Risk Taking by Mutual Funds as a Response to Incentives

Judith Chevalier

University of Chicago and National Bureau of Economic Research

Glenn Ellison

Massachusetts Institute of Technology and National Bureau of Economic Research

This paper examines a potential agency conflict between mutual fund investors and mutual fund companies. Investors would like the fund company to use its judgment to maximize risk-adjusted fund returns. A fund company, however, in its desire to maximize its value as a concern, has an incentive to take actions that increase the inflow of investments. We use a semiparametric model to estimate the shape of the flow-performance relationship for a sample of growth and growth and income funds observed over the 1982–92 period. The shape of the flow-performance relationship creates incentives for fund managers to increase or decrease the riskiness of the fund that are dependent on the fund's year-to-date return. We examine portfolio holdings of mutual funds in September and December and show that mutual funds do alter the riskiness of their portfolios at the end of the year in a manner consistent with these incentives.

We both wish to thank the National Science Foundation (SBR 9414141, SBR 94-96301, and 95-15076). Chevalier acknowledges research support from the Graduate School of Business of the University of Chicago, and Ellison acknowledges support from the Harvard Clark Fund and a Sloan Research Fellowship. Andrea Densham, Judy Lubin, Matthew Sevick, and Ramin Toloui provided excellent research assistance. We thank Yacine Ait-Sahalia, John Cochrane, Sara Fisher Ellison, Paul Joskow, Jack Porter, Rob Porter, and Andrei Shleifer; seminar participants at Boston College, Chicago, Harvard, Michigan, MIT, Northwestern, State University of New York at Albany, the U.S. Department of Justice, Yale, and the NBER meetings in Industrial Organization and Corporate Finance; two anonymous referees; and Lars Hansen for their comments.

[*Journal of Political Economy*, 1997, vol. 105, no. 6]
 0022-3808/97/0506-0003$02.50

I. Introduction

The potential conflict between mutual fund companies and the people who invest in them is a classic example of an agency problem. Consumers would like the fund in which they invest to use its judgment to maximize risk-adjusted expected returns. Mutual fund companies, however, are motivated by their own profits, and the information they possess and how they use it are not directly observable. As a result, if actions that maximize fund company profits differ from the actions that maximize risk-adjusted expected returns, we would expect some inefficiencies to arise.

The primary goal of this paper is to explore the risk-taking behavior of mutual funds in light of the agency relationship between mutual funds and their customers. Our analysis has two main components. The first of these is a detailed examination of the relationship between mutual fund performance and the inflow of investments. Given the structure of management fees in the industry, the flow-performance relationship can be thought of as an implicit incentive contract. Our analysis thus allows us to address such questions as whether funds that trail the market have an incentive to "gamble" and try to catch the market and whether funds that are ahead of the market have an incentive to index the market and "lock in" a winning year. Second, we use a new detailed data set to examine how mutual fund portfolios are in fact altered toward the end of the year. This provides us with descriptive evidence on risk changes and allows us to assess whether funds are reacting to the incentives we identify.

In its most general goals, our research is related to the large literature that studies the effects of performance evaluation schemes or career concerns on the behavior of workers. Theoretical models such as Holmstrom (1982), Scharfstein and Stein (1990), Zwiebel (1995), and Prendergast and Stole (1996) examine the distortions in managerial behavior that arise when (as we imagine here) the market is trying to learn the ability of a decision maker. Because direct observation of performance evaluation schemes is rare, there is only a small empirical literature on responses to incentive contracts. Among the papers that examine the effects of explicit incentive schemes in various contexts are Healy (1985), Bronars (1987), Asch (1990), Ehrenberg and Bognanno (1990*a*, 1990*b*), and Knoeber and Thurman (1994). Previous papers that (like this one) have investigated market demand as an implicit incentive scheme include Borenstein and Zimmerman (1988) and Berkowitz and Kotowitz (1993).

Several authors have previously documented a strong relationship

between the inflow of new investment into a mutual fund and the fund's past performance (see, e.g., Patel, Zeckhauser, and Hendricks 1990; Ippolito 1992; Sirri and Tufano 1993; Goetzmann and Peles, in press). Because mutual fund companies usually receive a fixed percentage of assets under management as compensation, they will have an incentive to take whatever actions increase the total assets of the fund. In effect, the flow-performance relationship then serves as an implicit incentive contract. On the most basic level the existence of the flow-performance relationship is a good thing: it provides funds with an incentive to perform well. What we attempt to do in the first part of this paper is to examine the flow-performance relationship in detail to better understand what types of incentives it might create for funds to alter the riskiness of their portfolios.

Our empirical approach in this part of the paper involves applying a semiparametric model to a data set containing flow and performance data for a sample of growth and growth and income funds containing 3,036 fund-years observed over the 1982–92 period. The model has been chosen to provide as detailed a view of the flow-performance relationship as is feasible, while putting few a priori restrictions on the shape of the relationship. We find significant nonlinearities in the relationship, with the overall sensitivity of the relationship and its shape being dependent on the age of the fund in question.[1]

While this work may improve understanding of consumer behavior in the mutual fund industry, we primarily focus on the incentives for fund risk taking the consumer behavior engenders. In particular, we use our analysis of the flow-performance relationship to derive estimates of how the market implicitly compensates funds for increasing or decreasing the riskiness of their portfolios toward the end of the year as a function of the funds' January–September year-to-date performance, age, and other characteristics. To see how such implicit incentives might be created, suppose that, because of limited information availability or some other reason, many consumers of mutual fund services react to year-end performance results, and given a particular fund's characteristics and year-to-date performance at the end of September, the fund company knows that its future inflows of investments will (over some range) be a convex function of its fourth-quarter performance. In such a case, the fund would be able to increase its expected growth by increasing the variance of its fourth-quarter return. Our estimates indicate that incen-

[1] Ippolito (1992), Sirri and Tufano (1993), and Goetzmann and Peles (in press) have previously noted that the relationship between flow and performance may be nonlinear.

tives to alter riskiness are stronger for young funds than for older funds. In line with popular wisdom, young funds appear to have an incentive late in the year to "gamble" and try to catch the market if they are a few points behind; they may also have an incentive to play it safe and act more like an index fund if they are ahead of the market. The strongest incentive we find, however, is an incentive of funds that are well ahead of the market to gamble (perhaps in an attempt to make year-end lists of "top performers").

The second main goal of the paper is to examine empirically how funds alter the riskiness of their portfolios toward the end of the year (and, in particular, whether they appear to respond to the incentives we identify in looking at the flow-performance relationship). In this task, our paper is related to a few other papers that have tested for distortions in behavior by investment managers, including Lakonishok, Shleifer, Thaler, and Vishny's (1991) study of "window dressing" by pension fund managers and Lakonishok, Shleifer, and Vishny's (1992) and Grinblatt, Titman, and Wermers's (1995) studies of herding. The two other papers we are aware of on risk taking are Brown, Harlow, and Starks (1996) and Roston (1996). Brown et al. look at whether mutual funds whose performance is behind the market in the first part of the year have more variable returns during the remainder of the year than mutual funds that were ahead of the market. They informally discuss why "tournaments" among mutual fund managers might produce this pattern of behavior. Roston examines whether the shape of the flow-performance relationship and the unsystematic risk level of a mutual fund change systematically as the fund ages. The contribution of our paper is to provide a more detailed view of risk taking and, most important, to explore the connection between risk taking and the incentives created by market demand.

Our analysis of actual risk-taking behavior exploits a newly constructed data set that contains the complete equity portfolios of mutual funds both at the end of September and at the end of December of a given year. We find that changes in the riskiness of funds' portfolios appear to be related to the incentives we have previously identified, with the pattern of actual changes corresponding to the estimated incentives in some detail. To verify that these changes appear to be reflected in measures of riskiness based on the complete portfolios of mutual funds as well, we also explore time-series data on fund returns.

The paper is organized as follows. In Section II, we describe the data prepared for the estimation. In Section III, we empirically estimate the flow-performance relationship. In Section IV, we estimate each fund's incentives to increase or decrease its level of risk. Finally,

in Section V, we examine the portfolio changes undertaken by firms between September and December in order to test whether these portfolio changes reflect the incentives to take risks estimated in Section IV. Section VI summarizes our conclusions.

II. Data

Virtually all the data used in this paper are obtained from Morningstar Incorporated. The primary data source is Morningstar's January 1994 *Mutual Funds OnDisc.* From the CD-ROM, we obtain data on mutual fund returns, assets under management, minimum initial purchase requirements, and expense ratios as well as information on whether the fund had ever been involved in a merger. While the data from *Mutual Funds OnDisc* include only mutual funds that were still in operation as of January 1994, Morningstar has maintained a list of funds deleted from the database since the beginning of 1989. Using this list, we reconstructed the returns and other information back to 1989 for funds that were not still in existence using the *Mutual Fund Sourcebook,* a Morningstar publication.

A large data set containing the complete equity portfolios of a large number of mutual funds was obtained directly from Morningstar. Portfolio reporting to Morningstar is voluntary, and we have portfolio data for only a minority of the fund-years in this database. In addition, the frequency with which portfolios are available to us varies with the fund: some portfolios are available quarterly or more frequently, some only at the end of the year, and some at sporadic intervals. Much of our analysis focuses on 839 cases (involving 398 different funds) in which we have the portfolios of a fund at both the end of September and the end of December of the same year.

