (1987, 1989a), Grinold and Rudd (1987), Kritzman (1987), and Starks (1987)) focus on the behavior elicited by incentive fee contracts. In this paper we argue that even without incentive fee contracts, the competitive nature of the mutual fund environment alone can affect a manager's portfolio decisions. Specifically, we suggest that viewing the mutual fund market as a tournament in which all funds having comparable investment objectives compete with one another provides a useful framework for a better understanding of the portfolio management decision-making process. Similar to the payoffs for golf and tennis competitions, the amount of remuneration that a fund receives for "winning" this tournament depends upon its performance relative to the other participants.

* Brown and Starks are from the University of Texas at Austin, and Harlow is from Fidelity Management & Research. The authors acknowledge the helpful suggestions of Brad Barber, Christopher Blake, Jeff Coles, Will Goetzmann, David Ikenberry, Chris James, Jon Karpoff, Paul Laux, Dan Quan, Henri Servaes, Meir Statman, Richard Thaler, Charles Trzcinka, Marc Zenner, and the participants of workshops at the University of Texas, Yale University, University of North Carolina, University of California-Davis, and University of Colorado-Denver. We would especially like to acknowledge the contributions of René Stulz (the editor) and an anonymous referee. Earlier versions of this paper have also been presented at the American Finance Association, Financial Management Association, and European Finance Association meetings. The opinions and analyses presented herein are those of the authors and do not necessarily represent the views of Fidelity Management & Research.

Past research substantiates that this is a reasonable characterization of the industry. In particular, citing evidence from separate surveys of households that had made recent mutual fund purchases, Goetzmann, Greenwald, and Huberman (1992) and Capon, Fitzsimons, and Prince (1992) note that past investment performance was the crucial input in the choice of which fund to acquire. Furthermore, Sirri and Tufano (1992) show that mutual funds earning the highest returns during an assessment period receive the largest rewards in terms of increased new investments in the fund. These additional contributions provide, in turn, increased compensation to the mutual funds' advisors as their rewards typically are determined as a percentage of the assets under management.

One interesting premise that this tournament interpretation of the fund industry suggests is that rational managers attempting to maximize their expected compensation may revise the composition of their portfolios depending on their relative performance during the year. While there will be times when such changes serve the best interest of the fund's investors, there will be other times when they may not. To this end, our goal in this paper is to test the hypothesis that given the profession's current system of assessing and reporting fund performance on an annual basis, managers with either extremely good or bad relative returns at mid-year have incentives to alter the investment characteristics of their portfolios. The central testable implication that emerges from our analysis is that the set of funds most likely to be "losers" in the final tournament results will see their risk levels increase relative to the group of probable "winners."

Testing this conjecture with monthly return data for more than 330 growth-oriented mutual funds dating from 1976, we find that funds classified as relative losers during an interim performance assessment period do indeed increase portfolio risk to a greater degree than do interim winners. This proved to be more uniformly true for the newer, less well-established funds in the sample, as well as for those funds that had been consistent losers and winners in past years. Furthermore, these behavioral differences, while present to some extent for the entire 1980–1991 test period we examine, become more pronounced during the last five years of the sample when the level of investment in the fund industry proliferated tremendously. Finally, we also demonstrate that these results are robust with respect to the timing of the interim performance review as well as any "window dressing" effects that may occur in the month of December.[1]

The structure of the paper is as follows. In the next section we review the economic tournament literature, with a particular focus on the qualitative implications for the investment management industry, and then motivate our testable hypotheses. Section II describes the data and methodology we employ in our empirical investigation, while Sections III and IV detail the findings of

[1] Studies that analyze why money managers engage in strategic portfolio rebalancing activities such as window dressing include Ritter (1989) and Lakonishok, Shleifer, Thaler, and Vishny (1991).

the classification and regression tests we use to support our conclusions. In a final section, we summarize the results and offer some concluding remarks.

I. Mutual Fund Performance and Tournaments

A. The Economics of Tournaments

As it has developed over the past decade, research on the economics of tournaments can be considered a subset of the literature on agency theoretic contracting. Specifically, the primary focus of this body of work has been on the normative aspects of performance-based compensation schemes. Examples of this include Nalebuff and Stiglitz (1983) and Rosen (1986), who analyze the economic incentives that different types of tournaments provide to participating agents. In fact, many compensation and reward structures can be viewed as tournaments. For example, most hierarchical organizations are such that employees at one level compete for a smaller number of positions at the next higher level, which invariably results in a reward system based on relative performance assessments. A tournament reward structure is particularly appropriate in situations where an agent's effort is not observable and the performance of all agents depend on a common economic "shock." In such a case, relative performance measures allow the principal to separate some of the variation in outcome due to the state of nature from the agent's contribution.[2]

Because of the normative approach adopted by the tournament literature, there is to date little empirical evidence on whether the incentive effects of these compensation structures actually elicit the desired behavior. Furthermore, those studies that do exist fail to reach a consensus. Two recent investigations typify this impasse. First, in an experimental study, Bull, Schotter, and Weigelt (1987) provide mixed support for the ability of the tournament model to converge to theoretic equilibrium levels. Second, using historical data on the prize money allocated at golf tournaments, Ehrenberg and Bognanno (1990) conclude that the level and structure of the compensation on the Professional Golfers Association tour does influence player performance.

B. Mutual Funds as Tournaments: Intuition and Testable Hypotheses

The underlying premise of this study is that the market for portfolio management services, and mutual funds in particular, can be viewed as a multiperiod, multigame tournament. Indirect support for this contention comes from observing the number of business publications and information services (e.g., *Business Week, Barron's, Forbes, Money*, Morningstar Mutual Fund Services, Lipper Analytical Services) that regularly rank funds according to their

[2] Lazear and Rosen (1981) and Green and Stokey (1983) examine the conditions under which a tournament structure will outperform other reward schemes in mitigating moral hazards. These conditions, which include risk aversion on the part of the participants, a common shock component, and a large number of agents, are satisfied by the mutual fund industry.

return performances. Although some of these publications and services rank funds every quarter, the most critical rankings are based on annual performance and are usually produced at the end of the calendar year.

Sirri and Tufano (1992) find that investors respond to these rankings. They document that flows of new capital by investors are related to the performance of the portfolio, with the mutual funds that rank highest in relative return receiving a larger share of new investment inflows in subsequent periods. Furthermore, they show that the relationship between fund inflows and performance is not symmetric; mutual funds that performed worse than the competition do not experience as significant an outflow of invested capital.[3] Fund inflows and outflows are important in this context because the compensation for mutual fund advisors is often structured as a flat fee plus a percentage of the level of assets under management. Thus, as new money comes into or leaves a fund, the compensation to the manager will change proportionally, but it will only decline until the base level is reached.

If managers view themselves as being participants in a series of tournaments in which their funds are ranked each year, their behavior should show certain systematic, measurable tendencies. One such tendency is for managers to change the risk profile of the fund during the course of the year (i.e., tournament) depending on the fund's relative accumulated performance. Specifically, the Sirri-Tufano results imply that managers who earn relatively higher returns receive larger proportional inflows of funds and managers who earn poorer returns are not penalized in an offsetting manner. Thus, the decisions made by a fund manager are predicated on a call option-like payoff structure that is influenced by their relative performance ranking.

This option-like compensation arrangement is similar in spirit to the incentive fee contract that Grinblatt and Titman (1987, 1989a), Kritzman (1987), and Starks (1987) discuss. Those studies note that the convexity of such a reward system provides incentives for managers to alter the risk of their portfolios. Thus, while this option-like compensation scheme may create different incentives for different funds, it can be argued that its presence is sufficient to entice any manager to consider changing the risk of his or her portfolio before the end of the assessment period.

For those who have performed poorly, the incentive will be to increase their relative risk level given that they can only benefit by trying to improve their ranking by year end. That is, managers finding themselves positioned at an interim assessment period as losers (i.e., having performed worse than average) will need to generate a return over the remainder of the tournament that is sufficient to make up their first period "deficit." One way in which they can attempt to achieve this result is by altering fund risk during the rest of the

[3] Goetzmann and Peles (1994) confirm these investment patterns and suggest that they can be explained by cognitive dissonance on the part of mutual fund investors. That is, faced with a difficult investment choice, people tend to buy funds that have done well in past performance rankings. However, after purchasing a fund that subsequently performs poorly, these same investors are reluctant to admit their mistake, choosing instead to revise their perceptions of the ranking data.

competition to a level that is larger than that expected for the interim winners (i.e., managers who performed better than average).

