WILEY-BLACKWELL

American Finance Association

Preferences for Stock Characteristics as Revealed by Mutual Fund Portfolio Holdings
Author(s): Eric G. Falkenstein
Source: *The Journal of Finance*, Vol. 51, No. 1 (Mar., 1996), pp. 111-135
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2329304
Accessed: 25/11/2009 10:15

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=black.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Blackwell Publishing and *American Finance Association* are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Finance.*

STOR®
http://www.jstor.org

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

Preferences for Stock Characteristics As Revealed by Mutual Fund Portfolio Holdings

ERIC G. FALKENSTEIN*

ABSTRACT

This investigation of the cross-section of mutual fund equity holdings for the years 1991 and 1992 shows that mutual funds have a significant preference towards stocks with high visibility and low transaction costs, and are averse to stocks with low idiosyncratic volatility. These findings are relevant to theories concerning investor recognition, a potential agency problem in mutual funds, tests of trend-following and herd behavior by mutual funds, and corporate finance.

THIS PAPER DOCUMENTS THE revealed preferences of U.S. open-end mutual funds for various stock characteristics. Using Morningstar's data on portfolio holdings, I examine the cross-section of the mutual fund sector's percentage ownership of a particular security; that is, the tests measure stocks of holdings and not flows as has been done previously (see Kraus and Stoll (1972), Lakonishok, Shleifer, and Vishny (1991, 1992), and Grinblatt, Titman, and Wermers (1995)). Using this data for the years 1991 and 1992, I am able to determine the percentage of shares outstanding held by these mutual funds for NYSE and AMEX listed stocks. The most indisputable conclusion that can be drawn is that mutual fund preferences for stocks are not totally driven by conventional proxies for risk alone. Mutual funds show significant preferences towards various security characteristics relevant for understanding the behavior of these institutions.

The empirical results are summarized as follows. Mutual funds display a nonlinear preference towards stocks with high volatility (ownership is concave in variance). Relevant to transaction costs, mutual funds show an aversion to low-price stocks, while demand is consistently increasing in liquidity (as measured by trading volume divided by the shares outstanding). Other than the small-cap sector, which specializes in small firms, funds show an aversion to small firms. Funds tend to avoid stocks with little information, as measured by the number of major newspaper articles or the number of months since listing on the exchange. The significance of these variables in explaining mutual fund ownership suggests relevant proxies for risk, transaction costs, and information generated by the firm.

* KeyCorp. I thank Phillip Braun, Ian Domowitz, Matt Jackson, Debbie Lucas, Mark Satterthwaite, René Stulz, and an anonymous referee for many helpful comments and suggestions. I am also grateful to seminar participants at the Kellogg Graduate School of Management. This paper draws on my dissertation, completed at Northwestern University in 1994.

Portfolio theory generates powerful predictions about the expected returns of assets as a function of their systematic variance, as well as the general nature of investor portfolios. Clearly some investors have comparative advantages buying certain types of assets, and these differentials imply that different groups of investors have portfolio holding biases toward different types of assets. Investigating this issue is interesting in itself as a microeconomic study of the firm, that is, what frictions influence portfolio decisions, and to what degree? In an example of this type of research, Kang and Stulz (1994) document that foreign investment holdings in Japan are biased towards large firms, among other variables. One potential friction in asset markets is firm information: the cost of procuring it and its precision and credibility. Since it seems reasonable that foreigners face relatively lower information asymmetries when investing in large firms, they hypothesize an information deficit drives this bias. This paper identifies several other relevant frictions.

This paper also directly addresses tests of herd behavior or trend following by mutual funds (see, for example Lakonishok, Shleifer, and Vishny (1991, 1992) or Grinblatt, Titman, and Wermers (1995)). Herding refers to any mass movement into particular stocks for whatever reason. Trend following is a specific type of herding by mutual funds that involves a large group of funds chasing stocks that have recently risen in value. The identification in this paper of several variables correlated with ownership implies that herding into these stocks must be occurring at some time, though I emphasize reasons unrelated to the models of herding by Welch (1992), Bannerjee (1992), or Scharfstein and Stein (1990). In those models, funds or their investors evaluate choices based in part on what others are doing, the net result being that investors tend to herd together. If mutual funds have a comparative advantage in holding securities with certain characteristics, however, then as a security acquires these characteristics funds in aggregate will purchase them. The strength of this empirical study is that it identifies several of these characteristics. In this way I generate and test an alternative interpretation of revealed trend-following or herding behavior: as stocks acquire specific characteristics, mutual funds are more likely to hold them. This may have nothing to do with under-weighing private information or responding disproportionately to short-run incentives. For example, I find that mutual funds show a strong aversion to low-priced stocks (priced less than \$5). Over time if a stock rises in price it will attract net mutual fund demand, evidence to those looking only at past stock return and net mutual fund demand as evidence of trend following. In reality the herding effect may be due to the comparative disadvantage of mutual funds buying low-priced stocks. Thus the empirical findings in this paper provide an alternative hypothesis to this line of research and at the least suggest variables to control for when testing for herding or trend following.

The remainder of the paper is organized as follows. Section I discusses the data used in the empirical study. Section II looks at the results of simple sorting of the data. Section III describes the estimation methodology for the censored regressions. Section IV describes the results relevant to separate

Table I

Summary Statistics on Mutual Funds

Below are characteristics of the mutual funds used in the censored regressions. Datasets were pulled from Morningstar's *Mutual Funds OnDisc*, January 1993 and January 1992. The January 1993 CD contains portfolio data for funds for the prior year, and similarly for the January 1992 CD.

	Year	
	1991	1992
Total number of funds covered by Morningstar	2,284	2,520
Number of funds utilized (w/ equity >50% of portfolio)	1,087	1,174
Total value of utilized funds equity holding	$275 billion	$340 billion
Mean value of utilized funds equity holding	$253 million	$290 million
Median value of utilized funds equity holding	$57 million	$70 million
Mean portfolio manager tenure	4.4 years	4.6 years
Median portfolio manager tenure	3 years	3 years
Mean portfolio turnover	95%	85%
Mean percentage of assets in equity	86%	86%
Mean percentage of assets in cash	8.8%	8.7%
Mean percentage of assets in bonds	3.7%	3.5%

hypotheses. In Section V, I summarize, interpret and discuss applications of the results.

I. Data

The main database was created by merging open-end mutual fund portfolio holdings information from the Morningstar company with security data from the Center for Research in Security Prices (hereafter CRSP). Since both databases have information on the tickers of securities (e.g., BUD is the ticker for Anheuser-Busch), I was able to determine various characteristics of a stock such as its price, beta, etc., as well as what percentage of its shares outstanding were held by mutual funds. In the tests that included book-to-market equity, I used only those firms that had COMPUSTAT's item 60 (book equity of common stock) the year before data on mutual fund ownership.

The Morningstar data set contains portfolio holdings for 2,284 and 2,520 mutual funds in 1991 and 1992, respectively (see Table I for summary statistics). I include only those mutual funds that hold more than 50 percent of their portfolio in equities, which leaves 1,087 and 1,174 funds in 1991 and 1992. This is done because funds with very small equity holdings relative to their bond portfolio frequently did not list their equity portfolios (they listed the bond holdings instead). Furthermore, it removes a rather small amount of net mutual fund equity holdings relative to the total.[1]

[1] The proportion of mutual fund equity holdings removed in this procedure was 2.6 percent in 1991 and 4.8 percent in 1992.

Table II

Summary Statistics on Mutual Fund's Total Equity Holdings

The variable own_i represents the estimated proportion of security i owned by the sample of mutual funds during the year. This is the dependent variable used in the censored regressions.

$$own_i = \sum_{m=1}^{M} \frac{\text{shares owned by fund m at time } t_m}{\text{shares outstanding at time } t_m}$$

where M is the total number of mutual funds in the specified sample.