In order to construct measures of the riskiness of each mutual fund at each point in time at which a portfolio was available, we matched the portfolio holdings to the database of the Center for Research in Security Prices (CRSP). For each fund-date for which portfolio data were available, the Morningstar database generally contained the name of the security, the number of shares of the security, and the value of the holding. Unfortunately, the database does not contain any security identifier numbers and frequently does not contain ticker symbols; the tickers are missing for 80,435 of the 121,895 security records in the 1,678 ($= 2 \times 839$) portfolios mentioned above. For each holding in the database, we attempted to generate (possibly more than one) potentially correct tickers by (*a*) trying to match the holding name electronically to the name of a security in the CRSP data, (*b*) trying to match the holding name

to the name of another holding in the Morningstar data for which Morningstar provided a ticker, and (*c*) trying to match manually the holding name to the complete listing of CRSP securities. Morningstar holdings were matched to CRSP security records, which matched both the tentatively assigned ticker and the price per share in the Morningstar data.[2] At the end of this process, we were able to find matches for 92.5 percent of the security records. A variety of things appear in the lists of unmatched securities: foreign securities, holdings of shares in other mutual funds, securities whose prices in the Morningstar data may be incorrect, and securities that may be in CRSP but for which we simply could not find the match. When we use the CRSP data on returns in the previous year to estimate betas and standard deviations of portfolios, additional securities must be dropped because they are new or otherwise lack sufficient historical return data. On average, we are able to use about 89 percent of the records for this purpose. In our analysis, we shall generally restrict ourselves to looking at funds for which estimates are based on matches to at least 85 percent of the September portfolio by value.

Clearly there are several potential areas for concern with the data. Because Morningstar did not keep records on funds that were dropped from its database prior to 1989, collecting data on such funds was decided to be prohibitively difficult, and our data set is survivorship-biased for the period prior to 1989. Another concern about the database is that the return and other information from *Mutual Funds OnDisc* may have a back-filling problem. At the time Morningstar began to provide information about a fund, Morningstar may have filled in back data for the fund. To see what effect this might have, the working paper version of this paper (Chevalier and Ellison 1995) reestimated some of the models on a smaller non-back-filled, non-survivorship-biased subsample (and found similar results). The voluntary reporting of portfolio data may create additional selection biases.

Despite the difficulties with the database that we have noted, the database we constructed from the Morningstar data is a unique resource for mutual fund portfolio data. The only other database of mutual fund holdings of which we are aware is the data on 274 mutual funds over the 1975–84 period used in Grinblatt and Titman (1989, 1992, 1993), Wermers (1993), and Grinblatt et al. (1995).

[2] Prices were required only to be within $1.00 per share. As a result, for 71 of the 41,603 ticker-date-price combinations, multiple CRSP matches were found and the match with the smaller price difference was selected.

III. The Flow-Performance Relationship

When consumers are faced with the decision of choosing a mutual fund, they must form beliefs about the suitability of each fund to their objectives and about the ability of each fund to generate excess returns. Previous research on the relationship between investment flows and past performance has demonstrated that consumers react strongly to historical returns. We are primarily interested in understanding what types of incentives this creates for funds to manipulate the riskiness of their portfolios, and how these incentives vary both over time as returns are realized and cross-sectionally with fund attributes. These incentives are affected both by the strength and by the convexity/concavity of the flow-performance relationship. Hence, we begin here with an attempt to analyze the flow-performance relationship in as much detail as the size of our data set allows.

A. *A Semiparametric Model of the Flow-Performance Relationship*

In this subsection we try to estimate the effects of past performance and other characteristics on the flow of investments into a fund. Our model primarily focuses on the effect of year t excess returns on investment flows in year $t + 1$. Specifically, we estimate both the coefficients γ_k, δ_k, and $\alpha_1, \ldots, \alpha_5$ and the shape of the function f in the model

$$\begin{aligned}\text{Flow}_{it+1} = &\sum_k \gamma_k \text{Age}\, k_{it} f(r_{it} - rm_t) + \sum_k \delta_k \text{Age}\, k_{it} \\ &+ \alpha_1(r_{it-1} - rm_{t-1}) + \alpha_2(r_{it-2} - rm_{t-2}) \\ &+ \alpha_3(r_{it+1} - rm_{t+1}) + \alpha_4 \text{IndustryGrowth}_{t+1} \\ &+ \alpha_5 \log(\text{Assets}_{it}) + \epsilon_{it+1}.\end{aligned}$$

The dependent variable, Flow_{it+1}, is the proportional growth in total assets under management for the fund between the start and end of year $t + 1$ *net* of internal growth in year t (under the assumption of reinvestment of dividends and distributions); that is,

$$\text{Flow}_{it+1} = \frac{\text{Assets}_{it+1} - \text{Assets}_{it}}{\text{Assets}_{it}} - r_{it+1}.$$

Our choice of the growth rate (as opposed to the absolute level of inflows or another measure) as the dependent variable reflects the results of preliminary regressions that seemed to indicate that at vari-

ous performance levels funds do tend to grow or shrink in proportion to their initial assets.[3]

The measure of performance in our model, $r_{it} - rm_t$, is the simple linear difference between a fund's return and the return on a value-weighted market index. We focus on the relationship between year t excess returns and year $t + 1$ flows, estimating the shape f of the flow-performance relationship and allowing the sensitivity of the relationship to vary with the age category to which a fund belongs. If consumers are attempting to infer the quality of a fund from historical data, we would expect flows for younger funds to be more sensitive to recent performance. We allow also for separate intercepts for each age category so that average growth rates may differ by age. We write Age k for the various dummy variables indicating whether a fund belongs to each of the seven age categories: 2, 3, 4, 5, 6–7, 8–10, and 11 or greater.

We also include as explanatory variables the excess return of the fund in years $t - 1$ and $t - 2$ (with this latter variable set to zero for 2-year-old funds). The market-adjusted return in year $t + 1$ is included to reflect both flows in response to intrayear returns and the fact that funds with high returns in year $t + 1$ exhibit additional growth due to the internal growth of investments that are made before the end of year $t + 1$. Additional variables included as controls are the growth in total assets under management by the equity mutual fund industry (taken from the 1994 *Mutual Fund Factbook*) and the natural logarithm of the ratio of total assets under management by the fund in question at the end of year t to the (geometric) mean of assets under management across all funds in the sample being analyzed.

B. Data

The data set on which we estimate our flow regressions contains information on flows into 449 growth and growth and income mutual funds during all years between 1983 and 1993 in which they were active. Our basic regression is run on a sample of 3,036 fund-years that consists of all observations meeting certain criteria. First, because we thought that the nature of their flows might be quite different, we removed funds that were closed to new investors, funds that are primarily institutional (which we defined as having a minimum

[3] A priori, one would expect withdrawals from a poorly performing fund to be proportional to its initial asset level. Growth might be proportional to assets for good performers as well if current investors are those most likely to add funds or if initial asset levels are a good proxy for the number of people who are aware of a fund or will hear positive news via word-of-mouth channels.

TABLE 1

SUMMARY STATISTICS FOR FLOW-PERFORMANCE REGRESSIONS

	FULL SAMPLE		YOUNG FUNDS		OLD FUNDS	
VARIABLE	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Flow_{t+1}	.122	.646	.251	.700	.079	.620
$r_t - rm_t$	−.005	.084	−.006	.085	−.005	.083
$r_{t-1} - rm_{t-1}$	−.002	.087	.0002	.091	−.003	.086
$r_{t-2} - rm_{t-2}$	−.001	.080	.001	.073	−.002	.083
$r_{t+1} - rm_{t+1}$	−.008	.077	−.006	.084	−.009	.075
$\text{IndustryGrowth}_{t+1}$	.283	.173	.272	.181	.286	.169
log(Assets)	4.936	1.427	4.254	1.173	5.172	1.431

	AGE CATEGORY						
	2	3	4	5	6–7	8–10	11+
Number of fund-years	221	199	189	171	238	168	1,850

initial purchase of at least $25,000), and funds that have very high expense ratios (in excess of 4 percent). We also eliminated funds that merged with other funds in year $t + 1$ (and hence have misleading growth rates) and two groups of funds for which the flow data are exceptionally noisy: funds less than 2 years of age at the end of year t and funds with less than $10 million in assets at the end of year t.[4] Table 1 contains summary statistics for all the variables.

C. *Estimation*

In our estimation we divide the data into "young" and "old" fund subsamples consisting of funds of ages 2–5 and funds of age 6 or more, respectively. The model is then estimated separately on each subsample. The shape of the flow-performance relationship may thus differ between young and old funds. Otherwise we have conserved degrees of freedom by assuming that the flow-performance relationship has the same basic shape for all age categories in a given subsample (albeit with age category–specific scale and shift effects). To identify the model, we omit the additive and multiplicative dummy variables for one age category in each subsample. If the

[4] When discussing the flow into a fund in year $t + 1$, we shall refer to a fund as being of age k if it is k years old at the start of year $t + 1$, i.e., if its inception date falls within year $t - k$. We exclude funds younger than 2 years of age and very small funds because we feel that initial fund size is probably a poor proxy for the potential to attract new customers, and thus our growth rate specification is probably inappropriate.

multiplicative coefficient γ_k is positive, then flow is more sensitive to performance for funds in the kth age category than for funds in the omitted age category. If the additive shift δ_k is positive, then the expected flow is greater for funds in the kth age category than for funds in the omitted category (at least at the value of r for which $f(r - rm) = 0$).

We estimate the model by a three-step process: first estimating the α coefficients, then the γ's and δ's, and finally the function f. For the first step, we note that, for each subsample of the data consisting of those observations in a single age category, we may obtain a consistent estimate of α by the nonparametric partialing-out procedure described by Robinson (1988): performing kernel regressions of both Flow_{it+1} and X_{it} on $r_{it} - rm_t$, and then regressing the residuals on the residuals. The estimate $\hat{\alpha}$ we use is obtained by computing separate estimates for each age category and taking a sample size weighted average of these estimates.