On the other hand, those managers who have high interim returns compared to their peers will want to maintain those high returns and will be less inclined to take risky positions and may even consider scaling back their risk positions in an attempt to "lock in" their present return level. Unfortunately, to the extent that they anticipate what those managers ranked below them might do, it may be necessary for them to increase risk as well, but they do not need to increase risk to the same extent as do the losers. Representing the interim loser and winner strategies by the subscripts L and W, respectively, and the corresponding portfolio risk levels in the first and second subperiods by σ_1 and σ_2, this reasoning leads to our central prediction:

$$(\sigma_{2L}/\sigma_{1L}) > (\sigma_{2W}/\sigma_{1W}). \tag{1}$$

That is, the "risk adjustment ratio" for the interim losers will be greater than that for the interim winners. That losers adopt higher risk-choice strategies than winners is consistent with the work of Bronars (1987), McLaughlin (1988), and Ehrenberg and Bognanno (1990); Bronars, for instance, finds that losing sports teams will often take greater risks toward the end of the game than they would under less time-constrained circumstances.[4]

As is true with athletic competition, the reality of the mutual fund tournament is complex and the precise strategic behavior of the losers is difficult to predict in isolation. Thus, it is important to note that equation (1) represents a general tendency rather than an exact prediction; it does not involve a forecast of whether risk in the second subperiod is larger or smaller for either type of manager. The exact adjustment to a given fund's risk level will depend on factors such as the size of the return deficit faced by the interim loser, the amount of information that the managers possess about each other's probable reactions, and, quite possibly, the amount of general market volatility. It should be the case, however, that the post-assessment relative risk increase (decrease) should be greater (smaller) for interim losers than for interim winners. In addition, notice that managers can alter the riskiness of their overall investment in several ways, including hedging with derivatives, increasing their relative allocation to cash equivalents, or physically rebalancing the portfolio to alter its inherent risk level.

[4] There are two other important advantages of employing a tournament approach to enhance our understanding of the mutual fund industry. First, benchmark and market-timing problems common to mutual fund performance evaluation studies (e.g., Grinblatt and Titman (1989b)) do not arise in the tournament context. By considering the effects of relative rankings in total returns across similar funds, we do not have to address the problem of whether the benchmark is mean-variance efficient. Second, the survivorship bias problem that Brown, Ibbotson, Goetzmann, and Ross (1992) and Brown and Goetzmann (1993) describe is not an issue in that if poorer performing funds drop out of our sample, the bias would be against finding the result that the model predicts.

A second hypothesis emerges from the demographic characteristics of the various funds themselves. In particular, there are at least three reasons to believe that the size and age of the fund will directly affect a manager's willingness or ability to alter the portfolio's investment characteristics. First, while a large-fund manager who is an interim loser may want to alter risk substantially, he or she may be unable to make the necessary revisions in a timely manner because of any number of investor clientele or liquidity reasons. Managers of smaller, newer funds might not be similarly constrained. Second, in order to survive, a smaller, newer fund has the incentive to pursue new investments more aggressively than would a portfolio with considerable existing assets to protect. Finally, investors are more likely to be influenced by bad short-term performance for a fund with a brief track record than for one with an extensive history. This might motivate new-fund interim losers to be more proactive in attempting to reverse mid-tournament losses. Thus, we also predict that (1) is more likely to hold for small, new funds than those that are large and well entrenched.

II. An Empirical Investigation

A. Data

The data for this paper consist of monthly returns to 334 growth-oriented mutual funds followed by Morningstar Mutual Fund Services and contained in their data base. We have restricted our analysis to funds with the growth classification for two reasons. First, as evidenced by the attention they receive from both the financial press and direct investor involvement, they are the most widely followed and often-ranked class of publicly traded funds. As such, they would appear to fit the qualitative criteria of the tournament model quite well. Second, as McDonald (1974) and Radcliffe (1990) note, it is also the fund class that is likely to have a high a priori tendency toward taking risky positions and therefore provides an excellent opportunity to either refute or establish the validity of our central conjecture.

For each of the funds in the sample, return data are available from the most recent of either 1976 or the inception of the portfolio and continue until the end of 1991. Following the most prevalent industry practice, we make the simplifying assumption that the tournaments are held on an annual basis and that all funds have their performance judged on the same cycle. However, we include a fund in a yearly tournament only if it has return data available for the entire year; funds initiated during a year are removed from that tournament. Table I summarizes several relevant aspects of this sample, such as the number of funds included by year, the total dollar investment in those funds, and the median return and median return standard deviation for each tournament. The median return statistics are particularly notable for their wide range of values. This degree of volatility is important because we use a variant of this measure to delineate the winner and loser samples in the annual tournaments. As shown in the next-to-last column of the display, in some years

Table I

Descriptive Statistics for a Sample of 334 Growth-Oriented Mutual Funds, 1976–1991

Summary information is reported for the sample of mutual funds with a growth orientation, as maintained by Morningstar Mutual Fund Services. A fund is only included in an annual sample if it has return data for the entire year.

Year	Number of Funds	Total Dollar Investment[a]	Median Return[b]	Median Std. Deviation[b]
1976	109	$ 8.93	21.97%	16.07%
1977	112	9.11	1.75	11.60
1978	118	10.71	12.12	21.27
1979	120	14.12	30.00	16.70
1980	121	13.99	35.32	21.48
1981	126	16.97	−1.56	16.07
1982	132	27.89	27.61	19.61
1983	142	29.62	21.11	12.54
1984	156	40.40	−2.29	15.87
1985	174	54.16	29.12	13.65
1986	201	60.01	14.50	17.04
1987	234	65.51	2.67	31.38
1988	266	83.59	14.86	10.81
1989	287	83.23	26.71	11.95
1990	311	129.24	−4.89	19.23
1991	334	134.25	34.31	17.08

[a] Total year-end dollar value of all funds classified as growth oriented, expressed in billions.
[b] Statistics are reported for the median values of the annual fund samples.

funds needed to increase their total values by more than one-third to achieve winner status while in other years breaking even was sufficient.

The most remarkable aspect of Table I, though, is the tremendous increase in the sample size during the whole 1976–1991 period in general and the last eight years in particular. Whether viewed in terms of the steep ascent in the number of funds in operation or the rapid appreciation of the total amount of assets under management, it is clear that this sector of the industry increased in size severalfold in the matter of just sixteen years. However, upon closer examination, it becomes apparent that the majority of this growth occurred during the last half of the period. For instance, the 1976 base of 109 growth funds had expanded by over 30 through 1983, but then more than doubled by the end of 1991. The figures for the total dollars invested in growth-oriented funds for the same two subperiods reflect this trend in an even more dramatic way. In the following section, we test whether the tournament incentive effects we have hypothesized became more pronounced with this increase in investment activity.[5]

[5] The pattern of expansion for growth funds documented in Table I was not unique in that it was mirrored by the entire fund industry. Data from the *1993 Mutual Fund Fact Book* reveals that between 1976 and 1983 the total net assets held by all equity funds changed from $34.3 to $73.9

B. Methodology

To test a generalized form of equation (1), we develop two variables from the fund return data base. First, for each of the sixteen annual samples, we create subgroups of interim winners and losers according to a fund's relative return performance between January and month M. Specifically, for each fund j in a given year y, we calculate the M-month cumulative return as follows:

$$RTN_{jMy} = [(1 + r_{j1y})\ (1 + r_{j2y}). . .(1 + r_{jMy})] - 1 \tag{2}$$

where r_j is the monthly change in the fund's net asset value plus distributions as reported by Morningstar. In our analysis, we allow month M to vary between April and August and so RTN is measured over periods ranging from four to eight months. After calculating a separate set of RTN for each sample year, the funds in that tournament are ranked from highest to lowest and the winner and loser appellations are attached according to the fund's ranking. For comparative purposes, we calculate two separate classifications systems: (i) whether funds are above ("winners") or below ("losers") the median value of RTN, and (ii) whether funds are in the upper quartile or lower quartile of the ranking.[6]

The second variable we need to test the hypothesis that winners and losers make different adjustments to the investment characteristics of their portfolios is a ratio of each fund's volatility measured before and after the interim assessment period. With the interim assessment date at month M, the fund j risk adjustment ratio, RAR, for a particular year y is calculated as:

$$RAR_{jy} = \sqrt{\left(\frac{\sum_{m=M+1}^{12} (r_{jmy} - \bar{r}_{j(12-M)y})^2}{(12 - M) - 1}\right)} \div \sqrt{\left(\frac{\sum_{m=1}^{M} (r_{jmy} - \bar{r}_{jMy})^2}{M - 1}\right)} \tag{3}$$

with the deviations in the numerator and denominator calculated relative to the mean return over the relevant subperiod. For each tournament y, equation (3) measures the ratio of the j-th fund's standard deviation after the month M interim performance assessment relative to its standard deviation before that date. Consequently, the empirical adaptation of the prediction in (1) is that this ratio should be significantly larger for funds labeled as losers at month M than for those designated as winners.