	Year	
	1991	1992
Total observations	2,133	2,221
Number with $own_i = 0$ (i.e., mutual funds held no shares)	356	380
Mean own_i	5.2%	5.7%
Median own_i	3.6%	3.9%
Max own_i	78%	94%

Table II displays summary statistics on mutual funds' total equity holdings of individual securities. Constructing the metric of mutual fund ownership was done in the following manner. The portfolio holdings are from various dates within the year for different funds, so that some of the 1991 portfolio holdings correspond to April while others correspond to June or October of that year. To get an estimate of the mutual fund industry's holdings of a particular stock in a particular year, for each fund's portfolio I divide the number of shares held of a particular company by the number of shares outstanding on that date (the date is listed in the fund's portfolio holdings menu), and then add up this quotient for all mutual funds that reported holdings within the year. This is my dependent variable: percentage ownership of a stock by open-end mutual funds. Specifically, mutual fund ownership for a specific stock, own_i, is defined as the following

$$own_i = \sum_{m=1}^{M} \frac{\text{shares owned of stock } i \text{ by fund } m \text{ at time } t_m}{\text{shares outstanding of stock } i \text{ at time } t_m} \tag{1}$$

where t_m represents the date to which the portfolio data correspond for fund m's holding of stock i, and M is the total number of mutual funds in the specific category of interest (e.g., "growth" funds only). Though the ownership data is taken at different points within the portfolio year, it represents a noisy yet unbiased estimate of mutual fund demand in that year, as fluctuations in aggregate stock price levels do not affect the measured share ownership percentage. Tests of subsets of the mutual fund universe utilize only funds within the specified subset. For example, I test the explanatory power of various sets of variables not only on the entire mutual fund universe, but on

Table III

Total Equity Holdings by Investment Objective

Investment objectives are stated by the fund and documented by Morningstar. Equity ownership for the difference objectives is calculated as of December.

	% of Total Morningstar Mutual Fund Equity Holdings	
	1991	1992
Growth	34.9	37.5
Growth-income	26.4	28.7
Small company funds	5.6	6.7
Equity-income	4.4	4.9
Balanced	4.5	3.3
Aggressive growth	2.5	2.5

small funds only, or only funds within the investment objective of "growth." In these cases, I sum only over these subsamples.

This ownership metric has the following relevance. If all asset demanders, non-mutual funds and mutual funds, held similar portfolios, then the mutual fund industry as a whole would own equal proportions of each stock. For example, given that the open-end mutual fund sector on average owns 5.2 percent of every stock on the NYSE and AMEX in 1991, if this same set of mutual funds owns only 3 percent of IBM's shares, we can say that for some reason this sample of mutual funds was averse to holding IBM shares.

Table II shows that the mutual fund sector as a whole had zero holdings of approximately 17 percent of the securities passing the relevant criteria to be included in these tests (described below). The mean ownership statistic indicates that the mutual fund sector owns an average of 5.2 percent and 5.7 percent of the shares outstanding of stocks in 1991 and 1992, respectively. Thus the entire sample of open-end mutual funds represents a small yet significant portion of total equity ownership of the NYSE and AMEX. Ownership is also positively skewed in distribution (maximum ownership is 94 percent in 1992). To keep the extreme ownership observations from dominating regression parameter estimates, I use the logarithm of ownership in the regressions, which also increases the fit of the regressions.

Table III shows the importance of several investment objectives. Although Morningstar lists over 20 objectives, the top 6 represent the majority of equity ownership (over 75 percent of equity holdings). Furthermore, growth and growth-income alone represent about 60 percent of the total sample of mutual funds. The small company funds represent a sizeable portion of investor demand and have obvious preferences for small companies. The largest 3 of these objectives are tested independently to determine if there is significant heterogeneity between funds with different objectives. Data relating to the fund size and the tenure of the fund manager are listed in Table I and are used

Table IV

Summary Statistics on NYSE and AMEX Securities Used in Censored Regressions

These are various statistics on stocks utilized with the Morningstar data in the censored regressions. Utilized stocks are a subset of CRSP's NYSE and AMEX stock database. 1992 utilized data have 24 months of uninterrupted returns for the period 1/90–12/91. This holds similarly for 1991 data. Size represents market equity. Beta is computed as the sum of the slopes in the regression of the return on the current and prior month's value-weighted market return. Return Variance is the total variance of monthly returns for the prior 2–5 years, depending on availability. News Stories is the number of news stories found in the UMI data file of major newspapers for the two-year period before 1991 and the one-year period before 1992. Age is the number of months since the stock was listed prior to the observation year. Monthly Volume/Shares Outstanding is a proxy for the security's liquidity.

	% Return Variance	Beta	Price ($)	Age (Months)	News Stories	Monthly Vol./Shares Out.	Size ($mill.)
	Panel A: 1991 Data (2,133 Observations)						
Mean	1.49	1.08	17.9	231	31	0.039	2,190
Median	1.01	1.07	11.6	187	15	0.029	144
Maximum	44.77	4.35	459.1	780	1641	4.84	410,000
Minimum	0.03	−2.80	0.016	24	0	0	0.15
	Panel B: 1992 Data (2,221 Observations)						
Mean	1.72	1.14	21.8	231	16	0.051	2,731
Median	1.04	1.11	15.4	162	8	0.036	201
Maximum	51.27	6.25	433.5	792	844	5.257	385,000
Minimum	0.03	−0.66	0.05	24	0	0	0.49

to test whether aggregate preferences were driven by the new entrants or established funds.

The independent variables were all available as of the December before the portfolio holdings year (i.e., 12/90 and 12/91 for years 1991 and 1992, respectively). Table IV shows summary statistics for variables of interest in explaining mutual fund demand. If at least 24 months of data or the price as of December was not available, I did not use the stock. Return variances are calculated as of December of the prior year, using the past 24–60 months depending on availability. Beta is derived by summing the slopes in a regression of the stock return on the current and prior month's market return (I used the NYSE and AMEX value-weighted return for the market return). Age represents the number of months the security has been listed on the NYSE or AMEX.[2] Given the criteria for determining variance, age is at least 24 months (since we need this many months of uninterrupted returns to calculate the variance of the issue's return). News stories is determined by searching of the

[2] CRSP starts its coverage of the NYSE in 1925, which puts a maximum limit on age below the maximum age of some companies. The nonlinear relation between age and other variables suggests this will not affect the tests significantly.

Table V

Correlation Statistics on NYSE and AMEX Securities Used in Censored Regressions for 1991

Utilized stocks are a subset of CRSP's NYSE and AMEX stock database. All stocks have 24 months of uninterrupted returns for the period 1/89–12/90. Size represents market equity. Beta is computed as the sum of the slopes in the regression of the return on the current and prior month's value-weighted market return. Variance is the total variance of monthly returns for the prior 2–5 years, depending on availability, while Stdev is the square root of the Variance. News is the number of news stories found in the UMI data file of major newspapers for the two-year period prior to 1991. Age is the number of months since the stock was listed prior to 12/90. Liq. is the liquidity of the stock as measured by the monthly volume divided by shares outstanding in 12/90.

Panel A

Price Category	% of Sample	Mean Size ($millions)	Mean β	Mean Variance (% Return)
Total	100	2,833	1.13	1.6
Price≤1	5.1	21	1.52	6.0
1<Price≤2	4.3	26	1.30	3.6
2<Price≤3	3.4	41	1.20	2.9
3<Price≤5	6.5	61	1.29	2.5
5<Price≤10	16.6	174	1.15	1.6
10<Price≤15	12.6	592	1.03	1.2
15<Price	51.5	5,282	1.07	0.9

Panel B: Correlation Matrix

	log(Prc + 1)	Log(size)	Stdev	Variance	Log(Liq)	Log (News)	Log(Age)	Beta
Log(Prc + 1)	1	0.83	−0.60	−0.49	−0.03	0.27	0.29	−0.18
Log(size)	0.83	1	−0.51	−0.43	0.04	0.37	0.26	−0.11
Stdev	−0.60	−0.51	1	0.91	0.12	−0.06	−0.07	0.51
Variance	−0.49	−0.43	0.91	1	0.08	−0.10	−0.07	0.34
Log(Liq.)	−0.03	0.04	0.13	0.08	1	0.17	0.07	0.24
Log(News)	0.27	0.37	−0.06	−0.10	0.16	1	0.28	0.12
Log(Age)	0.29	0.26	−0.07	−0.07	0.07	0.28	1	−0.02
Beta	−0.18	−0.11	0.51	0.35	0.24	0.12	−0.02	1

University Microfilms International (UMI) Newspaper Abstracts database and counting the number of stories in major newspapers that contained a particular company's name. For the 1992 and 1991 portfolio data, I use the number of stories in 1991, and 1989–1990, respectively (due to the nature of the UMI data).[3] Monthly volume divided by shares outstanding proxies a stock's liquidity, as it represents the turnover of a stock share.

