Mark Grinblatt
Sheridan Titman
University of California, Los Angeles

Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings*

This article employs the 1975–84 quarterly holdings of a sample of mutual funds to construct an estimate of their gross returns. This sample, which is not subject to survivorship bias, is used in conjunction with a sample that contains the actual (net) returns of the mutual funds. In addition to allowing us to estimate the bias in measured performance that is due to the survival requirement and to estimate total transaction costs, the sample is used to test for the existence of abnormal performance. The tests indicate that the risk-adjusted gross returns of some funds were significantly positive.

The evaluation of portfolio performance has generated a great deal of interest in academic circles. A variety of evaluation techniques have been proposed and implemented, but to date, there is no consensus about the ability of professional portfolio managers to earn abnormal returns. Previous studies of mutual fund performance have examined the actual returns realized by investors and, with few exceptions, found either negative performance or no performance for the average mutual fund. This is not surprising from an economic perspective: if mutual fund managers have superior investment talent, they may be able to capture the rents from their talent in the form of higher fees or perquisites obtained through higher expenses.[1] In this case, we can expect to observe abnormal performance only by examining gross returns, which do not have transaction costs, fees, or other expenses subtracted from them.[2]

* We wish to thank Jim Brandon, Pierre Hillion, and Erik Sirri for excellent research assistance, as well as Doug Diamond, an anonymous referee, and seminar participants at Columbia University for helpful comments on earlier drafts.

1. Admati and Pfleiderer (1989, in press) develop a model that provides conditions under which this is true.

2. Copeland and Mayers (1982) examined the perfor-

(*Journal of Business*, 1989, vol. 62, no. 3)

0021-9398/89/6203-0002$01.50

The present study compares the abnormal returns of active and passive investment strategies—both with and without transaction costs, fees, and expenses. It employs data that contain the quarterly equity holdings of a large sample of mutual funds that existed for part or all of the 1975–84 period. These holdings can be used to construct hypothetical mutual fund returns—returns that an investor would have realized by purchasing (without transaction costs) the portfolios reported in the funds' quarterly reports. Such returns do not have management fees, expenses, brokerage commissions, or other transaction costs subtracted from them.

The difference between the abnormal performance of the hypothetical returns and the actual mutual fund returns provides an estimate of the average magnitude of mutual fund transaction costs. In addition, the particular sample used here does not require that the included funds survive for the entire sample period. This is important since previous studies of mutual fund performance examined samples that consisted only of surviving funds.[3] A comparison of our sample without the survival requirement with one that includes only surviving mutual funds allows us to gauge the bias in studies with samples consisting only of surviving funds.

Transaction costs, survivorship bias, and abnormal performance are separately examined for subsamples of the funds, grouped by investment objective and net asset value. The more important findings include:

i) Survivorship bias is relatively small, on the order of .5% per year or less. The bias is somewhat larger for the smaller funds.

ii) Transaction costs, on average, are relatively large, on the order of 2.5% per year, and are inversely related to fund size.

iii) The abnormal performance of the funds, based on gross returns, is inversely related to fund size, but, since transaction costs are inversely related to fund size, the actual net returns are unrelated to the net-asset values of the funds.

iv) On average, actual returns do not display positive abnormal performance for any category of fund. However, the gross returns of both growth and aggressive growth funds are significantly positive on average, even after adjusting for risk with a benchmark that does not have size biases, dividend-yield biases, or beta-related biases. There is evi-

---

mance of five hypothetical portfolios formed every 6 months on the basis of securities ratings by the Value Line Investment Advisory Service. In addition, Jensen (1968, 1969) and Mains (1977) calculated gross returns by adding a rough estimate of transaction costs, fees, and expenses to observable net returns. However, we are unaware of studies that have used the portfolio holdings of actual funds to construct gross returns.

3. See, e.g., Jensen (1968, 1969); McDonald (1974); Mains (1977); Lehmann and Modest (1987); and Grinblatt and Titman (1988).

dence that this measured performance is at least partly generated by active management of the funds.