With these definitions, we are able to create a (RTN, RAR) pair for every fund in each of the twelve annual tournaments spanning the years 1980–1991.[7] The basic test procedure is then to generate a 2 × 2 contingency table

billion, and then to $367.7 billion by the end of 1991. Similarly, the total number of equity funds grew from 274 in 1977, to 374 by 1983, and then to 1,216 by 1991.

[6] Notice that the median-based definition of winners and losers has the advantage of using all of the funds in the sample, but may result in misleading inferences to the extent that some of the winner-loser distinctions in the middle of the rankings are spurious. On the other hand, the quartile-based definition does a much better job of isolating "extreme" winners and losers, but uses only half of the available data.

[7] An hypothesis we investigate in a subsequent section asserts that volatility adjustments are more pronounced for those managers who have been consistently classified as losers. One common

in which each pairing is placed into one of four cells: high *RTN* (i.e., winner)/high *RAR*; low *RTN* (i.e., loser)/high *RAR*; high *RTN*/low *RAR*; low *RTN*/low *RAR*. As with the earlier use of *RTN* to define interim winners and losers, high and low levels of *RAR* are determined as being above and below the median *RAR* value, respectively, for a given tournament. The null hypothesis in our tests is that the percentage of the sample population falling into each of these four cells is equal (i.e., 25 percent), which implies that the two classifications are independent. The alternative hypothesis consistent with equation (1) is that the low *RTN*/high *RAR* and high *RTN*/low *RAR* cells would have measurably larger frequencies than the other two outcomes. The statistical significance of these frequencies is established with a chi-square test having one degree of freedom (i.e., the product of one unrestricted row and one unrestricted column in the contingency table).

We test the risk adjustment prediction over three time frames. To get a sense of the longer-term dynamics of the industry, we group the annual tournament samples into four three-year intervals, two six-year intervals, and one twelve-year interval (i.e., the entire sample period). One problem with aggregating results from different years is that the annual ranking of the *RTN* and *RAR* variables might not transfer from one tournament to another. For instance, an M-month return of two percent might indicate a winner in one year but a loser in the next. As a solution to this problem we normalize both variables by netting the mean value of that variable from the fund-specific observations and dividing the difference by the variable's standard deviation. This allows us to extend the annual tournament concept to more meaningful time frames.

Finally, we also perform our tests with and without returns from the month of December. The rationale for excluding December returns is that, in addition to the compensation-related incentives for which we are testing, managers may have other motives for altering their portfolios in a way that could affect risk. One such reason involves the year-end manipulation of the portfolio, commonly known as "window dressing," whereby managers purportedly acquire new positions or liquidate existing ones to fine-tune the composition of the fund for accounting or reporting purposes. Of course, with maneuvering of this type comes the potential for increased fund volatility that has nothing to do with our investigation. Thus, to be conservative, we calculate *RAR* both with and without December returns.[8,9]

metric used by the industry to indicate sustained performance is to also rank managers on the basis of three- or five-year cumulative returns. Consequently, in order to mimic this procedure, we limit our analysis to the most recent twelve annual tournaments for which we have at least four prior years of return data.

[8] Procedurally, when we exclude December returns, we calculate an adjusted version of equation (3) by summing the squared deviations in the numerator over the interval $[M + 1, 12 - 1]$ and dividing that sum by $\{(12 - M - 1) - 1\}$. The mean return used in creating these adjusted deviations is also generated for the $[M + 1, 12 - 1]$ interval.

[9] Notice that we do not also exclude January returns in our analysis. The reason for this is that, within the tournament framework, the manager in month M would already know what had occurred in January and this would form part of his or her *RTN* ranking. Given that it is the

III. Empirical Results: Basic Tournament Findings

A. Altering Portfolio Risk: Whole Sample Period

Table II shows cell frequencies for each of several different experimental designs using the entire sample of data. We calculate separate contingency tables for all 20 combinations of performance assessment month $M = 4, 5, 6, 7$, and 8; median and quartile rankings of the *RTN* variable; and inclusion or exclusion of December returns. In interpreting these results, it should be kept in mind that merely rejecting the null hypothesis of equal frequency among all four cells in a particular table does not by itself constitute evidence in favor of the assertion in (1). If, for instance, the low *RTN*/high *RAR* and high *RTN*/low *RAR* cells have frequencies significantly less than 0.25, the results would indicate exactly the opposite of what we predict. Only when the frequencies of those cells are significantly above 0.25 does our proposition find support.

Panel A of Table II lists results for winners and losers categorized by the median value of the *RTN* variable, with Panels A1 and A2 indicating the consequences of either including or removing the set of December observations, respectively. Perhaps the most noticeable thing about these findings is that, with the exception of the interim performance rankings conducted at the end of April (i.e., $M = 4$), all of the cell frequencies are significantly different from their expectations under the null and in the predicted direction. Interestingly, the presence or absence of December returns appears to make little difference in this conclusion. Additionally, while every interim assessment month from May through August generates test statistics that are significant at conventional levels, it is clear that the July marking date (i.e., $M = 7$) is associated with the largest divergence in the cell values. This finding is consistent with the notion that managers revise their investment strategies within the month following the release of the second quarter performance rankings and consistent with the prima facie observation of when the financial press and information services actually report this information.

The interpretation of the results for the quartile-based winner/loser rankings in Panels B1 and B2 are very similar to those just described. In particular, we again see that the cell frequencies are uniformly significant in the predicted direction except for the case of performance assessment in April when we can't reject the null. Furthermore, the frequencies for the loser/high *RAR* cell are comparable to those in Panel A, and slightly larger for the July assessment date. Surprisingly, the quartile-based frequencies do not appear to be substantially different than the median-based ones, suggesting that the notion of being an "extreme loser" has no substantive content. As before, the frequencies in

post-interim ranking behavior that we attempt to predict, there is no reason to exclude any data from the pre-ranking period. Further, to the extent that any window dressing in January—even if it is simply an unwinding of window dressing from the previous December—raises fund volatility in the pre-ranking period, it will be harder to establish the validity of an empirical form of (1).

Table II

Frequency Distributions of a 2 × 2 Classification of the Risk Adjustment Ratio and "Winner/Loser" Variables, 1980–1991

Cell frequencies are reported for a 2 × 2 classification scheme involving the rank-ordered variables: (i) the Risk Adjustment Ratio (*RAR*); and (ii) the compound total return through the first *M* months of the year (*RTN*). We employ five different values for the interim performance assessment date: $M = 4, 5, 6, 7$, and 8. *RAR* is the ratio of standard deviations measured before and after month *M*. We construct and normalize the data for the classifications on a yearly basis from monthly returns to 334 growth-oriented mutual funds and aggregated for the 1980–1991 sample period. We divide the funds into four groups on a yearly basis according to (i) whether *RTN* is below ("low" or "loser") or above ("high" or "winner") the median (Panel A) *or* in the highest ("winner") or lowest ("loser") quartile (Panel B), and (ii) whether *RAR* is above ("high") or below ("low") the median. We also list results for *RAR* rankings calculated both with and without December returns.