A discussion of the relevance of these variables is important. Table V illustrates how these variables are correlated, transformed as they are in the regressions. Panel A in Table V highlights how price is correlated with market equity, beta, and variance. Note that a substantial number of observations

[3] In the regressions I divided the 1991 number of news stories in half in order to make it consistent with the 1992 coefficients.

have very low prices and very small size. For example, in 12/90 19.3 percent of the securities with at least 24 months listing on the NYSE or AMEX had a price less than \$5, with a mean capitalization of only \$39 million (see Table V). Since most funds are restricted from owning more than 10 percent of a firm's voting shares, there are at least institutional constraints on fund demand for these small cap stocks.[4] Furthermore, given the documented existence of upwardly sloped supply curves for stocks, the ownership of a large proportion of a stock may imply a premium must be paid to put on or take off a position (see Bagwell (1992) on the existence of positively sloped supply curves for equities). This implies the small size of some firms might discourage buying by large funds who perceive significant transaction costs in buying and selling shares relative to other, larger stocks.

Panel A of Table V highlights price because it shows up so clearly in the preferences of mutual funds. Low priced stocks are known to have greater bid-ask spreads as a percentage of their average price (see Blume and Stambaugh (1983)). The greater transaction costs of these small, low-priced stocks provides reason for an institutional aversion outside the realm of risk preferences. Another reason to control for price is that low-priced stocks have a bias in their measured variance due to their proportionally higher bid-ask spreads. A stock moving from bid to ask to bid adds a great deal to the return variance of low price stocks although this is not "real" stock price variance in the sense that it is not the variance of return relevant for a trader off the floor of the exchange. For example, a stock going back and forth from its bid of 1 to its ask of 1¼ will report a substantial variance, even though the true return remained constant at zero. Specifically, the bias is $\hat{\sigma}^2 = \sigma^2 + \varphi^2/4$ where σ^2 is the true variance, $\hat{\sigma}^2$ is the measured variance, $\varphi = P_A - P_B/[(P_A + P_B)/2]$, and P_B and P_A are the bid and ask prices.[5]

Conrad and Kaul (1993) and Kothari, Shanken, and Sloan (1995) highlight the correlation between low prices and large bid-ask spreads. This implies low-priced stocks have measured variance significantly above true variance. This bias probably helps explain the monotonic increase in monthly return variance among the price categories. Note that while the mean return variance was 1.6 percent for all stocks, for those stocks with a price less than or equal to \$1 (5.1 percent of the sample) the mean return variance was considerably higher (6.0 percent). Since the relationship between the price level of a stock and the bid-ask spread as a percentage of price is nonlinear and there are quite a few stocks with very low prices in this sample, I use the log of the price as an explanatory variable in the regressions.

[4] For example, under section 8 of the Investment Company Act (ICA) of 1940, a mutual fund must make a registration statement where it defines its investment policy. This statement declares whether the fund is "diversified" or "nondiversified," and if the fund declares itself "diversified" it can not invest in more than 10 percent of any firm's voting securities. Also, section 12(e) of the ICA of 1940 forbids the ownership of more than 10 percent of many financial service companies regardless of the fund's registration statement.

[5] See Stoll and Whaley (1983) or Conrad and Kaul (1993) for an exposition of the bid-ask spread bias.

Other variables of interest include monthly return variance and beta. I do not include residual variance because of its high correlation with total variance, so much so that residual variance added no information to total variance in unreported tests. Volume as a percentage of shares outstanding (i.e., liquidity) serves as a proxy for transaction costs. News stories and age can both proxy for aggregate information or the cost of information on particular stocks. For example, older stocks have a more established reputation and thus less estimation uncertainty of the riskiness of the stock, and the prevalence of news stories could proxy for the recent amount of useful information generated by a firm.

Panel B of Table V shows the correlation matrix for the independent variables. The collinearity of the explanatory variables shows up in the standard errors of the multivariate regressions and does not bias coefficients. Polynomial expansions of the explanatory variables were tried for each in order to pare down the list of potential influences on mutual fund holdings, but all of the remaining explanatory variables retained significance. The greatest correlations exist between price, size, and standard deviation of monthly returns. This points to the importance of separating the independent influence of these variables on the variable to be explained, mutual fund ownership.

II. Correlations Between Mutual Fund Ownership and Various Stock Characteristics as revealed by Sorting

Tables VI and VII introduce the relationships in the context of a simple sorting procedure. Sorting is done in two ways: by mutual fund ownership and by the stock characteristic. In this way, we can gauge the economic significance of the relationships. The high degree of correlation between some of the stock characteristics, especially price, return variance, and size, suggest that sorting may obscure a more complex, multivariate relation. Since inference from these sorts is therefore limited, and also because the 1991 and 1992 data are very similar, I only present the 1992 data.

Table VI sorts the stocks by their various characteristics and lists the mean value of each of the independent variables with the mean ownership percentage for that quintile. Ownership is increasing in all of the explanatory variables, and the range of ownership among the various quintiles range from 2.2 to 8.6 percent, with most mean ownership figures rising approximately two- to threefold from their smallest to largest quintile. Price, liquidity (volume as a percent of shares outstanding), and news stories show ownership rising consistently among the quintiles, while for size, age, beta, and variance the variables show a flattening relationship once the fourth quintile is reached. Ownership displays a hump-shaped relation with stock return variance, with the highest ownership percentage in the middle quintile.

Note the smallest size-sorted quintile of stocks has a mean market equity of only $12 million, and these stocks have a mean ownership of only 2.3 percent compared to an average of 5.7 percent in the entire sample. On a similar note, the lowest price quintile has a mean price of only $2, with a mean percent

Table VI

1992 Mean Mutual Fund Ownership and Security Characteristics, Sorted by Security Characteristics

Each quintile is rank ordered by the specific stock characteristic, with mean values and mutual fund ownership within each quintile listed in rows. All stocks have at least 24 months of uninterrupted returns prior to January 1992. Size represents market equity. Beta is computed as the sum of the slopes in the regression of the return on the current and prior month's value-weighted market return. Monthly Ret. Variance is the total variance of monthly returns for the prior 2–5 years, depending on availability. News is the number of news stories found in the UMI data file of major newspapers during 1991. Age is the number of months since the stock was listed prior to the observation year. % Mut. Ownership represents the mean mutual fund percentage ownership for the different quintiles of stocks, as each sorting is done by stock characteristic and thus each quintile will have a different mean ownership percentage. The stock characteristics are taken as of 12/91, while the ownership percentage is estimated using 1992 Morningstar portfolio data.

Panel A

Monthly Ret. Var.	% Mut. Ownership	Size ($mill)	% Mut. Ownership	Price	% Mut. Ownership	Age (months)	% Mut. Ownership
0.26	3.7	12	2.3	2.0	2.2	40	4.4
0.66	6.2	66	4.4	8.0	4.2	69	4.0
1.06	7.9	222	6.4	15.3	6.3	172	5.4
1.72	6.7	835	7.9	26.8	7.6	307	7.2
4.94	3.9	12,582	7.4	57.0	8.1	607	7.5

Panel B

News Stories	% Mut. Ownership	Beta	% Mut. Ownership	Volume/ # Shares	% Mut. Ownership
0	3.5	0.3	3.7	0.0078	3.0
3.8	3.7	0.8	5.4	0.0217	4.4
7.8	6.4	1.1	6.2	0.0359	5.3
11.9	7.0	1.4	6.8	0.0557	7.1
55.5	7.7	2.1	6.4	0.1321	8.6

Panel C: Sample with Price >$5 only

Monthly Ret. Var.	% Mut. Ownership
0.21	3.2
0.56	6.2
0.87	7.7
1.29	8.3
2.69	7.7

ownership of 2.2 percent. The multivariate tests suggest that price, surprisingly, is the driving force behind the aversion to extremely low-priced and small size firms. Also, monthly return variance rises sharply in the highest variance-sorted quintile while ownership falls substantially. Below I address the measurement problem created by low price in this finding.