I. Methodology

This article examines whether or not there are mutual fund managers who have superior stock selection abilities that generate abnormal returns. Abnormal returns are defined here as the intercepts from excess return regressions calculated with a benchmark that is mean-variance efficient from the perspective of an uninformed observer. With such a benchmark, passive investment strategies generate zero abnormal performance, and positive abnormal performance can only be generated by an active portfolio manager who buys and sells securities in response to superior information. Although the validity of this measure of abnormal performance, sometimes known as "the Jensen Measure," has been debated in the literature, our past work suggests that it is appropriate for this sample.[4]

The Jensen Measures of the mutual funds are calculated with four sets of benchmark portfolios: the monthly rebalanced equally weighted portfolio of all CRSP (New York and American Stock Exchange) securities, the CRSP value-weighted index, 10 factor portfolios created with factor-analytic procedures developed in Lehmann and Modest (1988),[5] and the eight-portfolio benchmark, formed on the basis of firm size, dividend yield, and past returns developed in Grinblatt and Titman (1988).

The appropriateness of these benchmarks for performance evaluation was analyzed in Grinblatt and Titman (1988). The eight-portfolio benchmark (denoted as P8) appeared to be the most appropriate for performance evaluation since the intercepts of 109 passive portfolios, constructed on the basis of securities characteristics and industry groupings, were closest to zero with this benchmark.[6] The other three

4. Criticisms of the Jensen Measure by Roll (1978), Jensen (1972), and Dybvig and Ross (1985), among others, are based on the sensitivity of the Jensen Measure to the choice of a benchmark portfolio and to timing ability. Our earlier work, Grinblatt and Titman (1988), indicated that for the sample analyzed here the benchmark issue is relevant but timing-related biases are not significant. See Grinblatt and Titman (1989, in press) for a discussion of the theoretical foundations of this analysis.

5. To construct these factor returns, the input portfolio weights for the entire 10-year period were derived from a 10-factor maximum-likelihood factor analysis of 750 securities over the 1978–82 period. The portfolios contain slightly fewer than 750 securities in the 1975–77 and 1983–84 periods since some of the securities from the middle period did not exist in the earlier and later periods. Although this method of forming factor portfolios can potentially create survivorship bias, (unreported) comparisons with the equally weighted index suggest that this bias is not large.

6. See Grinblatt and Titman (1987) for further discussion of mean-variance efficiency with multiple portfolio benchmarks.

benchmarks—the value-weighted index (VW), the equally weighted index (EW), and the 10 Lehmann-Modest factor portfolios (F10)—displayed size, dividend-yield, and beta-related pricing errors. In particular, funds that invest in large firms (which includes most funds) tend to exhibit negative performance with the EW and F10 benchmarks. In addition, because of the relatively poor performance of the value-weighted index in the 1975–84 time period, funds that invest in any but the very largest firms tend to exhibit positive performance with this benchmark. For these reasons, the latter three benchmarks are primarily used for comparison purposes.

II. The Data

Mutual fund data, consisting of two data sets, each with observations from December 31, 1974, to December 31, 1984, were obtained from CDA Investment Technologies, Inc., of Silver Springs, Maryland. The first data set, which was analyzed in Grinblatt and Titman (1988), contains cash-distribution adjusted monthly returns for those funds that existed on December 31, 1984. These returns are net of transaction costs, fees, and expenses but are gross of any sales charge for the load funds. The second data set contains the compositions of the equity portion of the funds' portfolios, as reported quarterly to the Securities and Exchange Commission (SEC). In contrast to the first data set, this second set is not subject to survivorship bias and is more complete, in the sense that it contains the compositions of all equity mutual funds that were registered with the SEC in a given quarter.[7] We were also able to obtain from CDA a list of investment goals for a subsample of the funds, as reported December 31, 1974.

Stock returns were obtained from the Center for Research in Security Prices (CRSP) daily returns file. The daily returns were compounded to create monthly returns. These were used to form hypothetical monthly mutual fund returns, constructed from the portfolio weights of the mutual funds, and to form returns for the four sets of benchmark portfolios described in the previous section. Excess returns were computed for funds and securities by subtracting the return of a 1-month treasury bill from the return of the relevant investment.

III. Hypothetical Portfolio Returns

The hypothetical mutual fund returns were computed as the return of the portfolio of NYSE- and AMEX-listed equities reported in the funds' quarterly reports. As mentioned earlier, information concerning the fixed income and cash positions were not available in the quarterly

7. These samples do not include bond, money market, or preferred stock funds.

holdings data set. Equity positions in over-the-counter securities were also ignored, although this is unlikely to have much of an effect on the funds' hypothetical returns in this time period because such positions were small.