Next, to obtain estimates of the γ and δ parameters, we note that for each age category k we may consistently estimate the function $\gamma_k f + \delta_k$ from the subsample of the data in that age category using a kernel regression of $\text{Flow}_{it+1} - \hat{\alpha}X_{it}$ on $r_{it} - rm_t$. Writing $\hat{g}^k$ for this estimate (and $\hat{g}^0$ for the estimate obtained from the subsample corresponding to the age category whose dummy is omitted), we are able to obtain in a large number of ways consistent estimates of γ and δ using the functions $\{\hat{g}^k\}$. For example, one very simple (and likely very inefficient) method would be simply to set

$$\tilde{\gamma}_k = \frac{\hat{g}^k(x_1) - \hat{g}^k(x_0)}{\hat{g}^0(x_1) - \hat{g}^0(x_0)} \quad \text{and} \quad \tilde{\delta}_k = \hat{g}^k(x_0) - \hat{\gamma}_k \hat{g}^0(x_0)$$

for some pair of points x_0 and x_1. The estimates, $\hat{\gamma}_k$ and $\hat{\delta}_k$, we use are obtained by evaluating each $\hat{g}^k$ on a grid with support $[-0.17, 0.17]$ and then regressing the values of each $\hat{g}^k$ on the values of $\hat{g}^0$ using $r_{it} - rm_t$ as an instrument.

Finally, to obtain an estimate $\hat{f}$ of the function f we perform a kernel regression of

$$\hat{y}_t \equiv \frac{\text{Flow}_{it+1} - \sum_k \hat{\delta}_k \text{Age}\, k_{it} - \hat{\alpha}X}{1 + \sum_k \hat{\gamma}_k \text{Age}\, k_{it}}$$

on $r_{it} - rm_t$. In this and all other kernel regressions, we have used an Epanechnikov kernel with a window width that varies across the data so that more smoothing is done near the edges than in the middle of the excess return distribution. Specifically, in any re-

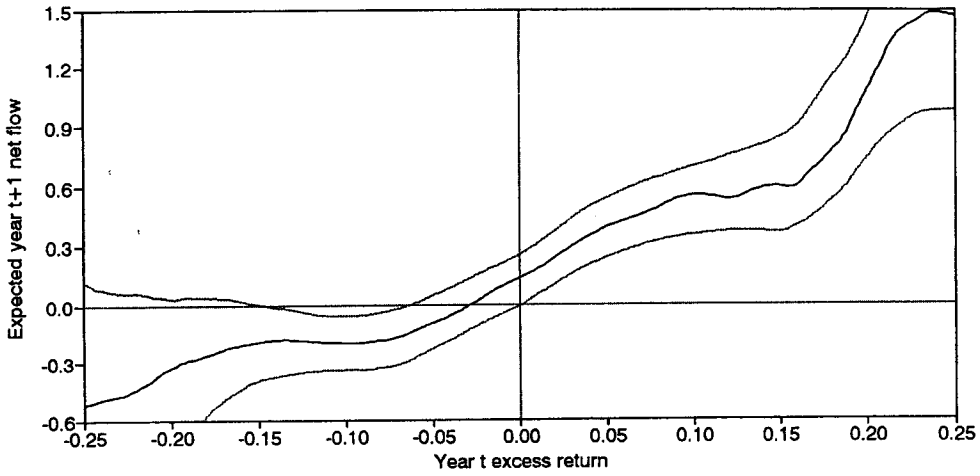

FIG. 1.—Flow-performance relationship $\hat{f}$ for young funds (age 2) with 90 percent confidence bands.

gression involving n data points, we use the window width $(0.3 + 0.1|r_{it} - rm_t|)(n/1{,}000)^{-1/5}$. Standard errors for all the parameter estimates were obtained by simulations that allowed for nonnormality of the error distribution and a degree of heteroskedasticity by sampling errors with replacement from the set of residuals corresponding to observations within the same age category.

D. Results

We begin with the estimated shapes for the flow-performance relationship. Figure 1 presents a graph of the function $\hat{f}$ obtained from the subset of young funds, along with pointwise 90 percent confidence bands. Given the normalizations we have made, the graph may be interpreted as presenting the expected growth rate in year $t + 1$ of a 2-year-old fund as a function of its year t excess return (under the assumptions also that the fund's excess return is zero in years $t - 2$, $t - 1$, and $t + 1$; that the fund's total assets match the geometric mean of our sample; and that the industry as a whole experiences zero growth). For example, such a fund would be expected to grow by approximately 15 percent in year $t + 1$ if its year t return matches the return on the market and to grow by approximately 55 percent if its return is 10 points greater than the market return.

While in this paper we do not attempt to test models of consumer behavior, the relationship in the figure appears to be roughly consistent with a model in which heterogeneously informed potential investors try to assess the quality of various funds. When the fund's

return is 15 or more points below the market, funds flow out quickly and the rate is sensitive to performance, as though increasingly even investors who pay little attention begin to take notice of the fund's poor performance.[5] For somewhat less disastrous results (say between 15 and eight points below the market), the curve appears to be largely flat, as though the fund attracts few or no new investors but does retain many of its old investors. At more typical performance levels, flow is increasing in year t excess returns. The marginal value of each unit of return appears perhaps to diminish once the excess return reaches 10 points and then increases sharply at excess return levels above 15 points. This pattern is consistent with a model in which very large returns bring sharply higher flows as a fund starts to make annual "best fund" lists and therefore comes to the attention of relatively uninformed potential investors.

For our purposes, though, why the flow-performance relationship has the shape it has does not really matter. We think of the shape as an empirical fact and shall be concerned in the next two sections with incentives to manipulate riskiness that derive from it.

When thinking about the statistical significance of the nonlinearities in the figure, one should keep in mind that both the intercept and the scale of the function f are identified only off differences between omitted age categories and the rest of the data. The confidence bands thus reflect more than just uncertainty about the shape of the function f. Subsection E presents a formal test of the significance of the departures from linearity.

The function $\hat{f}$ estimated from the sample of older funds and 90 percent confidence bands are presented in figure 2. Here, we have omitted the dummy for funds of age 11 and above, so that the graph can be interpreted as presenting the expected growth rate of a fund of this age group in year $t + 1$ as a function of its year t excess return. The estimated expected flows for these funds are clearly less sensitive to year t excess returns than those for 2-year-old funds, never falling below -15 percent or rising above 75 percent for a fund whose return is within 25 points of the market. The most striking feature of this graph is its generally convex shape, which suggests that incentives to carry unsystematic risk may be fairly universal. In contrast to the pattern for younger funds, we do not see outflows increase dramatically at the worst performance levels. Flows do, however, again increase sharply for the best-performing funds.

[5] The left side of the figure may be affected by selection bias. Funds that perform very poorly are more likely to "die" and have their flow omitted from our data set. Treating these deaths as substantial outflows would likely result in a picture that was more steeply sloped at low performance levels.

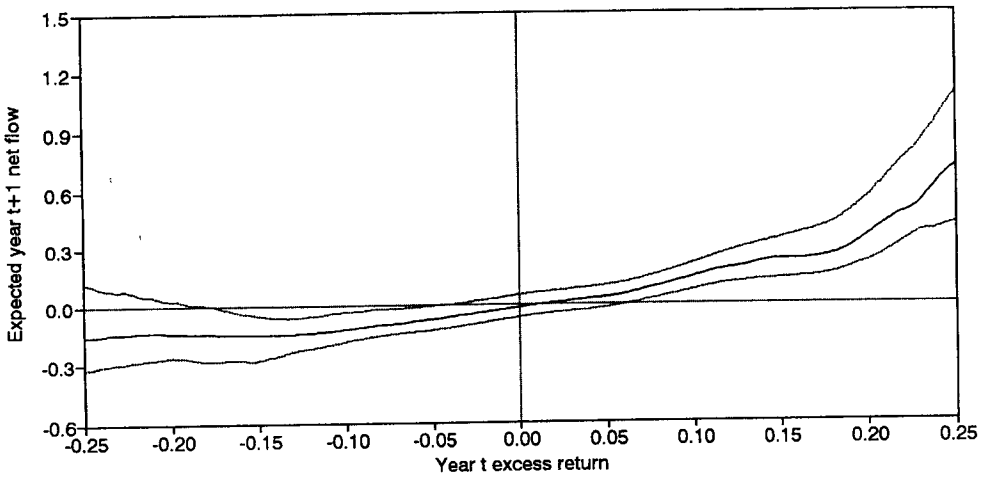

FIG. 2.—Flow-performance relationship $\hat{f}$ for old funds (age > 10) with 90 percent confidence bands.

We report estimates of the other parameters of the model obtained from the subsamples of young and old funds in columns 1 and 2, respectively, of table 2. To interpret the age category–specific scale and shift parameters, keep in mind that the omitted categories are 2-year-old and 11-year-old and older funds in the two subsamples. Hence, for example, to obtain a graph of the expected growth rate of a 4-year-old fund (having otherwise the standard characteristics), one would multiply the curve in figure 1 by a factor of .66 (= 1 − .34) and then shift it down by .02. A 4-year-old fund that matches the market return would thus be expected to grow by about 8 percent, with the expected growth increasing to about 36 percent if its return is 10 points above the market. In column 1 we see that the estimates of the multiplicative terms γ_3, γ_4, and γ_5 for funds of ages 3–5 are negative and monotonically decreasing. This indicates that the older funds' flows are increasingly less sensitive to their most recent performance. While these parameters are not very precisely estimated, the sensitivity of the 4-year-old and 5-year-old funds to year t returns is significantly smaller than that for 2-year-old funds at the 5 percent level in a one-tailed test. In the subsample of older funds, our point estimates are that flows into the 6–7 and 8–10-year-old funds are more sensitive to year t returns than flows into funds that are 11 years of age or more, although the differences fail to be significant. The additive effects are all small and insignificant.