Table VII

1992 Mean Mutual Fund Ownership and Security Characteristics, Sorted by Mutual Fund Ownership

Each quintile is rank ordered by mutual fund ownership, and summary statistics of the stocks within each quintile are listed in rows. All stocks have at least 24 months of uninterrupted returns prior to January 1992. Size represents market equity. Beta is computed as the sum of the slopes in the regression of the return on the current and prior month's value-weighted market return. Monthly Ret. Var. is the total variance of monthly returns for the prior 2–5 years, depending on availability. News is the number of news stories found in the UMI data file of major newspapers during 1991. Age is the number of months since the stock was listed prior to the observation year. The stock characteristics are taken as of 12/91, while the ownership percentage is estimated using 1992 Morningstar portfolio data. All quintiles are sorted by % Mutual Fund Ownership, which is listed in the leftmost column.

Panel A

% Mut. Ownership	Monthly Ret. Var.	Size ($mill.)	Price ($)	Age (months)	News Stories	Beta	Volume/ # Shares	Obs.
5.70	1.73	2732	21.8	208	15.5	1.14	0.0507	2221
0.01	1.98	494	12.0	95	6.1	0.91	0.0382	444
0.10	2.69	6204	15.1	161	9.7	1.24	0.0523	444
3.80	1.44	3178	24.2	270	21.4	1.10	0.0412	445
7.60	1.15	2297	29.4	259	19.9	1.17	0.0211	444
15.90	1.37	1486	28.3	253	20.3	1.29	0.0707	444

Panel B: Sample with Price >$5 only

% Mut. Ownership	Monthly Ret. Var. (%)
6.60	1.12
0.04	0.90
2.00	1.24
5.10	1.04
8.80	1.12
16.90	1.31

Table VII sorts stocks into quintiles based on their mutual fund ownership percentages, which have ownership means ranging from a low of 0.01 percent to a high of 15.9 percent. Inference from this table is similar to Table VI, which was sorted by the various stock characteristics. In general the explanatory variables appear increasing in ownership. While liquidity shows an irregular ownership correlation with peak ownership for the most liquid quintile, the other variables show a distinct flattening in the third or fourth quintile. Beta is generally increasing, but does not have a monotonic relation with ownership in this sort. It will be seen below that beta is not significant in multivariate tests that include variance and standard deviation, suggesting systematic variance is not a significant factor in explaining mutual fund ownership.

Both Table VI and Table VII include a separate variance sort that excludes stocks with prices lower than $5 in December 1991. This is done to show more accurately the relation between variance and ownership by abstracting from the bias on measured variance created by low price. This restriction trims the sample by 20 percent. As opposed to the total sample, variance shows a distinct positive correlation that is consistent through the quintiles in both sorts. This implies that the true relation between variance and ownership is positive, and highlights the importance of controlling for omitted variables in analysis of mutual fund preferences.

III. Estimation Technique

Although some of the funds hold short positions, the great majority do not. In fact for no stock does the entire sample have a net short position. Institutional constraints on short positions imply that a censored model would be most appropriate.[6] Thus the equation I estimated is a standard censored model of the form[7]

$$y^* = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \qquad \begin{array}{ll} y = 0 & \text{if} \quad y^* \leq 0 \\ y = y^* & \text{if} \quad y^* \geq 0 \\ y = \ln(1 + own_i) & \end{array} \tag{2}$$

where the explanatory variables x_i are various security characteristics, specifically: return variance, price level, months since listing on the exchange (age), liquidity, news stories, market equity, prior year's return, and beta. To linearize the variables, I take the natural logarithms of the dependent variable, mutual fund ownership, as well as the independent variables market equity, liquidity, age, and news stories.

For simple linear regressions heteroskedasticity leads to inefficient estimates, but with nonlinear regressions the estimates can be biased as well. Unreported analysis suggests that the errors indeed are not homoskedastic, especially across capitalization. For this reason, I estimated the parameters in equation (2) using Powell's censored least absolute deviations (CLAD) estimator (Powell (1984, 1986)). The CLAD estimation technique relies on symmetry conditions imposed by a censored model, and in this way is semi-parametric and does not require assumptions of homoskedasticity or normality. The CLAD parameter estimates are unbiased, consistent, and asymptotically normally distributed, with only a minimal restriction on the distribution of the error term (it must be median unbiased). Both years had to be estimated independently since both methods assume the samples are identically and independently distributed.

I also examine the changes in fund holdings as a function of the changes in the independent variables. Consistent with the CLAD technique, I use Honoré's fixed-effects Tobit estimator (Honoré (1992), Honoré and Campbell

[6] Most funds have charters that explicitly forbid short positions. Furthermore, all funds are restricted in the use of short sale proceeds, which must be placed in low-risk money market securities.

[7] See Maddala (1989) p. 149 for an exposition of censored regression models.

(1993)). This estimation technique eliminates between-firm variation so that only intertemporal changes in a security's characteristics and in the levels of mutual fund ownership for that security are used to influence the parameter estimates. In this way it is more like prior studies of mutual fund preferences since it looks at flows of mutual fund holdings. The fixed-effects Tobit estimation technique also relies on symmetry conditions imposed by a censored model, is semi-parametric, and does not require assumptions of homoskedasticity or normality. The fixed-effects Tobit technique has similar properties to the CLAD, but adds the assumption of homoskedasticity across time periods. It should be noted a priori that if variation in characteristics across firms is much greater than the variation in characteristics within firms over the sample period, there will be a significant loss of power. The results seem to support this conclusion, and thus the CLAD is the main estimation technique.

To get standard errors I use a bootstrapping methodology (see Efron (1979)). For both the CLAD and fixed-effects Tobit, estimating standard errors uses the method of kernels, a procedure that involves a subjective judgment concerning the appropriate bandwidth (see Silverman (1986)). The use of the bootstrap avoids this subjectivity while providing an unbiased and consistent estimate of the standard errors. The bootstrap method consists of the following steps. First, parameter estimates are made using a Monte Carlo technique that assigns a $1/n$ probability for each of the n sample observations. This involves creating a bootstrapped sample from the original sample by randomly pulling n observations with replacement. This procedure is repeated 500 times, creating 500 parameter estimates. The standard deviations of this sample of 500 parameter estimates is used to generate the standard errors of the parameter estimates.

IV. Estimation Results

There are five tables of regression results, each using either different samples or different estimation techniques. Table VIII lists CLAD estimates for the entire sample. Table IX shows the estimates from an OLS estimation of the truncated sample that only includes stocks with positive levels of mutual fund ownership. Table X is a fixed-effects Tobit estimation, which uses changes in the independent variables to explain the changes in mutual fund holdings. Table XI shows the estimates from a CLAD estimation on the subsample which includes only those stocks that had book equity data, and therefore represents a potentially biased subsample of Table VIII. Finally, Table XII contains parameter estimates for a CLAD estimation on the various subsectors of the mutual funds.

Table VIII reports the main cross-sectional results for years 1991 and 1992.[8] We see from the first regression in Table VIII that mutual funds prefer stocks with higher prices, age since listing on the exchange, volatility, liquidity, news stories, and smaller size. Table VIII also shows a polynomial expansion of

[8] Bo Honoré graciously supplied the Gauss program used to calculate the CLAD and fixed-effects Tobit estimates.

variance and size because these expansions generated significant coefficients. Unreported regressions using polynomial expansions of other variables, such as liquidity, did not diminish the significance of any other explanatory variables and were insignificant themselves and were left unreported.

Table VIII
Censored Least Absolute Deviations Model of Mutual Fund Demand (*t*-Statistics in Parentheses)

Explanatory variables were taken from the December prior to the year portfolio holdings were estimated. ln(Prc) is the natural logarithm of (1 + price). ln(Age) is the natural logarithm of the age of the security, where age is the number of months the security has been listed on the NYSE and AMEX. ln(News) is the natural logarithm of (1 + news stories), where news stories is the number of news stories found in the UMI data file of major newspapers. ln(Liq) is the natural logarithm of (liquidity + 1), where liquidity is defined as the monthly trading volume divided by the shares outstanding. Variance and Stdev are the variance and standard deviation of monthly returns for the prior 24–60 months, depending on availability. Beta is the sum of coefficients on the current and lagged value-weighted NYSE and AMEX market return for individual stocks (Dimson betas), using the past 24–60 months depending on availability. ln(Size) is the natural logarithm of market equity the December prior to year when mutual fund holdings are determined. Panel A contains 1991 portfolio holdings, while Panel B contains 1992 portfolio holdings. Censored Least Absolute Deviation standard errors were bootstrapped, using 500 random draws. Columns represent separate estimations, intercept omitted.

$$\text{Dependent Variable: } \ln(1 + \text{own}_i) \quad \text{where } own_i = \sum_{m=1}^{M} \frac{\text{shares owned by fund m at time } t_m}{\text{shares outstanding at time } t_m}$$

and M is the total number of mutual funds in the specified sample. own_i represents percent ownership of stock outstanding held by the entire sample of mutual funds (e.g., 14.5 percent ownership by the mutual fund sector is input as 145).