The portfolio weights were multiplied by the monthly excess returns of securities on the CRSP tape and summed to construct sets of hypothetical monthly excess returns for each of the mutual funds. The holdings were rebalanced monthly.

IV. Transaction Costs, Survivorship Bias, and Average Performance

The top half of table 1 describes the average excess returns and Jensen Measures of five samples of returns of which the first two have no survivorship bias and the first three have no transaction costs subtracted from them:

i) the hypothetical returns of an equally weighted portfolio of all funds in the quarterly holdings data set;

ii) the hypothetical returns of an equally weighted portfolio of 274 funds that existed on December 31, 1974, for which we were able to obtain initial investment objectives and initial quarterly holdings;

iii) the hypothetical returns of an equally weighted portfolio of 157 mutual funds that are available in both the quarterly holdings data set and the data set of actual monthly returns, for which we were able to obtain initial investment objectives;

iv) the actual returns of an equally weighted portfolio of the 157 funds in sample (iii); and

v) the actual returns of an equally weighted portfolio of the 279 funds in the data set of actual monthly returns.[8]

The Jensen Measures in the table are calculated with the (monthly rebalanced) equally weighted (EW) and value-weighted (VW) stock indices of all CRSP securities, the Lehmann-Modest (1988) 10-factor benchmark (F10), and the eight-portfolio benchmark (P8) used in Grinblatt and Titman (1988). The reported t-statistics in the top half of the table are calculated from a regression that uses the time series of the returns of an equally weighted portfolio of the funds.[9]

8. The sample of hypothetical funds with complete returns contained fewer funds than the corresponding sample of actual returns. This is because some of the funds did not report their holdings to the SEC in their early years of existence, and, in some instances, it was impossible to match returns in the holdings data set in different years because of name changes. (The vendor matched names for the data set that contained actual returns.)

9. If portfolio (abnormal) returns are serially independent, normally distributed, and homoscedastic, then t-statistics derived from a time series of portfolio (abnormal) returns provide valid inferences about the means of average (abnormal) returns. For further discussion, see Grinblatt and Titman (1988).

TABLE 1 **Actual or Hypothetical Monthly Returns in Excess of the Risk-free Rate, Betas, and Jensen Measures for Five Samples of Mutual Funds**

	All Funds on Quarterly Holdings Tape (Average No. of Funds = 578.1), Hypothetical[a]	274 Funds on Quarterly Holdings Tape as of January 1, 1975 (Average No. of Funds = 224), Hypothetical[a]	157 Funds on Quarterly Holdings Tape as of January 1, 1975, and Monthly Returns Tape as of December 31, 1984		279 Funds on Monthly Returns Tape, Actual[b]
			Hypothetical[a]	Actual[b]	
Average excess returns	.0083	.0085	.0087	.0063	.0067
	(.0045)[c]	(.0046)	(.0046)	(.0039)	(.0039)
Jensen Measures:					
EW index	−.0025	−.0023	−.0022	−.0031	−.0028
	(−1.19)[c]	(−1.07)	(−.99)	(−1.65)	(−1.59)
VW index	.0016	.0017	.0019	.0005	.0009
	(2.35)*	(2.30)*	(2.46)*	(.61)	(1.07)
10 factor	−.0020	−.0020	−.0019	−.0034	−.0033
	(−2.59)**	(−2.29)*	(−2.08)*	(−3.53)**	(−3.56)**
P8 portfolio	.0009	.0012	.0015	−.0005	−.0004
	(1.57)	(2.09)*	(2.53)**	(−.76)	(−.65)

Betas:					
EW index	.70	.71	.71	.61	.62
	(.032)[c]	(.034)	(.035)	(.029)	(.027)
VW index	1.08	1.11	1.12	.94	.95
	(.015)	(.017)	(.017)	(.017)	(.018)
	Survivorship Bias: 157 Hypothetical − 274 Hypothetical		Transaction Costs: 157 Hypothetical − 157 Actual		
EW index	.0001		.0009		
	(1.17)[c]		(1.65)		
VW index	.0002		.0014		
	(1.37)		(4.04)**		
10 factor	.0001		.0015		
	(1.25)		(4.47)**		
P8 portfolio	.0003		.0020		
	(2.64)**		(4.73)**		

[a] Data are for an equally weighted portfolio of funds. Returns were constructed by multiplying portfolio weights by CRSP securities returns and adding. Hence, these data have no expenses, fees, or transaction costs subtracted from them.