Turning to the control variables in the lower part of the table, we see that year $t - 1$ and $t - 2$ excess returns also have substantial and statistically significant effects on flows in year $t + 1$. For example, the 1.86 and 0.73 coefficients on $r_{it-1} - rm_{t-1}$ and $r_{it-2} - rm_{t-2}$ in the

TABLE 2

COEFFICIENTS FROM SEMIPARAMETRIC FLOW-PERFORMANCE MODEL

	DEPENDENT VARIABLE: $Flow_{t+1}$	
INDEPENDENT VARIABLE	Young Funds ($N = 780$) (1)	Old Funds ($N = 2,256$) (2)
Multiplicative:		
Age 3 (γ_3)	−.28	
	(.25)	
Age 4 (γ_4)	−.34	
	(.20)	
Age 5 (γ_5)	−.46	
	(.22)	
Age 6–7 (γ_{67})		.30
		(.79)
Age 8–10 (γ_{810})		.35
		(.38)
Additive:		
Age 3 (δ_3)	.06	
	(.07)	
Age 4 (δ_4)	−.02	
	(.06)	
Age 5 (δ_5)	−.01	
	(.06)	
Age 6–7 (δ_{67})		.03
		(.09)
Age 8–10 (δ_{810})		.02
		(.04)
$\text{Excess return}_{t-1}$	1.86	1.00
	(.25)	(.14)
$\text{Excess return}_{t-2}$	.73	.29
	(.31)	(.16)
$\text{Excess return}_{t+1}$	2.49	2.44
	(.26)	(.16)
$\text{Industry Growth}_{t+1}$	1.07	.85
	(.12)	(.07)
$\log(\text{Assets}/\bar{A})$	−.07	−.04
	(.02)	(.01)
R^2	.38	.22

NOTE.—Estimated standard errors are in parentheses.

regression for young funds indicate that beating the market by 10 points in 1990 (relative to tying the market) would result, on average, in an additional (relative to what it would have been had the fund tied the market in 1990) inflow of 18.6 percent of a fund's start-of-1992 assets in 1992 and an additional inflow of 7.3 percent of its start-of-1993 assets in 1993. The effects are smaller for older funds, and in each year $t - 1$ returns have a larger effect than year

$t - 2$ returns.[6] The coefficient on year $t + 1$ returns is substantially greater than one, which suggests that some investors at least react to year-to-date performance in making their investment choices. Industry growth has a nearly one-for-one effect on flows, as would be expected, and larger funds appear to grow slightly more slowly.

E. Departures from Linearity

Because the incentives to take risks we shall identify derive from nonlinearities in the flow-performance relationship, we discuss here briefly the extent to which the departures from linearity that our model identifies are statistically significant.

For a formal test of this hypothesis, we applied a version of the specification test of Ellison and Ellison (1993). Specifically, for $\hat{\boldsymbol{\epsilon}}$ the vector of residuals from a linear model with age category–year t excess return interactions and $\mathbf{W}$ a matrix of kernel weights reflecting differences in year t excess returns, we formed the test statistic $T = \hat{\boldsymbol{\epsilon}}'\mathbf{W}\hat{\boldsymbol{\epsilon}}/(\sqrt{2}\hat{\sigma}^2\|\mathbf{W}\|)$, where $\|\mathbf{W}\| = (\sum_{ij} w_{ij}^2)^{1/2}$. Critical values for the test were obtained by reestimating the regression and recomputing the test statistic on 1,000 simulated data sets created by sampling with replacement from the empirical distribution of the residuals. In the sample of old funds, we were able to reject linearity at the 1 percent level. In the sample of young funds, it is rejected at the 10 percent level.

Ippolito (1992) and Sirri and Tufano (1993) have previously noted that the flow-performance relationship appears nonlinear and have applied particular two- and three-segment piecewise linear models, respectively. Our decision to apply semiparametric methods instead derives primarily from a desire to avoid imposing the strong restrictions on risk incentives. For example, with a two-segment piecewise linear model the estimated incentive to increase or decrease risk will always be maximized at the point at which the flow-performance relationship has a kink and will decline smoothly to zero at extreme performance levels.

When a specification test such as that described above is applied to piecewise linear specifications modeled after those of Ippolito and Sirri and Tufano, it again rejects the null of correct specification at

[6] When we tried additional lagged returns in the regression, they were not significant. The working paper version of this paper (Chevalier and Ellison 1995) discusses a number of other specification issues in the context of a linear regression. Among our observations are that size-return interactions do not appear to be significant and that the parametric parts of the model are largely unchanged by attempts to risk-adjust results or reduce survivorship bias.

the 1 percent level in the subsample of old funds and at the 10 percent level in the subsample of young funds. We take this to indicate that we do have enough data to potentially derive useful information from semiparametric methods. A more complicated piecewise linear model could, of course, have fit the data fairly well.

IV. Estimation of Risk Incentives

In this section we discuss how the flow-performance relationship viewed as an incentive scheme may induce mutual funds to manipulate the riskiness of their portfolios toward the end of the year. We take the basic agency problem between a mutual fund and its investors to be that, while investors would like the fund to use whatever private information it may have to maximize risk-adjusted returns, the mutual fund itself will instead take whatever action maximizes its value as a concern given the incentives it faces.[7] Because management fees in the industry are usually charged as a percentage of assets (within some size range), the value of a mutual fund (with future expected growth and the level of management fees held constant) is to a first approximation proportional to its assets under management. Because many potential investors see and react to year-end returns, a fund may at times increase its expected inflow of investment (and hence its value) by altering the riskiness of its portfolio.

Given that a mutual fund wishes to maximize its value as a concern, how might it have an incentive to distort its portfolio? For a simple illustration, consider the case of a 2-year-old fund that by the end of September of year t has fallen eight percentage points behind the market and finds itself confronted by the flow-performance relationship of figure 1. If the fund indexes the market for the remainder of the year, it will finish the year eight points below the market and see an expected outflow of 14 percent of its value in year $t + 1$. If in the final quarter the fund loses five more points and finishes the year 13 points behind the market, the expected outflow will be about the same. If instead the fund outperforms the market by five points in the final quarter, it will finish the year three points behind the market—a position sufficient to provide (in expectation) a small positive expected inflow in year $t + 1$. Clearly, the fund's expected size at the end of year $t + 1$ is far greater if it holds a portfolio that is equally likely to gain or lose five points than if it indexes the market; that is, the fund is tempted to gamble and try

[7] Note that in doing so we assume away any agency problems between mutual fund companies and their managers.

to catch up with the market. In contrast, suppose that the fund were instead eight points ahead of the market by the end of September. Inflows would now be only slightly higher with a small improvement in performance, but would be sharply lower with an inferior performance. Hence, the fund benefits from "locking in" its gains by indexing the market throughout the fourth quarter.

The incentives to alter the riskiness of a portfolio described above are derived from the fact that flows (or our estimates of them) are a nonlinear function of calendar year returns.[8] If instead flows were a separable function of the returns in each quarter, a fund's position relative to the market at the end of September would be irrelevant to its behavior in the final quarter. Why should we believe that there is something "magic" about calendar year returns? Our view is that the return in the most recent calendar year is important because calendar year data appear to be most generally available to consumers. Listings of mutual funds, accompanied by calendar year returns, are published on an annual basis in many news, business, and financial publications. The annual return format is also used in the annual comprehensive fund listings produced by Morningstar and others. While higher-frequency return data are available from a number of sources, it is necessary for our argument only that a nonnegligible fraction of consumers rely on annual data. Some empirical support for the importance of calendar year returns in the evaluation of asset managers is provided by Lakonishok et al. (1991), who show that window dressing by institutional asset managers is more likely to occur at the end of the fourth quarter of the year than it is to occur at the end of any of the other three quarters.

How can we quantify the incentive of a mutual fund to increase or decrease the amount of risk it holds toward the end of the year? For the simplest formulation, suppose that the management fees a fund collects and its costs of operation are each directly proportional to its assets under management. Suppose also that inflows and outflows of funds occur only at the end of each year and that returns in year t affect the flow of funds only at the end of that year (and not in any subsequent years). In this case, the value of a mutual fund as a company would be directly proportional to its assets under management, and the benefits to a fund from changing the riskiness of its portfolio are directly proportional to the change in expected flow the change produces.[9]

[8] That such incentives exist when agents can alter their behavior in continuous time is the core of Holmstrom and Milgrom's (1987) argument that linear incentive schemes may be optimal in such situations.

[9] While administrative and trading costs are likely proportional to assets, research expenses seem likely to have a fixed component. With fixed costs the value of small

To construct a measure of the incentive to increase portfolio riskiness toward the end of the year, we thus note that at the end of September of year t a k-year-old fund's expected year $t + 1$ growth rate takes the form

$$E[\text{Flow}_{it+1}] = E[(1 + \gamma_k)f(r_{\text{Sept.}} + u) + \delta_k + \alpha X_{it}],$$

where $r_{\text{Sept.}}$ is the fund's year-to-date excess return and u is a random variable representing the fund's excess return in the final quarter of year t. If we assume that u is an appropriately distributed random variable with mean zero and standard deviation σ, and that v has mean zero and standard deviation $\sigma + \Delta\sigma$, then the expected increase in a fund's year $t + 1$ growth rate that results from increasing its quarterly standard deviation from σ to $\sigma + \Delta\sigma$ (or, more precisely, changing its fourth-quarter excess return distribution from u to v) is

$$h_k(r_{\text{Sept.}}; \sigma, \Delta\sigma) = E[(1 + \gamma_k)[f(r_{\text{Sept.}} + v) - f(r_{\text{Sept.}} + u)]].$$

Note that this measure of a benefit of increased risk is simply a linear functional of the function f we estimated in the previous section. We can thus estimate h_k consistently (on a bounded set of r values) by simply plugging our estimates of $\hat{f}(\cdot)$ and $\hat{\gamma}_k$ into the formula above (at least provided that u and v have bounded support).

Figure 3*a* provides our first clear look at the incentives to take risks created by the flow-performance relationship. The figure presents an estimate of the function $h_2(r_{\text{Sept.}}; \bar{\sigma}, 0.5\bar{\sigma})$ along with pointwise 90 percent confidence bands for the estimate. The function is interpreted as the expected increase in the year $t + 1$ growth rate of a 2-year-old fund that results from its increasing its risk in the fourth quarter of year t from the sample average to 50 percent above this average.[10] In accord with the intuition developed above, the point estimates graphed in the figure indicate that funds that are somewhat behind the market have an incentive to gamble and try to catch up, whereas funds that are somewhat ahead of the market have an incentive to lock in their gains. The figure suggests in addition that these incentives reverse at extreme positions: funds that are well behind the market may want to reduce their risk, whereas

---

funds may be largely their option value, which will create additional incentives to take risks. This is an additional reason behind our decision to drop from the analysis funds with less than $10 million in assets. When returns affect more distant future flows as well, the correlation between larger growth rates and larger initial asset levels creates a dynamic incentive to carry risk, which we also ignore.