Panel A: 1991 Portfolio Holdings							
ln(Prc)/100	7.64*	73.98*	82.32*	80.73*	75.26*		66.65*
	(9.64)	(9.52)	(10.53)	(8.22)	(7.81)		(12.43)
Stdev	2.45*	23.28*	24.16*	2.89*	23.34*		
	(9.28)	(10.06)	(9.62)	(3.66)	(10.56)		
Variance	−9.55*	−69.79*	−73.09*		69.86*		
	(−7.19)	(−7.33)	(−6.31)		(−7.41)		
ln(Age)/100	40.1*	25.8*	25.5*	26.3*	26.07*		
	(6.65)	(6.52)	(7.05)	(5.06)	(6.87)		
ln(News)/100	3.72*	1.30*	8.72*	14.67*	12.59*		
	(3.29)	(3.84)	(2.43)	(4.45)	(3.28)		
ln(Liq)/100	54.6*	26.4*	31.7*	43.7*	26.21*		
	(4.76)	(4.65)	(6.34)	(8.40)	(4.45)		
ln(Size)/100	8.65*	35.66*	−2.59	−5.71*	36.36*		
	(3.93)	(4.03)	(−1.87)	(−3.59)	(4.05)		
$[\ln(\text{Size})]^2/1000$	−3.24*	−15.11*		−15.32*			
	(−4.49)	(−4.64)		(−4.60)			
Beta	−5.09						
	(0.42)						
Ret(−1)/100					−3.00	36.8*	−243.2*
					(−0.40)	(7.2)	(−2.7)

Table VIII—*Continued*

Panel B: 1992 Portfolio Holdings

ln(Prc)/100	8.64*	74.62*	78.41*	80.28*	80.11*		66.45*
	(8.73)	(8.68)	(10.07)	(8.69)	(8.73)		(15.00)
Stdev	2.64*	21.51*	22.24*	1.55*	19.17*		
	(8.56)	(8.96)	(10.55)	(2.46)	(7.01)		
Variance	−10.59*	−63.43*	−68.41*		−50.31*		
	(−6.23)	(−5.91)	(−7.48)		(−4.38)		
ln(Age)/100	36.9*	26.8*	26.3*	22.2*	26.74*		
	(7.45)	(7.06)	(7.12)	(4.76)	(6.80)		
ln(News)/100	5.85*	14.10*	12.54*	22.8*	15.88*		
	(2.54)	(2.59)	(2.27)	(4.82)	(2.94)		
ln(Liq)/100	67.7*	33.1*	39.2*	50.7*	33.34*		
	(4.92)	(5.02)	(7.26)	(10.92)	(5.31)		
ln(Size)/100	8.60*	32.72*	−1.65	−4.86*	37.82*		
	(3.86)	(3.82)	(−1.01)	(−2.61)	(4.97)		
$[\ln(\text{Size})]^2/1000$	−15.11*	−12.98*	−14.96*				
	(−4.33)	(−4.24)	(−5.71)				
Beta	−4.71						
	(−0.99)						
Ret(−1)/100					−5.51*	9.5*	39.99
					(−2.15)	(3.1)	(1.14)

* Significant at the 5 percent level.

Table IX lists coefficients from a simple OLS regression on the variables after eliminating observations with zero mutual fund ownership. This is done to see if the results are driven primarily by stocks that are totally neglected by mutual funds, and also to show how much of ownership the explanatory variables explain. The adjusted R^2s are 0.27 and 0.32 for years 1991 and 1992, respectively. The general set of preferences is consistent with Table VIII, indicating that the neglected stocks are not driving any particular results.

Table X shows the coefficients in a fixed-effects Tobit model. This technique necessitates that we exclude age since the difference in age for each stock will be the same. It appears the loss in power from using within-firm changes in the explanatory variables is significant, as all of the previously relevant variables are now all insignificantly different than zero. This highlights how the use of flows of institutional ownership (as in Lakonishok, Shleifer, and Vishny (1991, 1992), and Grinblatt, Titman and Wermers (1995)) may be a weak method for generating tests of institutional preferences in certain circumstances. The weak power of these tests, however, may be due to my method of estimating ownership using various dates from the entire year, rather than the use of flows itself.

As various hypotheses are addressed by more than one table, below I address how tables VIII–XII affect these hypotheses.

Table IX

OLS Model of Mutual Fund Demand for those Securities with Positive Mutual Fund Ownership (*t*-Statistics in Parentheses)

Independent variables were taken from the December prior to the dependent variable year. ln(Prc) is the natural logarithm of (1 + price). ln(age) is the natural logarithm of the age of the security, where age is the number of months the security has been listed on the NYSE and AMEX. ln(News) is the natural logarithm of (1 + news stories), where news stories is the number of news stories found in UMI data file of major newspapers. ln(Liq) is the natural logarithm of (liquidity + 1), where liquidity is defined as the monthly trading volume divided by the shares outstanding. Variance and Stdev are the variance and standard deviation of monthly returns for the prior 24–60 months, depending on availability. ln(Size) is the natural logarithm of market equity the December prior to year when mutual fund holdings are determined. Ret(−1) is the prior year's net return. Ordinary least squares was performed on the sample of stocks with positive mutual fund ownership. Standard errors used in generating the *t*-statistics are heteroskedasticity consistent (White) standard errors. Intercept omitted.

$$\text{Dependent Variable: } \ln(1 + \text{own}_i) \quad \text{where } own_i = \sum_{m=1}^{M} \frac{\text{shares owned by fund m at time } t_m}{\text{shares outstanding at time } t_m}$$

and M is the total number of mutual funds in the specified sample. own_i represents percent ownership of stock outstanding held by entire sample mutual funds.

	1991 (1,777 Observations)	1992 (1,839 Observations)
Stdev	8.71*	2.71
	(3.82)	(1.82)
Variance	−27.7*	−9.93*
	(−4.15)	(−3.27)
ln(1 + Prc)/100	49.5*	46.2*
	(8.01)	(8.00)
ln(Liq)/100	37.2*	36.9*
	(10.6)	(10.8)
ln(News)/100	6.11*	7.86*
	(2.54)	(3.23)
ln(Age)/100	6.55*	5.26*
	(3.08)	(2.45)
ln(Size)/100	41.5*	61.1*
	(3.91)	(5.20)
$[\ln(\text{Size})]^2/1000$	−20.9*	−26.2*
	(−5.31)	(−6.19)
Ret(−1)/100	28.2*	−3.14
	(2.75)	(−0.64)
Adjusted R^2	0.27	0.32

* Significant at the 5% level.

A. Proxies for Transaction Costs

The coefficient on price is consistently highly significant in all the censored regressions. The use of the log of price suggests that this may more accurately be termed a strong aversion for stocks with low prices. Table VIII also shows that mutual funds prefer stocks with higher liquidity. Both coefficients suggest

Table X

Fixed-Effects Tobit Model of Mutual Fund Demand
(*t*-Statistics in Parentheses)

Independent variables were taken from the December prior to the dependent variable year. ln(Prc) is the natural logarithm of (1 + price). ln(News) is the natural logarithm of (1 + news stories), where new stories is the number of news stories found in the UMI data file of major newspapers. ln(Liq) is the natural logarithm of (liquidity + 1), where liquidity is defined as the monthly trading volume divided by the shares outstanding. Variance and Stdev are the variance and standard deviation of monthly returns for the prior 24–60 months, depending on availability. ln(Size) is the natural logarithm of market equity the December prior to the year when natural fund holdings are determined. Ret(−1) is the prior year's net return. Standard errors were bootstrapped using 500 random draws. Intercept omitted.

Dependent Variable: $\ln(1 + \text{own}_i)$ where $own_t = \sum_{m=1}^{M} \frac{\text{shares owned by fund m at time } t_m}{\text{shares outstanding at time } t_m}$

and M is the total number of mutual funds in the specified sample. own_t represents percent ownership of stock outstanding held by entire sample mutual funds.