[b] Returns are net of transaction costs, management fees, and other expenses, but not sales charges.

[c] *t*-statistics are in parentheses for the Jensen Measures, SEs are in parentheses for the betas and the excess returns.

* Significant at .05 level.

** Significant at .01 level.

A. *Transaction Costs*

Transaction cost estimates are presented in the bottom half of table 1. They indicate that the total transaction costs of the funds range between 1% and 2.5% per year, on average, depending on the benchmark. These estimates, derived by differencing the Jensen Measures of the hypothetical and actual return samples of 157 mutual funds, are all statistically different from zero, except for the estimate with the equally weighted index. The *t*-statistics are computed from a regression that uses the time series of differences between the returns of the two equally weighted portfolios of funds. The largest transaction cost estimate corresponds to the P8 benchmark. A perusal of Weisenberger (1975–85) indicates that the direct expenses of these funds average approximately 1% per year, which suggests that an additional 1.5% of transaction costs are generated from trading costs and additional costs due to the liquidity-motivated cash positions of the funds. (The differences between the betas of the actual and hypothetical returns is evidence of a large cash or other nonequity position for these funds.)

The difference in the compositions of the actual and hypothetical portfolios can generate the observed differences in the transaction cost estimates for the different benchmarks. This is because the risk-free position of a fund, which has zero performance regardless of the choice of benchmark, mitigages the benchmark errors associated with the equity portion of that fund's portfolio. In particular, the estimates of transaction costs with the F10 and EW benchmarks are biased downward since the equity portions of most mutual fund portfolios exhibit negative performance with these benchmarks (primarily due to a firm-size effect). This, of course, does not explain why estimated transaction costs are lower with the VW benchmark (for which the equity positions of funds exhibit upwardly biased performance) than with the P8 benchmark. However, the difference between the estimates of transaction costs with these two sets of benchmarks is small and probably can be attributed to sampling error.

An additional bias in the estimates of transaction costs arises because the portfolio weights that generate the hypothetical returns are only updated quarterly. To the extent that mutual funds realize superior performance from their more frequent updating, the hypothetical returns underestimate the true gross returns. Since the actual returns are not subject to this bias, the effective transaction costs of the funds are also underestimated.

An opposite bias might result from "window dressing"; that is, selling poorly performing stocks so that they do not show up on the fund's quarterly report. Stocks that are purchased and sold within a quarter never show up in the hypothetical returns. Hence, if funds sell stocks that perform poorly shortly after they are purchased, hypothetical re-

turns and transactions costs will be overestimated. However, if this bias were large enough to alter our conclusions, the turnover rates of most funds would be larger than they actually are and would be more highly correlated with performance measured from hypothetical returns than from actual returns. In fact, turnover rates have a higher correlation with performance from actual returns.

B. Survivorship Bias

The estimates of survivorship bias are computed by taking the differences between the Jensen Measures of the sample of hypothetical returns for the 274 funds, which are not subject to survivorship bias, and the sample of 157 hypothetical returns, which are subject to survivorship bias. These estimates indicate that the positive bias in performance estimates for samples that exclude nonsurviving funds is fairly small, on average between .1% and .4% per year, depending on the benchmark. With the exception of the estimates with the P8 benchmark, they are statistically insignificant.

C. Average Performance

The average performance of the actual and hypothetical returns differ substantially across the different benchmarks. The hypothetical returns exhibit negative performance with the EW and F10 benchmarks (although the former is not reliably different from zero) and positive performance with the VW and P8 benchmarks. Moreover, as we will later see in tables 2–4, the negative average performance with the EW and F10 benchmarks holds for almost all subsamples grouped by investment objectives. Since it is unlikely that fund managers would use superior information to generate negative performance, the negative performance of the hypothetical returns is probably indicative of the inefficiency of the EW and F10 benchmarks. In particular, the negative performance may be due to the size, beta, and dividend-related biases induced by these benchmarks.