[10] More precisely, $\bar{\sigma}$ (= 0.037) was set to the sample standard deviation of fourth-quarter excess returns in the complete set of growth and growth and income funds in *Mutual Funds OnDisc.* The distributions of u and v were taken to be truncated normals.

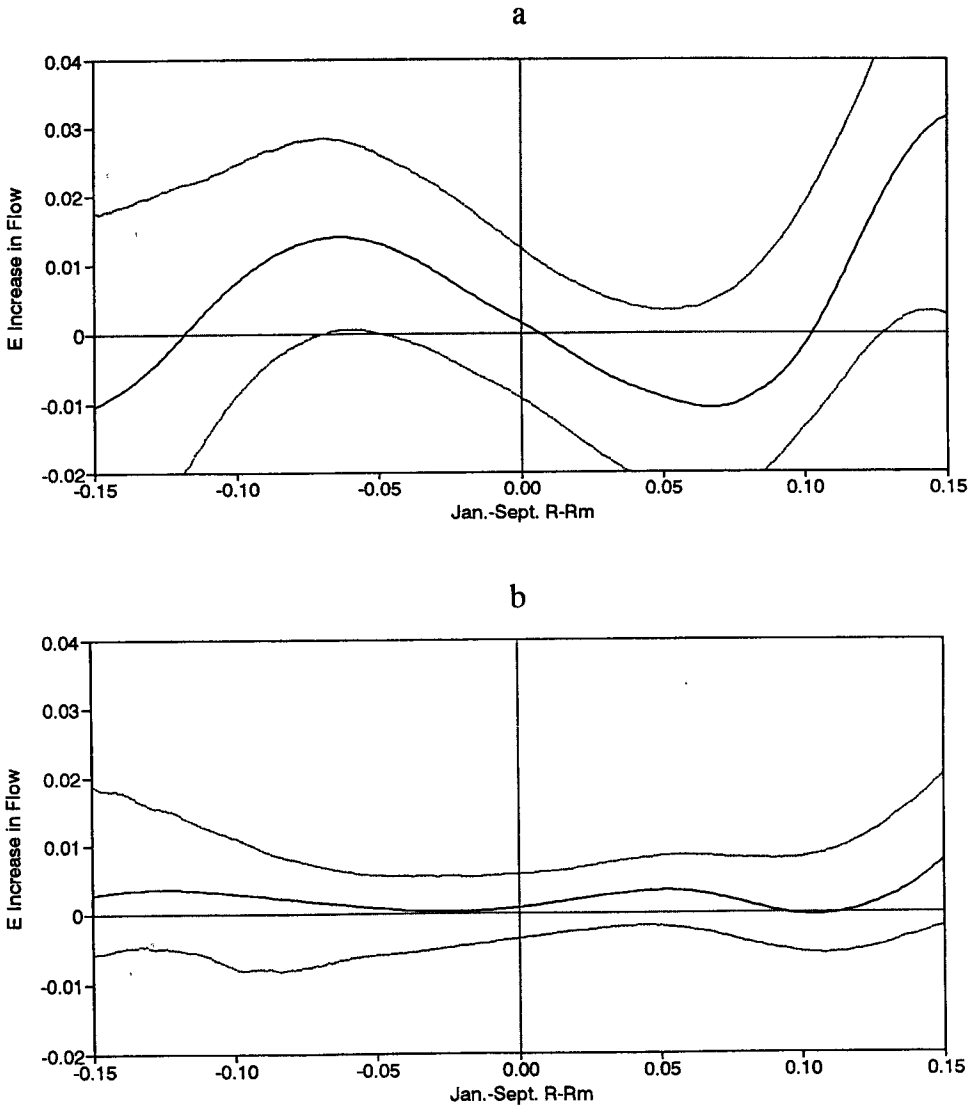

FIG. 3.—Risk incentives with 90 percent confidence bands. *a*, Funds of age 2. *b*, Funds of age 11 or more.

funds that are well ahead of the market may have a strong incentive to gamble. The figure also quantifies the incentives: such an increase in risk increases the expected growth rate of a fund that is somewhat behind the market by one percentage point and may increase the expected growth rate of a fund that is well above the market by over three percentage points. Note also that given our functional form assumptions and the estimated γ's, the incentives for 3-, 4-, and 5-year-old funds are identical in shape to the graph shown here, but scaled down by 28 percent, 34 percent, and 46 percent, respectively.

The confidence bands in the figure were calculated by estimating

the flow-performance relationship and deriving the risk incentive functions on 7,900 simulated data sets that were created as before by sampling with replacement from the age-specific empirical distribution of residuals from our base semiparametric models. The fact that the lower of these bands is positive in two places indicates that in pointwise tests the incentives are significantly different from zero at the 10 percent level for funds that are well ahead of the market and for funds that are a little more than five points behind the market. The incentive to lock in gains falls just short of significance at this level for funds that are about five points ahead of the market.

Figure 3*b* presents the corresponding graph for funds of age 11 or more. Note first that, as might be expected given that flows into these funds are much less sensitive to recent performance than flows into younger funds, the estimated incentives are much weaker than for the 2-year-old funds. The potential increase in the year $t + 1$ growth rate is estimated to exceed one-half of one percentage point only for funds that are very far above the market. In addition, the pattern of the estimated incentives is much less striking. The older funds appear frequently to have an incentive to increase their riskiness toward the end of the year, but we do not observe clearly defined regions in which the incentives are much stronger or weaker. While the confidence bands are tighter for older funds, we are unable in a pointwise test to reject that the incentive at any intermediate return level is zero. In a joint test, however, the average across all the return levels in the figure is significantly positive at the 10 percent level.

V. Do Funds Alter Their Risk in Response to These Incentives?

In this section, we come to the question that most interests us: investigating whether mutual funds adjust the riskiness of their portfolios in response to the incentives created by the flow-performance relationship. We have already seen that the incentive of a mutual fund to hold unsystematic risk in its portfolio is affected by the fund's position relative to the market portfolio at the end of September. Here, we explore the ways in which mutual funds actually alter their portfolios between the end of September and the end of December.

This section primarily focuses on detailed data on the *equity* portion of funds' portfolios—data that contain complete listings of the common stock holdings of 398 growth and growth and income mutual funds both at the end of September and at the end of December of certain years (for a total of 839 fund-years of data). With these data, we analyze how funds alter the riskiness of their portfolios to-

ward the end of the year and show that these changes appear to be related to the incentives to take risks we have previously identified. Subsequently, to allay fears that the changes we find in the equity portion of portfolios may not reflect the nature of overall portfolio changes, we look also at measures of riskiness constructed from time-series data.

We examine changes in fund riskiness that occur between the end of September and the end of December. In the real world, of course, opportunities to alter a portfolio arrive in continuous time. Funds that have a strong incentive to carry unsystematic risk in the final quarter may have already increased their riskiness before September is over. They might also begin to reduce their risk back toward the level that will be optimal in January before the end of December. Nonetheless, given the structure of the data available to us, it seems reasonable to hope that these effects are not so large as to eliminate any correlations entirely and to ask whether portfolio changes are in the direction we would predict from the incentives that we have identified.

To address this question, our primary measure of the riskiness of a fund's portfolio at a point in time is the sample standard deviation of the difference between the return on the portfolio and the return on the market portfolio (calculated from historical data on the component securities). Because of the implicit assumption in our flow regressions that investors react to the simple difference between a fund's return and the market return, it is in terms of this measure that we computed incentives to change risk levels. Note that the total variance of a portfolio in this sense can be decomposed into two parts:

$$\text{var}(r_i - rm) = \text{var}(r_i - \beta rm) + (\beta - 1)^2 \,\text{var}(rm).$$

The first term in this expression is the fund's unsystematic risk, the part of risk not associated with movements in the market. The second is increasing in the distance of a fund's beta from unity reflecting variance from implicit bets with or against the market as a whole. At times we shall also explore changes in each of these components of riskiness separately.

A. *Basic Tests of Reactions to Incentives*

In this subsection, we use a simple regression to show that risk changes are related to the incentives we have previously identified. We focus solely on the equity portion of funds' portfolios in this subsection, because our detailed security-level data are limited in two ways. First, our data contain the shares of fund assets that are

held in cash, bonds, and other securities only for a subset of the funds for which we have equity portfolio data. Second, even when we do know the shares of the portfolio held in cash or bonds, the data do not contain security-level descriptions of the bond holdings, derivative positions, and so forth, so that we can give only crude estimates of the riskiness of the fund's complete portfolio. Because the mutual funds in this study are garden-variety growth or growth and income funds, the vast majority of the funds' holdings are in common stocks. For the funds in our sample for which we have holdings data, the median fund has 88 percent of its holdings in common stocks (and no derivatives). Hence we may hope that changes in the equity portfolio provide a meaningful reflection of risk changes.

For each equity portfolio, risk measures were calculated from the CRSP daily return data for the prior year. Thus, for example, when calculating the beta for a stock held by a mutual fund in September or December of 1990, we examined the covariance of the stock's return with the return on the market portfolio for January through December of 1989.[11] Prior year data were used so that changes in portfolio risk reflect actual changes in the portfolio, not changes in the measured riskiness of the component securities. The beta of a portfolio was calculated by taking the weighted average of the betas for the component securities, and the standard deviation measures were calculated by taking the appropriate weighted sum of the variances and covariances of the returns of the individual securities. One difficulty with these calculations is that it was impossible to calculate the risk measures for every security in every portfolio because matches to the CRSP database could not always be made. In the regressions below, we use data only from those mutual funds for which we could match and obtain historical data on at least 85 percent of the total holdings (by value) in the September portfolio.