		Sample Frequency (% of Observations)					
		Low *RTN* ("Losers")		High *RTN* ("Winners")			
Assessment Period[a]	Observations	"Low" *RAR*	"High" *RAR*	"Low" *RAR*	"High" *RAR*	χ^2	*p*-value[b]
Panel A: Winners/Losers Ranked by Median							
		A1: All Monthly Returns					
(4, 8)	2484	24.32	25.60	25.56	24.52	1.36	0.244
(5, 7)		22.83	27.09	27.09	22.99	17.42	0.000
(6, 6)		22.46	27.46	27.46	22.62	23.97	0.000
(7, 5)		22.14	27.78	27.70	22.38	29.79	0.000
(8, 4)		23.43	26.49	26.45	23.63	8.58	0.003
		A2: Excluding December Returns					
(4, 8)	2484	24.52	25.40	25.36	24.72	0.58	0.446
(5, 7)		23.27	26.65	26.61	23.47	10.57	0.001
(6, 6)		22.79	27.13	27.13	22.95	18.10	0.000
(7, 5)		22.22	27.70	27.66	22.42	28.49	0.000
(8, 4)		23.39	26.53	26.53	23.55	9.30	0.002
Panel B: Winners/Losers Ranked by Quartile							
		B1: All Monthly Returns					
(4, 8)	1233	25.61	24.31	26.01	24.07	0.05	0.821
(5, 7)		23.20	26.68	28.95	21.17	15.67	0.000
(6, 6)		22.63	27.25	28.63	21.49	17.05	0.000
(7, 5)		21.65	28.22	27.41	22.71	15.69	0.000
(8, 4)		22.71	27.17	28.06	22.06	13.49	0.000
		B2: Excluding December Returns					
(4, 8)	1233	25.69	24.23	26.18	23.91	0.08	0.777
(5, 7)		23.20	26.68	28.39	21.74	12.67	0.000
(6, 6)		22.47	27.41	28.47	21.65	17.05	0.000
(7, 5)		21.57	28.30	27.58	22.55	17.07	0.000
(8, 4)		21.98	27.90	27.90	22.22	16.59	0.000

[a] The performance assessment period is listed as (*M*, 12-*M*), where *M* indicates the month of the interim assessment and 12-*M* is the rest of the year.

[b] The χ^2 statistic is calculated based on a null hypothesis that each cell should receive an equal distribution (i.e., 25 percent) of the sample.

Panel B2 indicate that nothing really changes when December returns are excluded.[10]

B. Altering Portfolio Risk: Temporal Dynamics

Although the preceding findings support the notion that losers increased portfolio risk more than winners over the entire twelve year assessment period, it is not necessarily the case that this is a pervasive result. In Table III we have chosen a single, conservative experimental design (i.e., variables classified by the median, performance assessment in month 7, and December returns excluded) and listed the cell frequencies corresponding to the several temporal sample partitions described above. To make the comparisons more accessible, Panel A replicates from the previous display the results for this specification over the entire period.

The striking conclusion suggested by the cell frequencies listed in the last two panels is that the propensity for interim losers to alter portfolio risk in no way spans the whole 1980–1991 interval. For instance, when the sample is divided into two six-year halves (Panel B), it is readily apparent that only in the most recent partition do the loser/high *RAR* and winner/low *RAR* frequencies turn significant and in the predicted direction. Conversely, the differences between cells are not significant during the 1980–1985 subperiod. Panel C breaks down this tendency further by showing that it is the two most recent three-year subintervals that drive the conclusion for the whole sample, with only the earliest three years acting in a reliably contrary manner.[11]

A different way of seeing this phenomenon can be inferred from a comparison of the average *RAR* statistics for the winner and loser fund groups. Over the entire sample period, the median ratios of the pre- and post-assessment date standard deviations for losers and winners are 1.20 and 1.12. This increment of 0.08, which produces a significant "difference in medians" chi-square statistic of 29.77, supports the proposition that, on average, managers with the worst interim performance increase portfolio risk by the largest amount. Interestingly, the typical interim winner also appears to increase

[10] One aspect of the results listed in Panel A of Table II that may seem curious is that the frequencies for the "loser"/high *RAR* and "winner"/low *RAR* cells are always identical. Rather than being an unlikely coincidence, this pattern is simply an artifact of the way in which we classify the sample. Specifically, given that we segment both the *RTN* and *RAR* variables by their median values, it is necessarily the case that two "loser" cells and the two "winner" cells must equal 0.50 for even sample sizes and approximately 0.50 for odd ones (depending on the cell to which the "odd" observation is allocated). Equivalently, the low *RAR* and high *RAR* cell pairs must also sum to 0.50. Consequently, when creating a 2×2 contingency table by first dividing the sample in half by *RTN* and then in half again by *RAR*, it must be the case that the diagonal cells are equal in value. For the quartile-based ranking of *RTN*, however, this need not be true, as evidenced by the unequal cell frequencies in Panel B.

[11] We also examine the cell frequencies for each of the annual tournaments separately and, not surprisingly, find that they are consistent with the aggregated findings listed in Table III. In particular, seven of the eight most recent years produce results supportive of our risk change hypothesis, including the last five tournaments where the chi-square statistics are significant beyond the 0.001 critical level.

Table III

Temporal Partitions of the Risk Adjustment Ratio and "Winner/Loser" Variable Classifications Using Median Values: December Observations Excluded

Listed below are the cell frequencies for a 2 × 2 classification scheme involving the rank-ordered variables: (i) the Risk Adjustment Ratio (*RAR*); and (ii) the total return (*RTN*) through the first seven months (i.e., $M = 7$) of the year. We construct and normalize the data for the classifications on a yearly basis from monthly returns to 334 growth-oriented mutual funds during the period 1980 to 1991. We divide the funds into four groups on a yearly basis according to whether *RTN* is below ("low" or "loser") or above ("high" or "winner") the median and whether *RAR* is above ("high") or below ("low") the median. Results are reported for four sample partitions: (i) the whole sample, (ii) two six-year periods, (iii) four three-year periods, and (iv) twelve annual periods. *RAR* values are calculated and ranked excluding December observations for each sample year.

		Sample Frequency (% of Observations)					
		Low *RTN* ("Losers")		High *RTN* ("Winners")			
Sample Period	Observations	"Low" *RAR*	"High" *RAR*	"Low" *RAR*	"High" *RAR*	χ^2	*p*-value[b]
Panel A: Whole Sample							
1980–1991	2484	22.22	27.70	27.66	22.42	28.49	0.000
Panel B: Six-Year Periods							
1980–1985	851	26.09	23.85	23.74	26.32	1.98	0.160
1986–1991	1633	20.21	29.70	29.70	20.39	57.72	0.000
Panel C: Three-Year Periods							
1980–1982	379	27.18	22.69	22.43	27.70	3.61	0.057
1983–1985	472	25.21	24.79	24.79	25.21	0.03	0.854
1986–1988	701	22.11	27.82	27.82	22.25	8.90	0.003
1989–1991	932	18.78	31.22	31.22	18.78	55.78	0.000

[a] The χ^2 statistic is calculated based on a null hypothesis that each cell should receive an equal distribution (i.e., 25 percent) of the sample.

fund volatility, but by a substantially smaller amount. Once again, however, the most recent six years are responsible for this result; the loser/winner *RAR* differentials for the 1980–1985 and 1986–1991 periods are 0.00 (= 1.32 − 1.32) and 0.11 (= 1.13 − 1.02), with the latter being significant at a critical level less than 0.001.

The complex dynamics of the risk adjustment process become even more apparent when the median *RAR* statistics are measured over successively shorter time intervals. For instance, dividing the sample period into four nonoverlapping three-year periods generates respective loser risk adjustment ratios of 1.29, 1.33, 1.29, and 0.98, while the corresponding winner ratios are 1.36, 1.32, 1.27, and 0.84. Notice that the difference in median values between losers and winners is positive in all but the initial 1980–1982 period, with differences during 1986–1988 and 1989–1991 being significant at the 0.003

and less than 0.001 critical levels, respectively. Furthermore, it is interesting to note that during the final subperiod the median risk change ratio for both types of funds is less than one. This means that over this three year period the average fund manager actually decreased risk in the aftermath of a mid-year performance ranking. As we mentioned earlier, however, this observation does not refute our central prediction, which only specifies that the losers' *RAR* should exceed the winners' *RAR*. Thus, a smaller decline in risk for fund managers classified as interim losers rather than winners (i.e., 0.98 versus 0.84) is consistent with the forecast that these managers will bear a relatively higher proportion of their original risk level during the second half of the year.