By contrast, positive performance is not necessarily indicative of benchmark inefficiency since it can be generated with superior information. An examination of performance with the P8 benchmark, which eliminates the biases of the other benchmarks, suggests that, on average, the sample of hypothetical returns without survivorship bias had slightly positive performance in the 1975–84 period. This performance, however, is less than 1.5% per year and does not exceed the transaction costs, fees, and expenses of the funds.

V. Investment Objectives and Performance

Tables 2–4 report the performance measures for subsamples, based on investment objective, for three of the five samples, ii, iii, and iv, re-

TABLE 2 157 Surviving Funds—Actual Returns

	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
No. of funds	45	10	46	37	13	3	3
Jensen Measures:[a]							
EW	−.0033	−.0027	−.0036	−.0031	−.0009	.0004	−.0038
	(−1.46)	(−1.60)	(−1.71)	(−1.74)	(−.64)	(.12)	(−1.67)
VW	.0010	−.0004	−.0001	−.0000	.0017	.0030	.0020
	(.69)	(−.44)	(−.10)	(−.10)	(1.51)	(.98)	(.86)
F10	−.0046	−.0021	−.0039	−.0026	−.0010	−.0002	−.0056
	(−3.19)**	(−2.16)*	(−3.94)**	(−3.57)**	(−1.04)	(−.08)	(−3.42)**
P8	.0003	−.0030	.0002	−.0009	−.0025	−.0002	−.0005
	(.35)	(−3.05)**	(.30)	(−1.46)	(−2.46)*	(−.06)	(−.38)
F-statistics:[b]							
P8-*F*1 statistic	1.66*	1.68	1.96**	1.02	1.32	.84	.52
	(.030)	(.096)	(.006)	(.462)	(.214)	(.475)	(.668)
P8-*F*2 statistic	1.65*	.99	2.01**	.89	.70	.95	.78
	(.031)	(.55)	(.005)	(.649)	(.745)	(.392)	(.460)

NOTE.—Tables 2–5 show Jensen Measures of actual and hypothetical returns in excess of the risk-free rate for a sample of 157 mutual funds that existed from 1975–84, and for the hypothetical returns of a sample of 274 funds that exhibits no survivorship bias, broken down by investment objective.

[a] Jensen Measure of actual returns of an equally weighted portfolio of funds with a given investment objective on 12-31-74. The 120 monthly returns used for the calculation are based on the redemption values available to an investor of the funds. The returns are net of transaction costs, management fees, and other expenses, but not sales charges. *t*-statistics are in parentheses.

[b] *F*2-statistic tests whether all funds in a category have equal Jensen Measures. Significance level below in parentheses. *F*1-statistic tests whether all funds in a category have Jensen Measures equal to zero.

* Significant at .05 level.

** Significant at .01 level.

TABLE 3 157 Surviving Funds—Hypothetical Returns

	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
No. of funds	45	10	46	37	13	3	3
Jensen Measures:[a]							
EW	−.0023	−.0022	−.0028	−.0024	.0005	−.0008	−.0014
	(−.88)	(−1.07)	(−1.45)	(−1.12)	(.27)	(−.25)	(−.58)
VW	.0025	.0010	.0011	.0011	.0040	.0028	.0057
	(1.70)	(1.47)	(1.22)	(2.18)*	(3.67)**	(.96)	(2.24)*
F10	−.0030	−.0005	−.0026	−.0013	.0012	.0010	−.0021
	(−2.13)*	(−.71)	(−2.68)**	(−1.79)	(1.41)	(.50)	(−1.33)
P8	.0028	−.0004	.0019	.0006	.0003	−.0008	.0032
	(3.05)**	(−.39)	(2.45)*	(.85)	(.36)	(−.25)	(1.90)
F-statistics:[b]							
P8-*F*1 statistic	1.71*	.53	1.79*	.81	.48	.28	1.80
	(.023)	(.865)	(.014)	(.756)	(.934)	(.841)	(.151)
P8-*F*2 statistic	1.73*	.53	1.67*	.84	.52	.40	.94
	(.021)	(.853)	(.028)	(.716)	(.898)	(.671)	(.394)

NOTE.—See note at table 2.