The basic hypothesis generated by the analysis of the flow-performance relationship is that mutual funds alter the variance of their portfolio returns around the return of the market in order to increase the expected flow of funds into the mutual fund. We test this hypothesis first via several simple cross-section regressions. The dependent variable in each specification is the change in a measure of the riskiness of each portfolio between the end of September and the end of December. The primary independent variable of interest, RiskIncentive, is a measure of the incentive of each fund to increase its riskiness calculated in much the same manner as we calculated the incentive to increase risk as a function of the end-of-September

[11] Our market return proxy is a value-weighted combination of New York Stock Exchange, American Stock Exchange, and NASDAQ returns.

position relative to the market in the previous section. The only differences are that the incentives are scaled up or down by the factor γ_k appropriate to the age of the fund and that the risk incentives are calculated as the incentive to increase the standard deviation not from the mean but instead from its end-of-September level. Note that the incentive variable has already been scaled to take into account each fund's age and so forth, so that it is natural for the variable to enter additively with a constant coefficient. Because the cost of increasing or decreasing the risk level of a portfolio may increase more than proportionally with a fund's assets under management, we include in the regression the risk incentive measure multiplied by the natural log of total fund assets. A negative coefficient for this variable would indicate that when faced with a comparable (in proportion to assets) incentive to take risks, larger mutual funds adjust their overall risk level less than smaller mutual funds. To allow for the possibility of mean reversion in measured portfolio riskiness, we include the September level of the risk measure in the regression. Finally, we also include in the regression the natural log of total assets of the fund.

The risk incentive measures are, of course, generated by our flow-performance estimation. To correct for the presence of generated regressors, standard errors for the regression were computed via a bootstrap procedure. This involved simulating pairs of data sets. The semiparametric model was estimated to generate the risk incentive measure on the first simulated data set, and the regression was then run on the second using the generated variable.

We exclude from the estimations the set of funds mentioned before for which we thought the flow-performance relationship might differ from that of standard retail growth and growth and income funds. Specifically, we excluded index funds, funds that were closed to new investors, funds with high minimum initial purchase requirements, funds with very high expense ratios, funds with total assets of less than \$10 million, and funds that merged during the fund-year. Summary statistics are reported in table 3.

The regression results are reported in table 4. Column 1 shows the results for the basic hypothesis test. The risk measure used, $\Delta SD(r - rm)$, is the change in the standard deviation of the difference between the fund return and the return on the market portfolio. The coefficient for the risk incentive measure is positive, whereas the coefficient on the interaction of the risk incentive measure and log(Assets) is negative, indicating that small funds at least do adjust their riskiness in the direction of the incentive we have measured and that the magnitude of the response is larger for smaller funds. The risk incentive coefficient is significantly different from zero at

TABLE 3

SUMMARY STATISTICS FOR RISK CHANGE REGRESSIONS

Variable	Mean	Standard Deviation
Sept. SD$(r - rm)$	.053	.022
Sept. SD$(r - \beta rm)$	.048	.020
Sept. $\lvert\beta - 1\rvert$	.153	.115
ΔSD$(r - rm)$	.0019	.008
ΔSD$(r - \beta rm)$	.0017	.007
$\Delta\lvert\beta - 1\rvert$	.006	.045
RiskIncentive	.0021	.003
log(Assets)	5.269	1.475
Sept. share matched	.936	
Dec. share matched	.918	

TABLE 4

SEPTEMBER–DECEMBER RISK CHANGES AND INCENTIVES ($N = 464$)

	DEPENDENT VARIABLE		
INDEPENDENT VARIABLE	ΔSD$(r - rm)$ (1)	ΔSD$(r - \beta rm)$ (2)	$\Delta\lvert\beta - 1\rvert$ (3)
RiskIncentive	.87	.73	2.22
	(.28)	(.24)	(1.40)
RiskIncentive − log(Assets)	−.15	−.11	−.57
	(.05)	(.04)	(.28)
log(Assets)	.00013	−.00003	.003
	(.0003)	(.0002)	(.002)
Sept. risk level	.05	.03	−.007
	(.02)	(.02)	(.02)
Constant	−.001	−.0002	−.006
	(.002)	(.002)	(.009)
R^2	.03	.03	.01

NOTE.—Estimated standard errors are in parentheses.

the 1 percent level. It is, however, important to note that the estimated response of portfolio risk changes to the risk incentive diminishes quickly as fund size rises. In a one-tailed test, we can reject at the 5 percent level the null hypothesis that the coefficient for the risk incentive measure plus the coefficient for the risk incentive measure multiplied by the log of assets equals zero for funds that are the size of the median fund in our sample.

The magnitude of the effect is fairly small but still has practical relevance. For example, for a 2-year-old fund RiskIncentive may take on a value of 0.01 for a fund that is a few points behind the market

and 0.03 for a fund that is well ahead of the market. For the smallest funds we are using (those with an asset value of approximately $10 million), the coefficients in the regression imply that the expected increases in the standard deviation are approximately 0.005 and 0.015 in the two situations. Such changes would require increasing the variance of the mean portfolio in our sample by about 17 percent and 48 percent, respectively. Contrary to what we might have expected, there seems to be no evidence of mean reversion in portfolio riskiness.

As we noted above, we may decompose the riskiness of a portfolio (in the sense we have used it) into the portfolio's level of unsystematic risk and the departure of its beta from unity. Column 2 of table 4 repeats the standard regression with our measure of the change in unsystematic risk, $\Delta SD(r - \beta rm)$, as the dependent variable. The results are very much like those discussed above, with each coefficient again significant at the 1 percent level.

If consumers do indeed react to the simple difference between a fund's return and the market return as they do in our econometric specification, then a fund that wishes to increase its level of risk may also do so by moving its beta away from unity. We would like to emphasize, however, that our specification of flows as a function of $r - rm$ instead of $r - \beta rm$ was dictated by data constraints, and we do not regard our results or anyone else's as providing any convincing evidence on the extent to which consumers take betas into account in judging performance. While we shall therefore be hesitant to claim that any regression explaining changes in betas is a clean test of reactions to incentives, we do feel that such a regression is at least of descriptive interest. Column 3 of table 4 repeats the basic risk change specifications of columns 1 and 2, with $\Delta|\beta - 1|$ as the dependent variable and Sept.$|\beta - 1|$ as the initial risk level. The results show a positive coefficient for the risk incentives measure and a negative coefficient for the risk incentive measure interacted with size. However, the coefficient on RiskIncentive is significantly different from zero at only the 11 percent confidence level.

B. A More Detailed Picture of Equity Risk Changes

While the simple regressions above provide a straightforward test of the hypothesis that funds react to the incentives we have identified, they may leave one wondering exactly how sharp the correspondence is between the patterns of actual incentives and behavior. We thus turn now to the task of providing a more detailed picture of actual risk changes.

Recall from Section IV that the pattern of incentives to manipulate

the riskiness of a mutual fund depends on the age of the fund. For "young" funds, the payoff to increasing risk (pictured in fig. 3*a*) is negative for funds that are far behind the market, increases with a fund's January–September excess return until reaching a local maximum at about six points behind the market, then decreases to a minimum for funds that are about seven points ahead of the market, and finally increases sharply to reach its highest level for funds that are far ahead of the market. For "old" funds, an incentive to increase risk (pictured in fig. 3*b*) appears to exist fairly generally, although the pattern of the incentive is not clear.

To assess whether actual risk changes follow such patterns in detail, we felt that it would be most valuable simply to produce pictures of actual changes that could be compared visually with figure 3. Because of the limited number of funds for which we have portfolio data, we chose not to attempt another semiparametric analysis, but instead simply to fit a piecewise linear model to the relationship between risk changes and January–September excess returns. Specifically, we used nonlinear least squares to estimate the equation

$$\begin{aligned}\Delta \mathrm{SD}(r_i - rm) = {} & \alpha_0 + \alpha_1 \log(\text{Assets}) \\ & + \alpha_2 \text{Sept. SD}(r_i - rm) \\ & + h(\text{Jan.–Sept. } r_i - rm) + \epsilon_i,\end{aligned}$$

where $h(x)$ was a continuous, piecewise linear function having five parameters: the locations of two kink points (KINK1 and KINK2) and the slopes in the regions to the left of both kinks (SLOPELEFT), between the two (SLOPEMID), and to the right of both (SLOPERIGHT). For the sample of young funds, for example, if the pattern of actual risk changes closely matched the estimated incentives to take risks shown in figure 3, we would expect to find that KINK1 was approximately −0.06 and KINK2 was approximately 0.07. Figure 3 also suggests that we would expect to find that SLOPELEFT and SLOPERIGHT are positive, whereas SLOPEMID should be negative.

Coefficient estimates from the piecewise linear model for the subsample of young funds are presented in column 1 of table 5. The correspondence between actual changes and incentives is fairly strong. The first estimated kink point, KINK1, is located at −0.06, with the estimate being highly significant. The second turning point, KINK2, is estimated very imprecisely, with the estimate of 0.001 not being significantly different from 0.07. The estimated slope of the first segment, SLOPELEFT, is positive as predicted and significantly different from zero at the 5 percent level. The estimated slope of the second segment, SLOPEMID, is negative as predicted but not significantly different from zero. However, SLOPEMID is signifi-

TABLE 5

ACTUAL RISK CHANGES AS A FUNCTION OF JANUARY–SEPTEMBER RETURNS

	DEPENDENT VARIABLE: $\Delta SD(r - rm)$	
INDEPENDENT VARIABLE	Young Funds ($N = 94$) (1)	Old Funds ($N = 375$) (2)
Constant	−.02	−.003
	(.006)	(.002)
KINK1	−.06	.03
	(.02)	(.03)
KINK2	.001	.08
	(.04)	(.03)
SLOPELEFT	.09	−.02
	(.04)	(.01)
SLOPEMID	−.07	.06
	(.07)	(.08)
SLOPERIGHT	.003	−.05
	(.02)	(.04)
log(Assets)	−.0007	.00002
	(.0006)	(.0003)
Sept. $SD(r - rm)$	−.02	.09
	(.04)	(.02)
R^2	.11	.08

NOTE.—Estimated standard errors are in parentheses.

cantly different from SLOPELEFT. The slope of the third segment, SLOPERIGHT, is positive although very small and again not significant. To facilitate comparisons with the incentives shown in figure 3, predicted values from this regression are graphed in figure 4*a*. From the graph it appears that the most salient feature of the actual risk change data is that funds that are somewhat behind the market do tend to increase the riskiness of their portfolios toward the end of the year as though trying to gamble and catch the market.