The median winner and loser *RAR* statistics from the twelve separate annual tournaments tell a similar, if considerably noisier, story. These values, which are not reproduced here to conserve space, are most notable for their considerable range of values. For example, the series of median ratios for the loser funds vary from 0.54 in 1980 to an extreme of 2.54 in 1987, with the winners' ratios ranging from 0.57 to 2.38 in those same years. In all, four (five) of the twelve loser (winner) median ratios fall below one, once again highlighting the fact that fund managers do not uniformly raise the absolute risk level of their funds after an interim assessment date. More importantly, though, the pattern of the difference in loser and winner ratios observed above still holds with two of the first four differentials being significantly negative at the 0.03 critical level or better and the last five being significantly positive. In fact, a "runs" test on the sequence of twelve ratio differentials indicates only four runs and confirms that this pattern of behavior is very likely not random.[12]

The most probable explanation for why the last part of the sample period provides the greatest support for our adverse-incentive hypothesis lies in the growth of the mutual fund industry and the increased investor scrutiny that attended that expansion. Ultimately, the rationale for our tournament framework rests on the empirical observation, reported by Goetzmann, Greenwald, and Huberman (1992) and Capon, Fitzsimons, and Prince (1992), among others, that investors focus primarily upon past investment returns in making their fund selection decisions. This information, however, did not enjoy wide, public dissemination until the mid- to late-1980s when most of the major business periodicals began to publish their mutual fund performance rankings (e.g., 1986 for *Business Week*, 1988 for *Money*). As a further indication of the

[12] One potential reason for the large dispersion in *RAR* values across the various subperiods (e.g., 1.32 versus 1.13 for the interim loser funds during 1980–1985 and 1986–1991, respectively) is that these statistics reflect changes in the volatility of general market returns. To investigate this possibility, for each sample subperiod we produce a corresponding set of market *RAR* statistics calculated in an analogous fashion using monthly returns to the Standard & Poor's 500 index. If anything, these market risk ratios are even more volatile than those for the winner and loser funds. For instance, the annual market *RAR* statistics range from 0.41 to 2.36, with a standard deviation for the twelve *RAR* observations of 0.659 compared to 0.535 for both of the fund samples. Further, the Spearman rank correlation coefficient between the market series and the winner and loser funds are 0.898 and 0.905, respectively, indicating the strong influence that market forces appear to have on the absolute levels of the fund risk adjustment ratios.

increased media attention to this issue during the sample period, a survey of the *New York Times* reveals that there were 334 articles concerning fund performance rankings from 1980–1985, but then 607 over the subsequent six years.

Of course, the heightened level of media awareness was itself most likely a consequence of the proliferation of both investors and available investment outlets that occurred during this era. In the discussion associated with Table I, we noted the dramatic increase in both the total assets under management and net flows into growth-oriented funds over the years we studied. Indeed, this pattern of expansion is reflective of a burgeoning financial services industry that became increasingly competitive over the last decade. In a study of the economic forces underlying the market for mutual funds, Baumol, Goldfeld, Gordon, and Koehn (1990) establish substantial decreases in the concentration ratio over the latter part of their 1982–1987 sample period, which lead them to conclude that competition in the mutual fund industry was intensifying. Thus, given the result that those funds with the best performance attract the most new assets, it is natural to suggest that as investor awareness of fund performance increased, so too did the pressure for a rapidly expanding group of competitors to be counted among the set of winners.[13]

C. Fund Size, Entrenchment, and Load Effects

Our second hypothesis is that managers in newer and smaller funds have more incentives to change risk and be less constrained by market forces from taking action than do larger, more established operations. We test this notion in two ways. First, we segment the 109 growth funds that were in the data base since its inception in 1976 from the rest of the sample as being indicative of "well-entrenched" portfolios. We then compare their cell frequencies with those corresponding to the remaining collection of "new" firms. Second, we also divide the funds according to their total asset values in a particular annual tournament, with "small" ("large") funds being defined as those with assets falling below (above) the median value for that year. These results are listed in Table IV for both sample stratifications using the entire 1980–1991 period and the two six-year subintervals.

Panel A of this display reports cell frequencies for the entrenched versus new funds. Consistent with our prediction, the results indicate that it is the newer firms that provide more of the separation between the portfolio risk adjustments of winner and loser funds. For instance, a "difference in proportions" test for new versus entrenched funds during the 1980–1991 period yields a

[13] To confirm the ultimate impact of this trend, we estimate separate regressions of the series of twelve annual chi-square statistics from the "difference in medians" tests reported earlier against three different variables: a time indicator, the change in the annual number of sample funds, and the change in the annual dollar amount invested in those funds. The coefficients on each of these regressors are positive and significant at critical levels no higher than 0.02, demonstrating that differences in loser and winner risk adjustment behavior did become more meaningful over time as the size and scope of the mutual fund industry increased.

Table IV

Comparative Frequency Distributions for New versus Entrenched Funds and for Small versus Large Funds

Listed below are the cell frequencies for a 2 × 2 classification scheme involving the rank-ordered variables: (i) the Risk Adjustment Ratio (*RAR*); and (ii) the total return (*RTN*) through the first seven months (i.e., $M = 7$) of the year. We divide the funds into four groups on a yearly basis according to whether *RTN* is below ("loser") or above ("winner") the median and whether *RAR* is above ("high") or below ("low") the median. Entrenched funds are those that were in existence for the whole sample period while "new" funds enter during the period. We define "small" ("large") funds as having a net asset value below (above) the median for a particular tournament year. Results are listed for both the entire 1980–1991 period and the two six-year subsamples. *RAR* values are calculated and ranked excluding December observations for each sample year.

		Sample Frequency (% of Observations)					
		Low *RTN* ("Losers")		High *RTN* ("Winners")			
Period	Sample (No. of Obs.)	"Low" *RAR*	"High" *RAR*	"Low" *RAR*	"High" *RAR*	χ^2	*p*-value[b]
	Panel A: New vs. Entrenched Funds						
1980–91	Entrenched (n = 1341)	23.56	25.80	26.92	23.71	3.96	0.047
	New (n = 1143)	20.65	29.92	28.52	20.91	32.53	0.000
1980–85	Entrenched (n = 669)	26.31	24.22	22.57	26.91	2.79	0.095
	New (n = 182)	25.27	22.53	28.02	24.18	0.01	0.913
1986–91	Entrenched (n = 672)	20.83	27.38	31.25	20.54	19.74	0.000
	New (n = 961)	19.77	31.32	28.62	20.29	37.75	0.000
	Panel B: Small vs. Large Funds						
1980–91	Large (n = 1382)	20.26	24.82	29.59	25.33	10.95	0.001
	Small (n = 1102)	24.68	31.31	25.23	18.78	19.03	0.000
1980–85	Large (n = 482)	24.69	22.82	23.24	29.25	2.85	0.091
	Small (n = 369)	27.91	25.20	24.39	22.49	0.01	0.919
1986–91	Large (n = 900)	17.89	25.89	33.00	23.22	28.19	0.000
	Small (n = 733)	23.06	34.38	25.65	16.92	29.02	0.000

[a] The χ^2 statistic is calculated based on a null hypothesis that each cell should receive an equal distribution (i.e., 25 percent) of the sample.

z-statistic of 2.86, which is significant at the 0.004 level.[14] Beyond this, for both new and entrenched funds it is again the case that the relationship between interim performance and volatility ratios does not become statistically significant until the most recent subperiod. On balance, though, despite this evidence that newer and well-established funds respond slightly differently to being classified as an interim winner or loser, the overriding result is that both groups react in a manner that conforms to the premise that the mutual fund industry is an economic tournament.

[14] In conducting this test under the specifications of our central prediction, we combine the loser/high *RAR* and winner/low *RAR* cell frequencies for both new and entrenched fund samples and treat these pooled values as "successful" trials. These success frequencies are then differenced and subjected to the standard statistical procedure for proportional data.

The cell frequencies for the small versus large fund comparisons are shown in Panel B of Table IV. Although these findings are of a different nature than those in Panel A—as evidenced by an overall "difference in proportions" z-statistic of only 1.06—they nevertheless support the notion that losers adjust portfolio risk more than winners and that it is the smaller funds that tend to alter portfolio risk by the largest amount. An interesting facet of these results that makes drawing this last conclusion directly somewhat difficult is that the larger funds tend to be winners more frequently than do their smaller counterparts; for the entire sample period, the proportion of big fund to small fund winners is roughly 55 to 45 percent. However, once the cell frequencies are "normalized" to the size of the respective winner and loser populations, the small funds are indeed the group that produces the largest adjustments.[15] Consistent with all of the temporal dynamics reported so far, these tendencies became quite a bit stronger with the growth in the industry and increased investor awareness.