Ret(−1)/1000	−6.64
	(−0.04)
ln(Prc)/100	6.08
	(0.84)
Stdev	−1.15
	(−1.76)
Variance	1.29
	(1.19)
ln(News)/100	0.75
	(0.59)
ln(Liq)/100	−0.66
	(−0.33)
ln(Size)/100	3.07
	(0.42)
$[\ln(\text{Size})]^2/1000$	−4.93
	(−0.18)

* Significant at the 5 percent level.

that mutual funds are sensitive to transaction costs, and are not acting on a comparative advantage in purchasing "costly" stocks.

B. Information Proxies

Table VIII shows that two proxies for information, age and news stories, are significantly positively correlated with mutual fund ownership.[9] Firms with low profiles might require greater search costs in highlighting them as securities desirable within a portfolio. They may also have greater uncertainty in

[9] Specifically, these newspapers included the *Atlanta Constitution, Boston Globe, Chicago Tribune, Christian Science Monitor, Los Angeles Times, New York Times, USA Today, Wall Street Journal,* and the *Washington Post.*

Table XI

Censored Least Absolute Deviations Model of Mutual Fund Demand in a Sample with Book/Market Equity (t-Statistics in Parentheses)

Independent variables were taken from the December prior to the dependent variable year. ln(Prc) is the natural logarithm of (1 + price). ln(Age) is the natural logarithm of the age of the security, where age is the number of months the security has been listed on the NYSE and AMEX. ln(News) is the natural logarithm of (1 + news stories), where news stories is the number of news stories found in the UMI data file of major newspapers. ln(Liq) is the natural logarithm of (liquidity + 1), where liquidity is defined as the monthly trading volume divided by the shares outstanding. Variance and Stdev are the variance and standard deviation of monthly returns for the prior 24–60 months, depending on availability. ln(Size) is the natural logarithm of market equity the December prior to the year when mutual fund holdings are determined. Ret(−1) is the prior year's net return. BE/ME is the natural logarithm of the book equity divided by the market equity. Standard errors were bootstrapped using 500 random draws. Intercept omitted.

$$\text{Dependent Variable: } \ln(1 + \text{own}_i) \quad \text{where } own_i = \sum_{m=1}^{M} \frac{\text{shares owned by fund m at time } t_m}{\text{shares outstanding at time } t_m}$$

and M is the total number of mutual funds in the specified sample. own_i represents percent ownership of stock outstanding held by entire sample mutual funds.

	1991	1992	1991	1992
ln(Prc)/100	46.64*	41.90*	44.44*	38.37*
	(6.70)	(7.67)	(7.06)	(5.78)
Stdev	6.29*	4.31*	6.51*	3.82*
	(3.43)	(2.68)	(3.94)	(2.08)
Variance	−20.02*	−16.03	−22.22*	−14.71
	(−2.53)	(−2.40)	(−3.08)	(−1.87)
ln(Age)/100	0.28	0.13	0.31	0.11
	(0.99)	(0.68)	(1.10)	(0.57)
ln(News)/100	0.01	4.00	−0.19	5.49
	(0.01)	(1.31)	(−0.07)	(1.66)
ln(Liq)/100	38.38*	40.17*	39.60*	42.36*
	(6.80)	(9.34)	(7.23)	(10.45)
ln(Size)/100	45.78*	38.99*	43.48*	35.99*
	(5.11)	(5.20)	(5.26)	(5.33)
[ln(Size)]^2/1000	−18.61*	−15.37*	−17.88*	−14.61*
	(−5.79)	(−5.88)	(−5.98)	(−6.13)
Ret(−1)/100	5.87	−0.90	2.40	−2.59
	(0.88)	(−0.42)	(0.45)	(−1.14)
BE/ME	0.037	0.057*		
	(1.55)	(3.03)		

* Significant at the 5 percent level.

the estimation of their risk. Funds show a significant preference towards stocks that are discussed in newspapers, and also stocks that have been listed for a significant time period. The use of the log of news stories and the log of age suggests that this may be more accurately characterized as an aversion to stocks that are very rarely in the news and are newly listed.

C. Volatility

The coefficient on standard deviation is positive while the coefficient on variance is negative, indicating a nonlinear preference towards volatile stocks. This finding is consistent throughout the subsectors of mutual funds and stocks. Looking at Table VIII we see that the coefficient on beta is insignificant in a regression including standard deviation and variance. This implies that mutual fund ownership is concave in idiosyncratic variance, or perhaps more accurately, mutual funds are averse to very low idiosyncratic volatility. Table VII shows that in a regression excluding variance, the coefficient on standard deviation is positive, indicating that the general preference is for higher volatility.

D. Trend Following and Herding

In Table VIII prior return is insignificant in 1991 and negative and significant in 1992 in regressions with other variables. Using OLS on the sample of those stocks with positive ownership, the coefficient is positive and significant in 1991, but insignificant in 1992 (Table IX). It is important to use caution in interpreting prior return coefficients given the estimation technique. As the mean turnover of stocks for the funds in this sample is 85 percent in 1992 (see Table I), this necessarily implies a considerable stability of mutual fund ownership between 1991 and 1992. Given the inertia in the variation of mutual fund ownership across stocks, a stock characteristic that is short term such as prior annual return may give confusing results and reflect concentration in a particular sector (e.g., technology stocks) that has correlated returns. Therefore, it is more appropriate to use prior return in the fixed-effects Tobit estimation where we look at flows, instead of stocks of ownership, where changes in ownership are influenced by changes in the independent variables.

While positive-feedback trading (i.e., trend following) may be difficult to detect using the fixed-effects estimation procedure (for reasons relating to low power mentioned above), these tests highlight an important omitted variables problem. Table VIII shows that when prior return is used alone in explaining mutual fund ownership, its coefficient is positive and significant in both years: mutual funds tend to hold proportionally more of stocks that have recently appreciated, ceteris paribus. This correlation could easily be interpreted as a preference for stocks that have risen in price, or alternatively an aversion to stocks that have fallen in price. When we include other variables, however, we see that this inference is mistaken. Using price and prior return (Table VIII), the coefficient on prior return is significant and negative in 1991 and positive and insignificant in 1992. This implies that the aversion to low-priced stocks could be driving the simple correlation between prior return and mutual fund demand in prior studies. Low prices are correlated with low prior returns, and the strong aversion to low-priced stocks by the mutual fund sector will show up in a simple correlation between ownership and prior return, providing the unwary econometrician with evidence of positive-feedback investing (i.e., buying stocks that have gone up in value).

E. Subsectors of Mutual Funds

Table XII breaks down the basic set of parameters by mutual fund subsets to determine if some coefficients are driven by the size of the fund or the tenure of the portfolio manager. Large (Small) funds are those with assets greater

Table XII

Censored Least Absolute Deviations Model of Mutual Fund Demand (*t*-Statistics in Parentheses)

Each column refers to a different subset of mutual fund portfolio holdings. Large funds have assets greater than the median, while Small refers to funds that have assets less than the median. Old refers to funds run by managers with greater than median tenure (i.e., 3 years as reported in Table I), while Young is the subset of funds run by managers with less than or equal to median tenured managers. Growth contains only Growth funds while Gr/Inc. contains only Growth and Income funds. Small Cap refers to those funds targeting small-sized firms, while Ex.Sm.Cap is the universe of Morningstar funds excluding the small cap funds. Explanatory variables were taken from the December prior to the year portfolio holdings were estimated. ln(Prc) is the natural logarithm of (1 − price). ln(Age) is the natural logarithm of the age of the security, where age is the number of months the security has been listed on the NYSE and AMEX. ln(News) is the natural logarithm of (1 + news stories), where news stories is the number of news stories found on the UMI data file of major newspapers. ln(Liq) is the natural logarithm of (liquidity + 1), where liquidity is defined as the monthly trading volume divided by the shares outstanding. Variance and Stdev are the variance and standard deviation of monthly returns for the prior 24–60 months, depending on availability. ln(Size) is the natural logarithm of market equity the December prior to the year when mutual fund holdings are determined. BE/ME is the natural logarithm of the book equity divided by the market equity. Panel A contains 1991 portfolio holdings, while Panel B contains 1992 portfolio holdings. Standard errors were bootstrapped using 500 random draws. Intercept omitted.

$$\text{Dependent Variable: } \ln(1 + \text{own}_i) \quad \text{where } own_i = \sum_{m=1}^{M} \frac{\text{shares owned by fund m at time } t_m}{\text{shares outstanding at time } t_m}$$

and M is the total number of mutual funds in the specified sample. own_i represents percent ownership of stock outstanding held by the entire sample mutual funds.