Our analysis of risk incentives for older funds did not reveal a clear pattern. Therefore, it is not surprising that no clear patterns of actual risk changes appear in column 2 of table 5, which presents estimates of the same piecewise linear function on the subsample of old funds. Neither of the turning points is estimated precisely, and only the first of the estimated slopes is significantly different from zero at the 10 percent level. Predicted values from this regression are graphed in figure 4*b*. The one potentially interesting regularity in the figure is that we find that over a wide range of January–September excess return levels, funds do tend to increase their riskiness toward the end of the year. While this is very much in line with

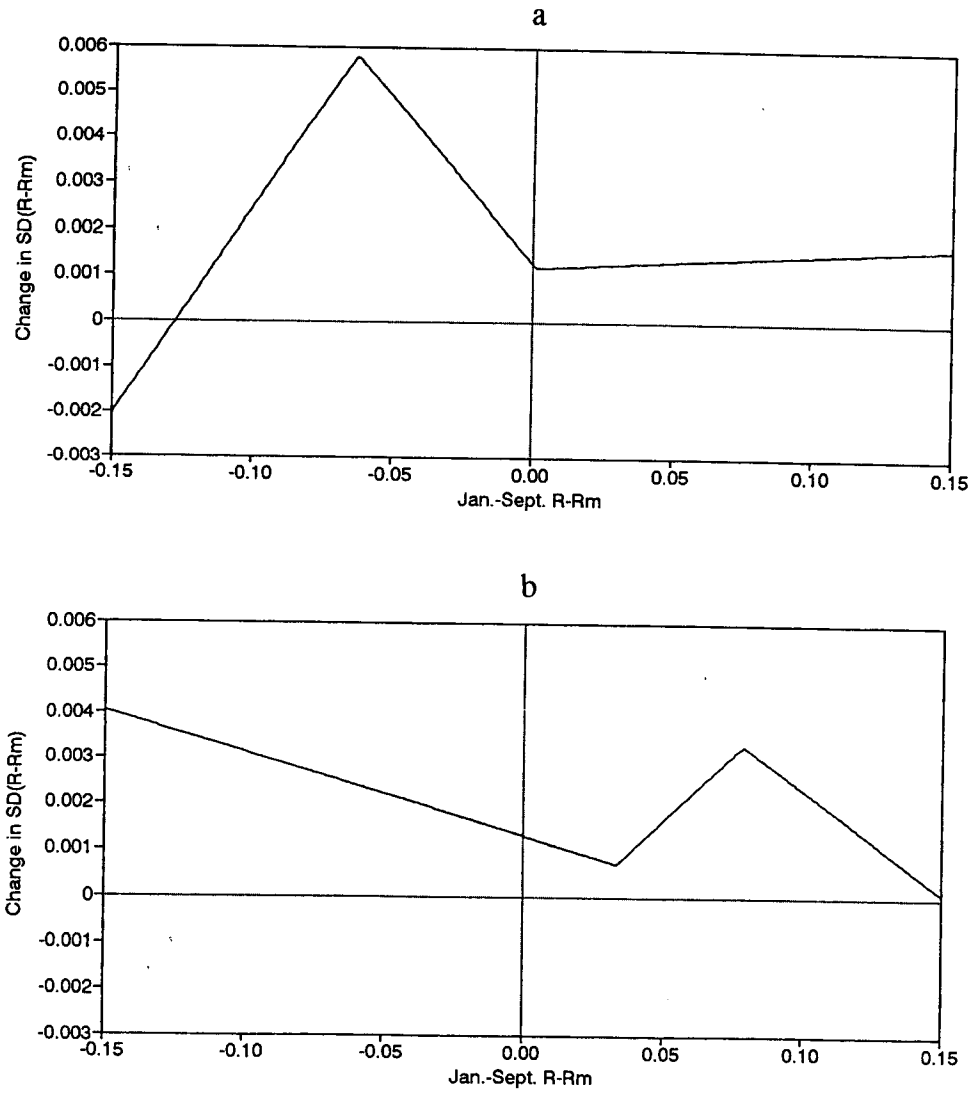

FIG. 4.—Pattern of actual risk changes: fitted values from piecewise linear model. *a*, Young funds' equity portfolios. *b*, Old funds' equity portfolios.

the incentives we have identified, we hesitate to emphasize the result because of the potential that our sample selection has biased the results.

C. *Changes in Total Portfolio Riskiness*

So far, we have used the detailed data available to us to show that changes in the riskiness of the equity portion of funds' portfolios reflect the incentive to attract investment flows. While we lack data of a similar quality on the remainder of funds' holdings, we attempt here to look at measures of risk of the broader portfolios and pro-

vide some assurance that changes in the riskiness of equity holdings are not undone by other changes. To do this, we examine changes in portfolio risk measures constructed using time-series data on fund returns.[12]

An obvious alternative to our portfolio construction procedure (used, e.g., by Brown et al. [1996]) is to try to estimate the riskiness at separate points in time using time-series data on fund returns. From Morningstar's *Mutual Funds OnDisc* we have available monthly return data for a large sample of funds and can construct noisy estimates of the change in a fund's riskiness by comparing the sample variance of a fund's monthly excess returns for the January–September and October–December periods.

Examining risk changes in this way introduces a number of measurement error problems. First, in the absence of high-frequency return data, the risk measures that such an approach relies on will be quite noisy (being estimated from as few as three data points). This problem is, however, ameliorated by our being able to use a much larger sample because we require much less data about each individual fund. Second, and perhaps more important, because mutual funds change their composition over time, an estimate of the variance of the September portfolio computed from January–September returns will be biased, with the bias correlated with the level of January–September excess returns—our primary explanatory variable.[13] Ignoring these problems, we now proceed with such an analysis.

The data available to us contain monthly returns for a sample of growth and growth and income funds between 1983 and 1993 (for a total of 3,163 fund-years). For each fund, we construct a measure of the change in the variance of the fund's simple excess return by

$$\Delta\text{TS var}(r_i - rm) = \frac{1}{3}\sum_{j=\text{Oct.}}^{\text{Dec.}} (r_{ij} - rm_j)^2 - \frac{1}{9}\sum_{j=\text{Jan.}}^{\text{Sept.}} (r_{ij} - rm_j)^2.$$

To explore how end-of-year risk changes measured in this way vary with a fund's January–September excess return, we computed non-

[12] The working paper version of this paper (Chevalier and Ellison 1995) contains many additional robustness checks, including an analysis of changes in funds' allocations of assets between stocks, bonds, derivatives, etc., estimates obtained from subsets of the data in which measurement errors and survivorship biases should be less important.

[13] One can regard Brown et al.'s (1996) decision to simply compare the end-of-year riskiness of the groups ahead and behind the market in September as an attempt to overcome this problem if one assumes that the biases will be equal for funds that are symmetrically ahead and behind the market so that the bias cancels from the comparison. Such an approach, however, is less helpful if we want to explore the pattern of risk changes in more detail.

TABLE 6

RISK CHANGES FROM TIME SERIES AS A FUNCTION OF JANUARY–SEPTEMBER RETURNS

	DEPENDENT VARIABLE: ΔTS var$(r - rm)$	
INDEPENDENT VARIABLE	Young Funds ($N = 869$) (1)	Old Funds ($N = 2,294$) (2)
Constant	.001	.001
	(.0008)	(.0005)
KINK1	−.032	−.036
	(.030)	(.032)
KINK2	.054	.217
	(.037)	(.043)
SLOPELEFT	.024	.024
	(.010)	(.006)
SLOPEMID	−.004	.011
	(.013)	(.003)
SLOPERIGHT	.024	.085
	(.010)	(.040)
R^2	.02	.03

NOTE.—Estimated standard errors are in parentheses.

linear least-squares estimates of a piecewise linear model similar to that used in Section V*B* (but omitting controls for fund size and initial riskiness).

Coefficient estimates from the subsamples of young and old funds are presented in table 6, with the predicted values of the functions graphed in figure 5 so that they may be compared with both the incentives pictured in figures 3 and 4 and the estimates of risk changes from the equity data pictured in figure 4. For the young funds, the point estimates on the three slopes again follow a positive, negative, positive pattern, with one kink to the left of zero and the other to the right of zero. However, the middle slope is not nearly as negative as the risk incentive picture suggests it should be, and the locations of the kinks are not estimated precisely.