Finally, as an additional investigation, we also test whether the sales fee structure adopted by a fund influenced a manager's decision to adjust portfolio risk after the interim ranking date. Because no-load funds generally rely on the print media (i.e., advertisements) rather than brokers to sell their products, they would presumably be the group most heavily affected by published performance rankings. To test this notion, we employ load-fee information from *Wiesenberger's* for the 135 funds in our sample that are covered in that source and assign each to one of two subsets according to whether it had a front-end sales charge. For both the load and no-load subsamples we then generate separate 2×2 contingency tables over the entire sample period. These results (which are not reported in detail here) show that there was indeed a statistically significant difference in the tendency for no-load winners and losers to increase portfolio risk in the second part of the year. In contrast, we do not find a significant difference in the cell frequencies for the load fund subgroup.

Before concluding that there is a tractable relationship between a fund's load structure and the post-assessment risk adjustments of the manager, however, an important potential conflict must be considered. Specifically, load fee status may be correlated with the other fund characteristics discussed above (i.e., new versus entrenched, small versus large). We examine this possibility in two steps. First, we calculate separate rank correlation statistics between the fund's sales fee arrangement and each of the characteristic variables. These results indicate that while no-load funds show a significant tendency to be new (e.g., 71.7 percent of the no-load observations were entrenched compared to 82.6 percent for the load sample, with a correlation p-value of 0.001), there is

[15] To see this, consider the adjustments made by small and large funds that were winners in the 1980–1991 time frame. In this category, the percentage of large funds that have a low risk adjustment ratio is 54 percent (= 29.59 ÷ (29.59 + 25.33)) while the comparable statistic for small funds is 57 percent (= 25.23 ÷ (25.23 + 18.78)). In the loser category, the analogous adjustment percentages for the high risk adjustment ratio are 55 and 56 percent, respectively.

no apparent connection between a fund's size and its commission structure given the p-value of 0.414. Second, we recalculate the load and no-load cell frequencies after further segmenting the observations into new and entrenched funds. This analysis produces four separate 2×2 contingency tables from which only the new fund samples display the predicted risk-revision behavior for both no-load and load funds (i.e., chi-square p-values of 0.046 and 0.030, respectively); the load subsample generates significant results only in the entrenched group. Thus, we conclude that the load structure results cannot be distinguished from the fund's age dimension.

IV. Empirical Results: Robustness Tests

A. *The Influence of Cumulative Performance*

To this point in the analysis, we have implicitly assumed that investors make their investment decisions solely on the basis of annual fund performance rankings. There is evidence, however, that this is too simple a view. Sirri and Tufano (1992), for instance, find that both current- and past-year performance measures were important in explaining new fund inflows. By extension, then, it is also possible that a manager's record of past performance might influence his or her decision to alter portfolio risk at the interim date of a current tournament. In particular, we would predict that the more consistently the manager has been a loser (winner) in the past the more (less) likely it is that he or she will have an above average volatility ratio.

To examine this prediction, we first define the following measure of cumulative return performance for the j-th fund who finds itself at the interim assessment date in year y:

$$\text{CUM}RTN_{jy} = \prod_{t=y-1}^{y-n} \left[\prod_{m=1}^{12} (1 + r_{jmt}) \right]. \tag{4}$$

Here the subscript t represents the years prior to the current tournament that are being used in the ranking. Consistent with industry practice, which often considers a manager's three- and five-year relative performance, we calculate (4) for both two years and four years (i.e., $n = 2$ and 4) before the current period. These statistics are then used, along with the current year's partial return calculated by (2), as independent variables in a series of logistic regressions designed to predict the probability that a fund will adjust its volatility ratio following an interim ranking date. Accordingly, we specify a dependent binary variable having a value of 1 if fund j's volatility ratio is below the median level (i.e., low RAR), and 0 otherwise. These results are summarized in Table V.

This display lists two sets of findings, depending on the exact definition of the independent variables. In Panel A, these variables are used in their ranked form; that is, the actual return values are translated into binary ranking variables assuming the value of 1 if the particular return measure was above

Table V

The Relationship Between Risk Adjustments and Past Performance

Listed below are the estimated coefficients for various forms of the following logistic regression: *DVOLRAT* = f(Interim Annual Return, 2-year Cumulative Return, 4-year Cumulative Return), where *DVOLRAT* is defined as a binary variable assuming the value of 1 if a fund's volatility ratio is below the median value in an annual tournament, 0 otherwise. We use two sets of independent variable definitions. In Panel A, they are defined by assigning a value of 1 to any fund whose return is above the median in the interim annual, 2-year cumulative, or 4-year cumulative assessment periods (i.e., "winners"), 0 otherwise. Panel B then reports results using actual values for the independent variables.[a]

No. of Obs.	Intercept	Interim Perform. Measure	Cum. 2-yr Perform. Measure	Cum. 4-yr Perform. Measure	χ^2
		Panel A: Binary Variable (i.e., Ranked) Values			
2484	−0.220 (0.000)	0.430 (0.000)			28.49 (0.000)
2079	−0.336 (0.000)	0.365 (0.000)	0.225 (0.011)		24.40 (0.000)
1752	−0.391 (0.000)	0.390 (0.000)		0.367 (0.000)	32.27 (0.000)
		Panel B: Actual Variable Values			
2484	−0.005 (0.909)	0.200 (0.000)			23.37 (0.000)
2079	−0.039 (0.378)	0.149 (0.001)	0.118 (0.008)		18.73 (0.000)
1752	−0.009 (0.846)	0.163 (0.002)		0.136 (0.006)	20.01 (0.000)

[a] *P*-values for the Wald chi-squared statistics associated with each regression coefficient as well as for the covariates are listed parenthetically beside each statistic.

its median (i.e., "winners"), 0 otherwise. Panel B then uses each of these returns in their original form. The results indicate that whether the return data were measured in raw or ranked form, both cumulative and current year performance significantly influence the probability that a fund will increase its volatility ratio subsequent to the mid-year performance assessment. It is also interesting to note that irrespective of whether it is measured over the prior two or four years, cumulative performance has almost as large an impact on the likelihood that the risk ratio is in the largest group as does the interim return in the current tournament.[16]

B. How Managers Alter Portfolio Risk

The preceding results indicate quite strongly that funds generating the lowest returns at a mid-year ranking alter portfolio risk to a greater degree

[16] To help interpret Table V, notice that an estimate of the probability that a fund will increase its volatility ratio can be recovered from the logit procedure with the function $p = (1 + e^{-z})^{-1}$, where $z = X\phi$, with X representing the independent variables and ϕ the estimated vector of parameters. For example, using the first set of results listed for the binary ranking variables in Panel A, the probability of an interim loser increasing risk is 0.55 while the same estimate for an interim winner is 0.45. Thus, these logistic regression results confirm the findings from the previous sections supporting the prediction in equation (1).

than do those funds whose returns are the greatest. An interesting issue to consider as well is the way in which managers affected these changes. Generally speaking, there are two possibilities. The first, and the one suggested by our tournament story, is that managers actively revise the composition of their portfolios to achieve an explicitly riskier profile. This could entail either turning over some of their existing security holdings or, if allowed, altering the entire fund synthetically with derivative positions. On the other hand, the second type of decision is more reactive in nature; what if an interim loser continues to hold the same portfolio after the assessment date but the securities in that portfolio are now riskier than they were when the fund was first formed?[17] Under this scenario, the increase in risk occurs at the asset level and the only decision the manager makes is to maintain what he or she presumably recognizes to be a more volatile fund.

To determine which of these two strategies the sample of fund managers tend to employ, we create a collection of simulated control portfolios and replicate the previous experimental design. Our goal in this simulation is to establish the behavior of a group of unmanaged stock portfolios to contrast with our sample of actual funds, where any difference between them would indicate the degree of active mid-year risk manipulation. We conduct the experiment with two control fund samples. First, as the average number of stocks contained in our actual fund collection was 70.86, we attempt to match this by forming simulated portfolios of 75 stocks. We then contrast these results with a separate set of simulated portfolios containing 150 stocks apiece.

For both control groups, we randomly select securities from the CRSP data base and calculate equally-weighted averages of returns on a monthly basis from the beginning of 1980 until the end of 1991. No additions or deletions are made to these funds during the sample period; if a company represented in a given portfolio stops listing returns for any reason, it is not replaced. We then repeat this selection process 250 times for each portfolio size, which generates the samples we use in the analysis. As before, we classify all of the respective 250 control funds as interim winners and losers on an annual basis based on their cumulative returns through July of a particular year. A subsequent calculation of risk adjustment ratios allows us once again to summarize behavior in the simulated samples with sets of 2×2 contingency tables.