	Small	Large	Young	Old	Growth	Gr/Inc.	Ex.Sm.Cap	Small Cap
Panel A: 1991 Portfolio Holdings								
ln(Prc)/100	45.17*	80.63*	68.86*	81.36*	77.03*	43.39*	89.23*	120.65*
	(5.39)	(10.21)	(7.29)	(8.92)	(9.07)	(3.60)	(6.79)	(10.40)
Stdev	8.24*	25.22*	23.77*	21.66*	39.31*	14.32*	23.61*	33.47*
	(3.00)	(9.96)	(6.73)	(10.51)	(8.99)	(3.21)	(5.17)	(5.47)
Variance	−25.96	−76.24*	−64.89*	−66.00*	−149.40*	−66.10*	−77.18*	−116.11*
	(−1.77)	(−6.63)	(−4.65)	(−6.34)	(−6.59)	(−2.75)	(−3.05)	(−4.65)
ln(Age)/100	8.20*	27.20*	12.20*	25.90*	5.00*	28.70*	15.20*	17.40*
	(2.90)	(6.64)	(3.02)	(6.51)	(1.07)	(5.99)	(3.48)	(3.44)
ln(News)/100	0.54	10.65*	2.00	11.84*	−0.58	3.94	3.55	3.88
	(0.22)	(2.89)	(0.48)	(2.85)	(−0.19)	(0.90)	(0.87)	(0.74)
ln(Liq)/100	33.45*	29.71*	67.54*	23.29*	86.23*	79.63*	68.25*	−8.49
	(11.46)	(6.04)	(11.93)	(4.08)	(14.41)	(11.70)	(12.62)	(−1.59)
ln(Size)/100	−0.32	−2.52	7.92*	−5.12*	11.23*	16.97*	12.59*	−38.56*
	(−0.26)	(−1.75)	(4.33)	(−2.89)	(7.05)	(8.72)	(5.25)	(−15.97)

Table XII—*Continued*

Panel B: 1992 Portfolio Holdings								
ln(Prc)/100	33.16*	81.89*	58.26*	74.26*	73.69*	41.24	73.58*	118.37*
	(4.88)	(9.31)	(10.20)	(8.99)	(6.70)	(4.34)	(8.84)	(11.79)
Stdev	13.27*	23.50*	21.40*	21.30*	37.29*	19.76*	23.67*	28.82
	(7.34)	(9.32)	(13.06)	(8.25)	(11.22)	(4.98)	(6.58)	(5.81)
Variance	−48.65*	−73.54*	−76.24*	−66.00*	−143.37*	−101.05*	−85.40*	−91.62*
	(−5.39)	(−6.44)	(−10.14)	(−5.45)	(−8.75)	(−4.79)	(−4.57)	(−4.57)
ln(Age)/100	8.60*	25.70*	13.90*	28.50*	08.70*	29.30*	15.10*	8.40*
	(2.81)	(6.89)	(5.22)	(6.07)	(2.21)	(7.08)	(3.92)	(1.52)
ln(News)/100	1.65	12.09*	0.95	14.90*	−5.47	3.63	−1.36	11.4
	(0.45)	(2.16)	(0.21)	(3.20)	(−0.90)	(0.76)	(−0.29)	(1.83)
ln(Liq)/100	43.62*	38.19*	57.82*	35.26*	98.71*	81.03*	77.65*	8.25
	(13.92)	(5.83)	(13.79)	(6.70)	(14.99)	(11.32)	(15.40)	(1.44)
ln(Size)/100	4.13*	−1.30	8.46*	−2.14	15.24*	17.85*	16.01*	−34.99*
	(3.17)	(−0.74)	(6.58)	(−1.31)	(8.44)	(9.69)	(8.85)	(−16.82)

* Significant at the 5 percent level.

(smaller) than the median fund. Old and Young pertain to manager tenure, which is reported (though with considerable error) in the Morningstar data. Old (Young) funds are those with tenure longer (shorter) then the median fund. Table XII also breaks the mutual fund sector down by objective. The CLAD estimation technique has difficulty converging for panels with few positive observations, and thus funds with less prominent objectives are not examined since they hold so few of the total number of available stocks. I break down the mutual funds into growth (about 35 percent of the fund sample), growth-income (about 28 percent), small cap (about 6 percent) and all funds but the small cap funds (about 94 percent).

Perhaps the major surprise is the lack of differentiation. The coefficients on the various stock characteristics are very stable across these subsectors, suggesting little diversity in preferences. Predictably, the only striking deviation is by the small cap subsector, which shows an affinity towards smaller firms as opposed to the general preference towards larger stocks. This preference is further manifested in the coefficient on liquidity, in that the small cap subsector shows a much weaker preference towards higher liquidity. The general preference by the old tenure managers for smaller stocks reflects the tendency for the small cap sector funds to have been in existence for a while. Excluding the small cap sector, funds clearly prefer larger sized firms.

F. Book-to-Market Equity

Table XI shows the influence of the book-to-market equity ratio (BE/ME) on mutual fund preferences. The subset of the sample that includes data on book equity is approximately 75 percent of the original sample. This sample has been documented as biased towards older and more successful firms by Kothari, Shanken, and Sloan (1995) and Breen and Korajczyk (1993). We can see this by performing the same estimation as in Table VIII, without the BE/ME ratio. This results in a loss of significance for the coefficients on news and age. Since the effect of news and age on ownership is nonlinear (i.e., mainly an aversion to very new and lightly mentioned stocks), a sample biased towards older and more successful firms would be consistent with the observed reduction in the significance of these coefficients.

Lakonishok, Shleifer, and Vishny (1994) document that institutions seem to prefer "glamour stocks," which are analogous to growth stocks that have low BE/ME ratios. They hypothesize that previous success of the stocks helps institutions justify their portfolios to investors, and also that trend following may also bias institutions toward these types of stocks. This holding preference may be relevant to Fama and French's (1992) finding that high BE/ME stocks appear to produce high risk adjusted returns (using a two factor model that includes size and the firm's market beta). Specifically, perhaps the high demand for low BE/ME stocks causes them to be overpriced, which induces lower returns in the future. The coefficient on BE/ME is positive in both years, however, and significant in 1992. Thus there appears to be some evidence of a mutual fund preference towards the opposite strategy (i.e., value stocks) which targets high BE/ME ratios. Prior return is also included in this table because it is related to the hypothesis that funds concentrate on low BE/ME stocks (i.e., the trend-following hypothesis), and its coefficient is not significant. The sample bias created by using only those stocks with book equity data in COMPUSTAT, however, makes any conclusion premature.

V. Conclusions and Implications

This paper documents for the first time several specific security characteristics that mutual funds prefer. Using a comprehensive data set on open-end mutual fund portfolio holdings for the years 1991 and 1992, I calculate the relative holdings of shares by the open-end mutual fund sector as a whole, and by specific fund objective. Although it is only 2 years of data, it is free of the survivorship bias since it was not constructed retroactively by Morningstar. The semiparametric regression results indicate that share price level, volatility, liquidity, news stories, age, size, and idiosyncratic volatility are all significant in explaining aggregate mutual fund holdings of individual securities.

The breakdown of mutual funds by sector, age, and size suggests that mutual fund preferences are relatively consistent across all sectors except for firm size. Small-cap funds show, not surprisingly, a preference toward small cap stocks which seems most concentrated in those stocks with relatively low liquidity.

The aversion to low variance stocks may be related to a strategy outlined by Peter Lynch in his popular book *One Up on Wall Street* (1992). Lynch suggests buying potential "ten-baggers," or stocks that have the potential of rising ten-fold. If one has the ability to identify undervalued stocks, one might as well choose the stocks that can outperform the market by a significant amount rather than stocks that can at most offer only a modest overperformance. Most funds, distinguishing themselves from passively managed index funds, justify their active management and corresponding higher fees by their ability to spot undervalued stocks. Interest from these managers may only follow after a threshold of volatility is reached.