In the data on older funds, we see that risk changes do again exhibit something of a tendency to be positive. The one clear regularity in the data that is somewhat puzzling in contrast with our earlier results is that higher January–September excess returns are clearly correlated with larger risk increases. This pattern is quite compatible with the results of Brown et al.'s (1996) analysis of monthly return time series. Looking back at the top panels of figures 4 and 5, we see that the time-series approach produces a more positive correlation

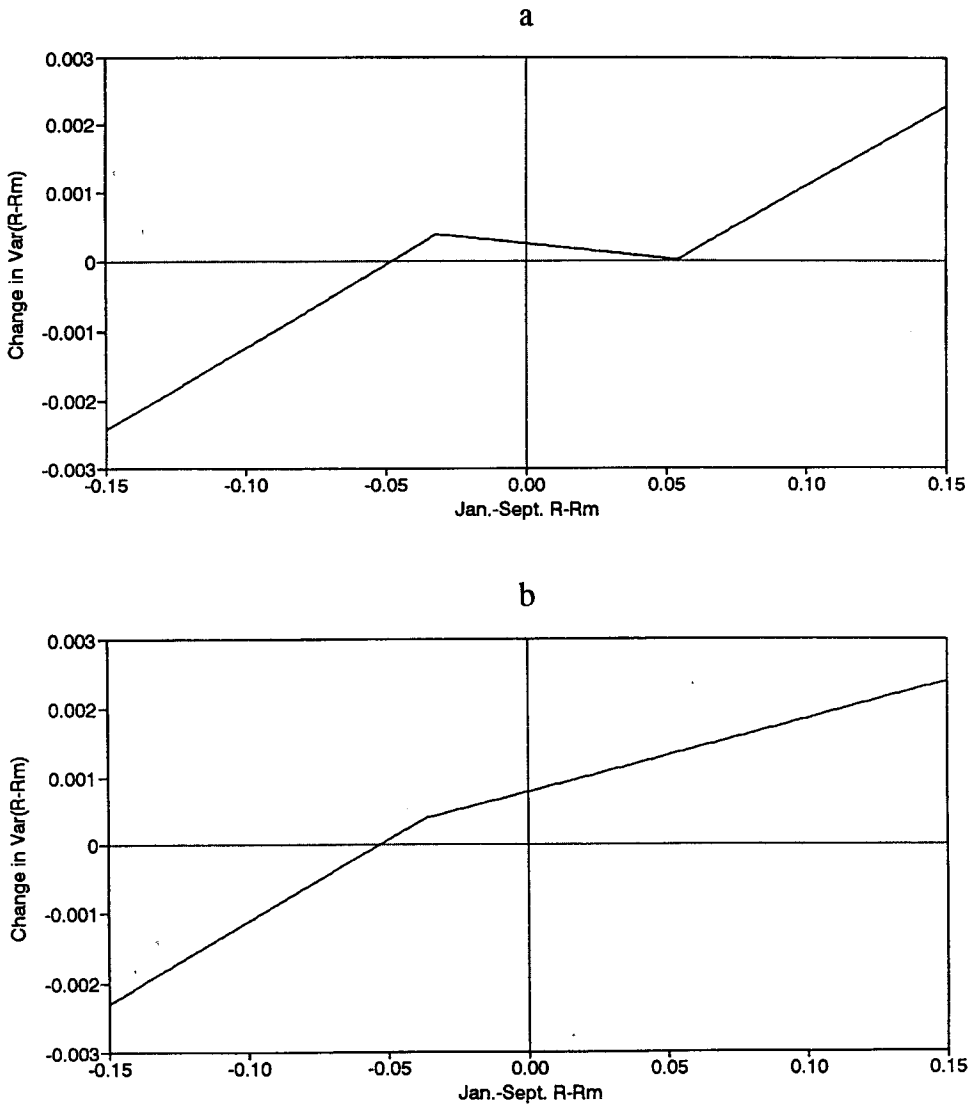

FIG. 5.—Pattern of actual risk changes via time series: fitted values from piecewise linear model. *a*, Young funds. *b*, Old funds.

between excess returns and risk increases than is visible in the equity portfolios for the young funds also.

VI. Conclusion

In this paper we interpret the flow-performance relationship as an incentive scheme implicitly given to mutual fund companies by mutual fund investors. We show that the flow-performance relationship can generate incentives for mutual fund companies to increase or

decrease the riskiness of their portfolios. Finally, we show that mutual fund companies respond to this incentive scheme: funds alter their portfolios between September and December in a manner consistent with the September incentive to take risk calculated from the flow-performance relationship.

The methodology of treating the flow-performance relationship as an incentive scheme could be used to examine other hypotheses about the behavior of mutual funds or institutional asset managers. For example, Scharfstein and Stein (1990) and Zwiebel (1995) present models in which optimal performance evaluation gives managers an incentive to "herd." Lakonishok, Shleifer, and Vishny (1992) and Grinblatt et al. (1995) have used data on institutional asset managers and mutual funds, respectively, to examine whether asset managers exhibit herd behavior. An interesting extension of the methodology in this paper would be to test the performance evaluation component of Scharfstein and Stein's and Zwiebel's models directly and see whether mutual funds receive a larger payoff if they herd. If, for example, we were to find no evidence that institutional investors disproportionately invest with asset management firms in a manner that would encourage them to herd, this could help to explain why Lakonishok, Shleifer, and Vishny (1992) found little evidence that these investment managers have a tendency to herd.

Similarly, Lakonishok et al. (1991) have examined whether institutional asset managers engage in window dressing, selling off poor-performing assets from their portfolios prior to issuing year-end holdings reports. Their results suggest that very little window dressing is undertaken by their sample of managers. An extension of the methodology of this paper would be to examine whether investors believe that window-dressing managers are of higher quality than managers with equal performance who have not window-dressed. For example, controlling for the overall returns of the portfolios, one could test whether investors are less likely to invest with a management company that reports holding many "losers" in its year-end report. If investment flows do not systematically differ for funds with the same performance but different amounts of window dressing, then we should not expect funds to engage in the costly activity of window dressing. Our preliminary forays in this direction suggest that retail mutual fund companies are not rewarded for window dressing.

Finally, our examination of the flow-performance relationship could serve as a starting point for further examination of decision making by investors. A comparison of the shape of the flow-performance relationship for the retail mutual funds we have studied to the shape of the flow-performance relationship for products pur-

chased exclusively by institutional investors may provide insight into the evaluative procedures used by different types of investors.

References

Asch, Beth J. "Do Incentives Matter? The Case of Navy Recruiters." *Indus. and Labor Relations Rev.* 43 (suppl.; February 1990): S89–S106.

Berkowitz, M. K., and Kotowitz, Yehuda. "Incentives and Efficiency in the Market for Management Services: A Study of Canadian Mutual Funds." *Canadian J. Econ.* 26 (November 1993): 850–66.

Borenstein, Severin, and Zimmerman, Martin B. "Market Incentives for Safe Commercial Airline Operation." *A.E.R.* 78 (December 1988): 913–35.

Bronars, Stephen. "Risk Taking Behavior in Tournaments." Manuscript. Santa Barbara: Univ. California, 1987.

Brown, Keith C.; Harlow, W. V.; and Starks, Laura T. "Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry." *J. Finance* 51 (March 1996): 85–110.

Chevalier, Judith A., and Ellison, Glenn D. "Risk Taking by Mutual Funds as a Response to Incentives." Working Paper no. 5234. Cambridge, Mass.: NBER, August 1995.

Ehrenberg, Ronald G., and Bognanno, Michael L. "Do Tournaments Have Incentive Effects?" *J.P.E.* 98 (December 1990): 1307–24. (*a*)

———. "The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour." *Indus. and Labor Relations Rev.* 43 (suppl.; February 1990): S74–S88. (*b*)

Ellison, Glenn D., and Ellison, Sara Fisher. "A Simple Framework for Nonparametric Specification Testing." Discussion Paper no. 1662. Cambridge, Mass.: Harvard Inst. Econ. Res., 1993.

Goetzmann, William N., and Peles, Nadav. "Cognitive Dissonance and Mutual Fund Investors." *J. Financial Res.* (in press).

Grinblatt, Mark, and Titman, Sheridan D. "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings." *J. Bus.* 62 (July 1989): 393–416.

———. "The Persistence of Mutual Fund Performance." *J. Finance* 47 (December 1992): 1977–84.

———. "Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns." *J. Bus.* 66 (January 1993): 47–68.

Grinblatt, Mark; Titman, Sheridan D.; and Wermers, Russ. "Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior." *A.E.R.* 85 (December 1995): 1088–1105.

Healy, Paul M. "The Effect of Bonus Schemes on Accounting Decisions." *J. Accounting and Econ.* 7 (April 1985): 85–107.

Holmstrom, Bengt. "Managerial Incentive Problems—a Dynamic Perspective." In *Essays in Economics and Management in Honor of Lars Wahlbeck.* Helsinki: Swedish School Econ., 1982.

Holmstrom, Bengt, and Milgrom, Paul R. "Aggregation and Linearity in the Provision of Intertemporal Incentives." *Econometrica* 55 (March 1987): 303–28.

Ippolito, Richard A. "Consumer Reaction to Measures of Poor Quality: Evi-

dence from the Mutual Fund Industry." *J. Law and Econ.* 35 (April 1992): 45–70.

Knoeber, Charles R., and Thurman, Walter N. "Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production." *J. Labor Econ.* 12 (April 1994): 155–79.

Lakonishok, Josef; Shleifer, Andrei; Thaler, Richard; and Vishny, Robert W. "Window Dressing by Pension Fund Managers." *A.E.R. Papers and Proc.* 81 (May 1991): 227–31.

Lakonishok, Josef; Shleifer, Andrei; and Vishny, Robert W. "The Impact of Institutional Trading on Stock Prices." *J. Financial Econ.* 32 (August 1992): 23–43.

Patel, Jayendu; Zeckhauser, Richard; and Hendricks, Darryll. "Investment Flows and Performance: Evidence from Mutual Funds, Cross-Border Investments, and New Issues." Manuscript. Cambridge, Mass.: Harvard Univ., Kennedy School Government, 1990.

Prendergast, Canice, and Stole, Lars. "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning." *J.P.E.* 104 (December 1996): 1105–34.

Robinson, Peter M. "Root-N-Consistent Semiparametric Regression." *Econometrica* 56 (July 1988): 931–54.

Roston, Marc. "Mutual Fund Managers and Life Cycle Risk: An Empirical Investigation." Manuscript. Chicago: Univ. Chicago, Dept. Econ., 1996.

Scharfstein, David S., and Stein, Jeremy C. "Herd Behavior and Investment." *A.E.R.* 80 (June 1990): 465–79.

Sirri, Erik R., and Tufano, Peter. "Buying and Selling Mutual Funds: Flows, Performance, Fees, and Services." Manuscript. Boston: Harvard Bus. School, 1993.

Wermers, Russ. "Herding, Trade Reversals, and Cascading by Institutional Investors." Manuscript. Boulder: Univ. Colorado, 1993.

Zwiebel, Jeffrey. "Corporate Conservatism and Relative Compensation." *J.P.E.* 103 (February 1995): 1–25.