These findings are shown in Panels A and B of Table VI for the 75- and 150-stock funds, respectively. For comparative ease, under each of the control portfolio cell frequencies in Table VI, we also list the difference between this value and its corresponding counterpart in the sample of actual funds (the difference is calculated as actual cell frequency less control cell frequency). Whether viewed over the whole twelve-year period or during the more recent

[17] Hamada (1972) shows under stylized conditions, that a change in a firm's capital structure directly impacts the systematic risk of its stock. Thus, when an individual stock declines in market value (which, by definition, is more likely to happen for securities in interim loser portfolios than in interim winner funds), the firm's debt-equity ratio will change in a way that increases its risk level. This, in turn, can lead to a riskier fund if such changes are both pervasive and not diversified across the entire portfolio.

Table VI

Risk Adjustment Ratio Classifications for an Unmanaged Control Portfolio Sample

Listed below are the cell frequencies for a 2 × 2 classification scheme involving the rank-ordered variables: (i) the Risk Adjustment Ratio (*RAR*); and (ii) the total return (*RTN*) through the first seven months (i.e., $M = 7$) of the year. We construct and normalize the data for the classifications on a yearly basis from monthly returns to a set of 250 simulated portfolios. Each control portfolio contains stocks randomly selected during the period 1980 to 1991 and the composition of a fund is not altered once formed. We divide funds into four groups on a yearly basis according to whether *RTN* is below ("low" or "loser") or above ("high" or "winner") the median and whether *RAR* is above ("high") or below ("low") the median. Results are reported for three sample partitions: (i) the whole sample, (ii) the last six-year period, and (iii) the last two three-year periods. Differences in cell values relative to the actual fund sample are also reported.

		Control Sample Frequency (% of Observations)					
		Low *RTN* ("Losers")		High *RTN* ("Winners")			
Sample Period	Observations	"Low" *RAR*	"High" *RAR*	"Low" *RAR*	"High" *RAR*	χ^2	*p*-value[a]
		Panel A: 75-Stock Control Portfolios					
1980–1991	3000	23.97	26.00	26.00	24.03	13.72	0.000
	Difference[b]	−1.75	1.70	1.66	−1.61		
1986–1991	1500	25.00	25.00	25.00	25.00	54.98	0.000
	Difference	−4.79	4.70	4.70	−4.61		
1986–1988	750	27.60	22.40	22.40	27.60	35.71	0.000
	Difference	−5.49	5.42	5.42	−5.35		
1989–1991	750	22.40	27.60	27.60	22.40	16.76	0.000
	Difference	−3.62	3.62	3.62	−3.62		
		Panel B: 150-Stock Control Portfolios					
1980–1991	3000	23.73	26.17	26.17	23.93	11.12	0.001
	Difference	−1.51	1.53	1.49	−1.51		
1986–1991	1500	25.07	24.87	24.87	25.20	58.12	0.000
	Difference	−4.86	4.83	4.83	−4.81		
1986–1988	750	25.87	24.13	24.00	26.00	17.13	0.000
	Difference	−3.76	3.69	3.82	−3.75		
1989–1991	750	24.27	25.60	25.73	24.40	39.46	0.000
	Difference	−5.49	5.62	5.49	−5.62		

[a] The χ^2 statistic is calculated based on a null hypothesis that each cell should receive a distribution equal to that of the actual sample (cf. Table III).

[b] Indicates the difference in respective cell frequencies between actual and control fund samples, calculated as the former minus the latter.

subperiods, the results show rather convincingly that our overall results do not hold for a group of randomly generated unmanaged asset portfolios. Perhaps the best indication of this is that in both simulations the collective control frequencies differ from actual fund frequencies in the predicted way (i.e., a negative difference between actual and control for the loser/low *RAR* and winner/high *RAR* cells). As importantly, the reported chi-square statistics

indicate that these control frequencies are reliably different from the null hypothesis that each cell contains the same proportion of the distribution as the real sample. Interestingly, a comparison of the two panels suggests only slight differences associated with the number of stocks contained in the control funds. Thus, we conclude that the mid-year risk adjustments we observe were not passive in nature.[18]

C. Risk-Change Behavior in Volatility Subtournaments

Another implicit assumption in our initial analysis is that every fund listing "growth" as its investment objective would be included by investors in the same performance tournament. While this presumption is consistent with the way performance rankings are actually reported in the financial press (e.g., *Barron's*), it may be the case that investors consider additional fund investment characteristics beyond the stated objective when assessing relative performance. Indeed, Brown and Goetzmann (1995) show that traditional objective classification categories, such as "growth" and "income," are not necessarily accurate indicators of a mutual fund's true investment style or future performance. Thus if, for instance, managers of "low volatility" and "high volatility" growth funds are viewed by investors as being fundamentally different entities, commingling their performance in the same tournament could produce spurious overall results.

To address this issue, as of the beginning of every year during the last half of the sample period (i.e., 1986–1991), we calculate two different measures of volatility for each fund using its returns over the prior two years: (i) total risk, as indicated by return standard deviation; and (ii) systematic risk, calculated as an individual fund's beta coefficient relative to equally-weighted returns for the entire fund universe. The decision to use the previous 24 months of return data is based on the fact that this would be the information set available to investors at the start of a new competition. After calculating separate rank orderings of the funds by each of the volatility measures, we next divide the ranked samples into three non-overlapping groups, representing "high," "middle," and "low" volatility subgroups. Treating these volatility subgroups as different tournaments, we then reproduce the relevant *RAR/RTN* performance

[18] As both winners and losers, on average, altered their post-interim assessment risk levels by different amounts, it is relevant to consider whether the interim losers were successful in changing their ultimate tournament standing or whether the strategic reactions of the interim winners was sufficient to maintain their positions. To test this directly, we first define the tournament-end total return to the j-th portfolio in year y as $TRTN_{jy} = [((1 + r_{j1y})(1 + r_{jM+2y}) \cdots (1 + r_{j12y})) - 1]$ and then specify an interim-to-final "rank change" variable as $[\text{Rank}(TRTN_{jy}) - \text{Rank}(RTN_{jy})]$. To test the hypothesis that those funds with the largest risk adjustments were able to increase their relative ranking, we estimate a logistic regression of the form (Rank Change) = $f(RAR)$. That is, we investigate the direct link between increases or decreases in the rank change variable and the risk adjustment ratio for each of the fund managers, regardless of their initial winner/loser classification. We find the estimated parameter on *RAR* to be positive and significant beyond the 0.001 critical level. This is consistent with the notion that it was the interim losers who tended to increase their status, even if not necessarily at the expense of the interim winners.

Table VII

Comparative Frequency Distributions for Volatility-Ranked Subtournaments, 1986–1991

Listed below are the cell frequencies for a 2 × 2 classification scheme involving the rank-ordered variables: (i) the Risk Adjustment Ratio (*RAR*); and (ii) the total return (*RTN*) through the first seven months (i.e., $M = 7$) of the year. We initially divide funds into four groups on a yearly basis according to whether *RTN* is below ("loser") or above ("winner") the median and whether *RAR* is above ("high") or below ("low") the median. We then further sort the funds into one of three non-overlapping "volatility" subtournaments at the beginning of each year according to the relative level of their return volatility from the previous two years. We use two different measures of volatility: (i) total risk, measured by return standard deviation; and (ii) systematic risk, measured by an individual fund's beta coefficient relative to an equally weighted average of the entire sample. Results are listed for the six-year period 1986–1991. *RAR* values are calculated and ranked excluding December observations for each sample year.