The sensitivity of mutual funds to proxies for transaction costs (e.g., price and liquidity) and information flows (e.g., new stories and months since initial listing) has an interesting implication. It suggests that mutual funds behave less like "insiders" and more like the general public. Intuitively, individual investors are at a comparative disadvantage in trading illiquid stocks and stocks that are not frequently in the news, but this would imply that someone else is holding disproportionate amounts of these illiquid, obscure stocks. That someone else is not the open-end mutual fund sector.

A separate question raised by frictions is their relevance to the expected returns of these assets. Even though groups may have comparative advantages buying different types of assets, arbitrage should eliminate any differential in risk-adjusted returns. Tests of market efficiency usually set forth a model of asset pricing, and then add an extraneous asset characteristic to see if it has statistical significance in explaining expected returns. By highlighting security characteristics that generate unrepresentative ownership, we highlight variables useful in further tests of market efficiency. In another vein, we may alter models of equilibrium asset returns in the presence of certain frictions. For example Merton (1987) derives a model where expected returns are decreasing with the degree of investor recognition. The idea is that greater investor recognition leads to more investors who demand the stock, raising its equilibrium price and lowering its expected return. Proxies for investor recognition have included listing on a major exchange (see Barry and Brown (1982) or Kadlec and McConnell (1994)), size (Merton (1987)), or institutional holdings (Arbel, Carvel, and Strebel (1982)). I also examine a new characteristic directly related to investor recognition—media coverage—and find it significant in explaining mutual fund preferences.

Related to the Merton model, the finding that funds other than the small cap sector prefer larger firms suggests a straightforward resolution to the small size effect seen in cross-sectional asset pricing (see Banz (1981)). If previous to the growth of the small cap investment objective all institutions had a strong preference towards larger firms, this perhaps drove up larger firm stock prices relative to the smaller firm stock prices. This would explain why the ignored smaller stocks enjoyed positive abnormal returns until the growth of the small cap sector in the eighties, which, interestingly, is about the time when small cap stocks stopped outperforming the market.

These findings are also relevant in examining a potential agency problem within mutual fund companies and their investors. Sirri and Tufano (1993) document that mutual fund performance affects net capital inflows mainly for the highest ranking funds. Top funds (ranked by non-risk-adjusted portfolio returns) receive large net inflows of new money, yet funds that perform average or poorly do not lose or gain many assets. Since most funds receive compensation as a fixed percentage of assets managed, this creates a nonlinear payoff function to fund managers similar to that of a call option. As with call options, where higher variance of the underlying asset implies greater value to the option owner, this payoff suggests fund managers and fund complexes (which manage a group of funds) have an incentive to engage in volatile strategies. In other words, given the preference of investors toward the top-performing funds, a revealed preference by fund managers towards volatile strategies would seem rational. Evidence that funds are choosing volatile portfolios in order to take advantage of this fact comes from several sources. Related to this hypothesis, this paper finds that in aggregate, funds have a preference towards securities with higher volatility, and this does not appear beta related. The nonlinear relation between idiosyncratic volatility and ownership, and the lack of significance of beta has several implications. First, mutual funds seem more poised to take advantage of relative rather than absolute performance, as funds are not simply buying high beta stocks. Second, mutual funds do not seem to prefer the most highly volatile stocks as much as they seem to avoid the lowest volatility stocks. If fund managers are taking advantage of the option-like payoff to their relative performance, a subtle constraint must exist that prevents it from being pursued outright, since otherwise they would strictly prefer the most volatile securities. As most funds diversify somewhat (no fund holds only one stock), clearly the riskiest strategies which would maximize an option-like payoff are constrained by an acknowledgment of risk. It seems general investor wariness of risk does put a constraint on fund behavior.

Several extensions are promising. First, to re-evaluate prior tests of trend following by controlling for price. Second, to gauge the significance of the explanatory variables in explaining expected returns. This can be in tests of market efficiency or tests of Merton's theory of an investor recognition factor. There are also corporate finance implications. The price of a stock is somewhat a choice variable of the firm, and since it clearly seems to influence mutual fund ownership (and probably other institutions) it is unclear why firms would choose a low or high price. Finally, the avoidance of newly-listed stocks may be relevant to the studies of IPOs, specifically, what is it about the newly listed stocks that discourages mutual fund ownership?

[10] Chevalier and Ellison (1995), Goetzman and Peles (1993), and Ippolito (1992) also document a primarily convex relation between mutual fund inflows and risk-adjusted returns.

REFERENCES

Arbel, Avner, Steven Carvell, and Paul Strebel, 1983, Giraffes, institutions and neglected firms, *Financial Analysts Journal* May/June, 2–6.

Bagwell, Laurie Simon, 1992, Dutch Auction Repurchases: An analysis of shareholder heterogeneity, *Journal of Finance* 47, 71–106.

Banerjee, A. V., 1992, A simple model of herd behavior, *Quarterly Journal of Economics* 107, 797–818.

Banz, Rolf, 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 1–18.

Barry, Christopher, and Steven Brown, 1984, Differential information and the small firm effect, *Journal of Financial Economics* 13, 407–422.

Blume, Marshal, and Robert F. Stambaugh, 1983, Biases in computed returns: An application to the size effect, *Journal of Financial Economics* 12, 387–404.

Breen, William J., and Robert Korajczyk, 1993, On selection biases in book-to-market based tests of asset pricing models, Working paper 167, Northwestern University.

Chevalier, Judith, and Glenn Ellison, 1995, Risk taking by mutual funds as a response to incentives, Working paper, University of Chicago.

Conrad, Jennifer, and Gautam Kaul, 1993, Long-term market overaction or biases in computed returns? *Journal of Finance* 48, 39–64.

Efron, B. 1979, Bootstrap methods: Another look at the jackknife, *Journal of the American Statistical Association* 76, 312–319.

Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–466.

Goetzmann, William N., and Nadav Peles, 1993, A simple framework for nonparametric specification testing, Discussion paper 1662, Harvard Institute for Economic Research.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review*, Forthcoming.

Honoré, Bo, 1992, Trimmed LAD and least squares estimation of truncated and censored regression models with fixed effects, *Econometrica* 60, 533–566.

Honoré, Bo, and Jeffrey R. Campbell, 1993, Median unbiasedness of estimators of panel data censored regression models, *Econometric Theory* 9, 499–503.

Ippolito, Richard, 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law and Economics* 35, 45–70.

Kadlec, Gregory B., and John J. McConnell, 1994, The effect of market segmentation and illiquidity on asset prices: Evidence from exchange listings, *Journal of Finance* 49, 611–636.

Kang, Jun-Koo, and René M. Stulz, 1994, Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan, Working Paper, The Ohio State University.

Kraus A., and Hans Stoll, 1972, Parallel trading by institutional investors, *Journal of Financial and Quantitative Analysis* 7, 2107–2138.

Kothari, S. P., Jay Shanken, and Richard Sloan, 1995, Another look at Expected Stock Returns, *Journal of Finance* 50, 185–224.

Lakonishok, Josef, Andrei Shleifer, Richard Thaler, and Robert Vishny, 1991, Window dressing by pension fund managers, *American Economic Review* 81, 227–231.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23–44.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

Lynch, Peter, 1989, *One Up On Wall Street* (Penguin Books, New York).

Maddala, G. S., 1984 Limited dependent and qualitative variables in econometrics, (Cambridge University Press, Cambridge).

Merton, Robert C., 1987, Presidential address: A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483–510.

Powell, J., 1984, Least absolute deviation estimation for the censored regression model, *Journal of Econometrics* 25, 303–325.

Powell, J., 1986, Symmetrically trimmed least squared estimation for Tobit models, *Econometrica* 54, 1435–1460.

Sharfstein, David, and Jeremy Stein, 1990, Herd behavior and investment, *American Economic Review* 80, 465–479.

Silverman, B. W., 1986, *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, London).

Stoll, H. R., and R. E. Whaley, 1983, Transaction costs and the small firm effect, *Journal of Financial Economics* 12, 57–79.

Sirri, Erik R., and Peter Tufano, 1993, Competition and change in the mutual fund industry, in Samuel L. Hayes III, Ed.: *Financial Services: Perspectives and Challenges* (Harvard Business School Press, Boston, Mass.).

Welch, Ivo, 1992, Sequential sales, learning and cascades, *Journal of Finance* 47, 695–732.