		Sample Frequency (% of Observations)					
		Low *RTN* ("Losers")		High *RTN* ("Winners")			
Volatility Subgroup	No. of Obs.	"Low" *RAR*	"High" *RAR*	"Low" *RAR*	"High" *RAR*	χ^2	*p*-value[b]
Panel A: Total Risk							
High	492	20.93	28.86	28.86	21.34	11.74	0.001
Middle	492	21.54	28.25	28.25	21.95	8.33	0.004
Low	489	20.25	29.45	29.45	20.86	15.49	0.000
Panel B: Systematic Risk							
High	492	20.93	28.86	28.86	21.34	11.74	0.001
Middle	492	20.73	29.07	29.07	21.14	13.01	0.000
Low	489	22.09	27.61	27.61	22.70	5.33	0.021

[a] The χ^2 statistic is calculated based on a null hypothesis that each cell should receive an equal distribution (i.e., 25 percent) of the sample.

cell frequencies. Table VII details these findings for both the total risk (Panel A) and systematic risk (Panel B) volatility measures.[19]

These results show that, regardless of volatility definition or subtournament stratification, differences in the reported cell frequencies are all in the hypothesized direction and statistically significant at conventional levels. That is, even after allowing for the possibility that investors predicate their fund performance assessments (and, in turn, their subsequent investment decisions) on ex ante volatility measures, it remains the case that interim losers increase their funds' post-assessment risk levels to a greater extent than do interim winners. Further, given that this conclusion holds for both the total and systematic risk measures, we can infer that investors are not basing their

[19] As reported in Panel B of Table III, there are 1,633 annual fund observations available during the six-year period from 1986 and 1991. The results in Table VII, however, are based on only 1,473 observations, a reduction attributable to the additional requirement that each fund have two years of return history prior to 1986.

decisions on the diversifiable component of fund risk in a way that would make our overall conclusions suspect. In fact, both risk statistics produce identical "high" volatility rankings and hence identical cell frequencies. Consequently, our earlier conclusions remain valid and can be interpreted as if investors consider growth funds with disparate absolute risk levels to be part of the same tournament.

These new results also help dispel the prospect that differences between winner and loser fund *RAR*s are merely statistical artifacts of the way funds respond to changes in aggregate market volatility. That is, given the wide range of movement in the market volatility ratios during the sample period, it is possible that the two types of funds had greatly different systematic risks which, in turn, led to natural differences in their postassessment reactions that had nothing to do with managerial incentives. For example, if interim loser funds tended to have the lowest betas during the 1989–1991 subperiod when the average *RAR* was less than one, we would expect their volatilities to decline less when aggregate volatility declined. The findings in Panel B of Table VII, however, make this explanation unlikely. Even after separating all of the funds into three nonoverlapping beta subgroups before identifying them as winners and losers, it is still the case that the prediction in equation (1) is confirmed. Thus, we conclude that differences in winner and loser *RAR*s are not being driven by differential responses to aggregate volatility shocks.

V. Conclusions and Implications

Although a great deal has been written about the moral hazard problems associated with incentive fee structures for money managers, little has been documented as to the actual behaviors these schemes evoke. In this paper we argue that the mutual fund industry is appropriately viewed as a tournament in which the funds compete with one another for new assets (and, hence, higher compensation as a percentage of those assets) based on their relative performance. We propose that managers in such an environment would have different incentives to alter the volatility of their portfolios after an interim performance assessment depending on whether they were ranked as winners or losers.

Our central prediction is that interim losers would increase the risk of their funds more than would interim winners over the balance of the annual tournament period. Using the monthly returns from more than 330 growth-oriented mutual funds over the 1980–1991 period, we present evidence consistent with this prediction in that, at the margin, losers (winners) did indeed shift their investments so as to increase risk by a greater (lesser) degree. This effect is more pronounced during the last six years of the sample period as investor awareness of relative performance and the number of new funds in the industry escalated. Analysis of a simulated set of unmanaged stock portfolios suggest that these risk changes are due to explicit managerial actions and not generated entirely at the asset level. Furthermore, we demonstrate that our results are not an artifact of the ex ante level of fund volatility.

Perhaps the most important implication of this research is that it is possible that the current tournament structure of the mutual fund industry truly does provide adverse incentives to fund managers. That is, by focusing so much attention on relative return performance that is assessed annually, the industry may be effectively changing managerial objectives from a long-term to a short-term perspective. Of course, this is exactly the point that Kritzman (1987) raises. Given our empirical findings, we cannot argue against the notion that the tournament system provides incentives for managers to alter the essential nature of their portfolios in an effort to salvage or secure their yearly performance. To the contrary, we show that the managers with the poorest initial performance during an assessment period appear to act in a manner consistent with their own best interest, but not necessarily that of their shareholders. In the context of McDonald's (1974) results, this may well be equivalent to reclassifying the portfolio from, say, a growth fund to an aggressive growth fund. This presumably is not what the funds' investors want inasmuch as they can make this change more cheaply on their own.

REFERENCES

Baumol, W., S. Goldfeld, L. Gordon, and M. Koehn, 1990, *The Economics of Mutual Fund Markets: Competition Versus Regulation* (Kluwer Academic Publishers, Boston).

Bronars, S., 1987, Risk taking behavior in tournaments, Working paper, University of California at Santa Barbara.

Brown, S., W. Goetzmann, R. Ibbotson, and S. Ross, 1992, Survivorship bias in performance studies, *Review of Financial Studies* 5, 553–580.

Brown, S., and W. Goetzmann, 1993, Attrition and mutual fund performance, Working paper, Columbia University.

Brown, S., and W. Goetzmann, 1995, Mutual fund styles, Working paper, Yale University.

Bull, C., A. Schotter, and K. Weigelt, 1987, Tournaments and piece rates: An experimental study, *Journal of Political Economy* 95, 1–33.

Capon, N., G. Fitzsimons, and R. Prince, 1992, An individual level analysis of the mutual fund investment decision, Working paper, Columbia University.

Cohen, S., and L. Starks, 1988, Estimation risk and incentive contracts for portfolio managers, *Management Science* 34, 1067–1080.

Ehrenberg, R., and M. Bognanno, 1990, Do tournaments have incentive effects? *Journal of Political Economy* 98, 1307–1324.

Goetzmann, W., B. Greenwald, and G. Huberman, 1992, Market response to mutual fund performance, Working paper, Columbia University.

Goetzmann, W., and N. Peles, 1994, Cognitive dissonance and mutual fund investors, Working paper, Columbia University.

Golec, J., 1992, Empirical tests of a principal-agent model of the investor-investment advisor relationship, *Journal of Financial and Quantitative Analysis* 27, 81–96.

Green, J., and N. Stokey, 1983, A comparison of tournaments and contracts, *Journal of Political Economy* 91, 349–364.

Grinblatt, M., and S. Titman, 1987, How clients can win the gaming game, *Journal of Portfolio Management* 13, 14–20.

Grinblatt, M., and S. Titman, 1989a, Adverse risk incentives and the design of performance-based contracts, *Management Science* 35, 807–822.

Grinblatt, M., and S. Titman, 1989b, Portfolio performance evaluation: Old issues and new insights, *Review of Financial Studies* 2, 393–421.

Grinold, R., and A. Rudd, 1987, Incentive fees: Who wins? Who loses? *Financial Analysts Journal* 43, 27–38.

Hamada, R., 1972, The effect of the firm's capital structure on the systematic risk of common stocks, *Journal of Finance* 27, 435–452.

Kritzman, M., 1987, Incentive fees: Some problems and some solutions, *Financial Analysts Journal* 43, 21–26.

Lazear, E., and S. Rosen, 1981, Rank order tournaments as optimum labor contracts, *Journal of Political Economy* 89, 841–864.

Lakonishok, J., A. Shleifer, R. Thaler, and R. Vishny, 1991, Window dressing by pension fund managers, *American Economic Review* 81, 227–232.

McDonald, J., 1974, Objectives and performance of mutual funds, 1960–1969, *Journal of Financial and Quantitative Analysis* 9, 311–333.

McLaughlin, K., 1988, Aspects of tournament models: A survey, *Research in Labor Economics* 9, 225–256.

Nalebuff, B., and J. Stiglitz, 1983, Prizes and incentives: Towards a general theory of compensation and competition, *Bell Journal of Economics* 14, 21–43.

Radcliffe, R., 1990, *Investment: Concepts, analysis, strategy, 3e* (Scott-Foresman, Homewood).

Ritter, J., 1989, Portfolio rebalancing and the turn-of-the-year effect, *Journal of Finance* 44, 149–167.

Rosen, S., 1986, Prizes and incentives in elimination tournaments, *American Economic Review* 76, 701–715.

Sirri, E., and P. Tufano, 1992, The demand for mutual fund services by individual investors, Working paper, Harvard University.

Starks, L., 1987, Performance incentive fees: An agency theoretic approach, *Journal of Financial and Quantitative Analysis* 22, 17–32.