[a] Jensen Measure of an equally weighted portfolio of funds with a given investment objective. The 120 monthly returns were constructed by multiplying the fund's portfolio weights by CRSP securities returns and adding. Hence, the returns have no expenses, fees, or transaction costs subtracted from them. *t*-statistics are in parentheses.

[b] *F*2-statistic tests whether all funds in a category have equal Jensen Measures. Significance level below in parentheses. *F*1-statistic tests whether all funds in a category have Jensen Measures equal to zero.

* Significant at .05 level.

** Significant at .01 level.

TABLE 4 274 Initial Funds—Hypothetical Returns[a]

	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
Initial no. of funds	73	19	81	57	31	6	7
Jensen Measures:							
EW	−.0024	−.0024	−.0031	−.0025	.0000	−.0014	−.0029
	(−.91)	(−1.13)	(−1.24)	(−1.19)	(.03)	(−.37)	(−1.04)
VW	.0025	.0008	.0008	.0010	.0036	.0025	.0031
	(1.73)	(1.36)	(.94)	(2.01)*	(3.23)**	(.78)	(1.38)
F10	−.0031	−.0011	−.0028	−.0015	.0008	.0014	−.0032
	(−2.30)*	(−1.46)	(−3.02)**	(−1.99)*	(.97)	(.64)	(−1.81)
P8	.0027	−.0001	.0016	.0002	−.0001	−.0009	.0017
	(3.07)**	(−.18)	(2.00)*	(.29)	(−.15)	(−.21)	(.98)

NOTE.—See note at table 2.

[a] Jensen Measure of an equally weighted portfolio of funds that existed in December 1974 with the stated investment objective. The portfolio was reweighted each quarter to account for funds that left the sample or had missing data. Returns were constructed by multipling portfolio weights by CRSP securities returns and adding. Hence, these data have no expenses, fees, or transaction costs subtracted from them. *t*-statistics are given in parentheses.

* Significant at .05 level.

** Significant at .01 level.

spectively. Tables 2 and 3 report Jensen Measures for the actual and hypothetical returns of the 157 surviving funds. Across all categories, transaction costs, which are the difference between the corresponding numbers in the two tables, seem to be substantial. Of the three categories with large numbers of funds (aggressive-growth funds, growth funds, and growth-income funds), transaction costs appear to be the largest for the aggressive-growth funds. Table 5, which provides summary statistics on fund characteristics, indicates that the aggressive-growth funds have the highest average turnover, fees, and expenses and the smallest average net asset value.

Table 4 reports Jensen Measures for the 274-fund sample that is not subject to survivorship bias. These figures are very similar to those reported for the hypothetical returns in table 3, suggesting that the impact of survivorship bias is small for each classification of fund objective.

Using the P8 benchmark, the hypothetical returns of the growth and aggressive-growth funds exhibit positive performance. The aggressive-growth funds do particularly well, with abnormal performance of approximately 3% per year. This performance is statistically significant even when accounting for the multiple comparison, which arises because the abnormally high performance measure was discovered in an examination of seven different categories.[10] Hence, the evidence supports the hypothesis that superior investment talent exists within this group of funds. However, the performance of the actual returns of these funds does not significantly differ from zero, so any superior investment talent possessed by the managers of aggressive-growth funds do not offer individual investors an opportunity to achieve abnormal returns by purchasing fund shares.[11]

The F-statistics in tables 2 and 3, which test the joint hypothesis that the Jensen Measures of each of the funds with a particular investment goal are equal to zero ($F1$) or equal to each other ($F2$), provide further

10. Using the Bonferroni inequality, the probability that the largest t-statistic in a sample of seven exceeds 2.90 is less than .03. The Bonferroni inequality states that under the null hypothesis, $N\alpha$ is an upper bound on the probability that the largest t-statistic in a set of N t-statistics has a significance level of α. See Miller (1981).

11. The aggressive-growth funds exhibit relatively poor performance when measured with the equally weighted index and factor-based benchmark. This is probably due to the beta and dividend-yield biases of these benchmarks. These funds generally invest in securities with higher than average betas and lower than average dividend yields. As demonstrated in Grinblatt and Titman (1988), investments with high betas against the equally weighted index, and/or low dividend yields, tend to have low Jensen Measures in the 1975–84 time period when measured with the EW or F10 benchmarks. It should be noted that these negative Jensen Measures are not robust with respect to changes in the sample period. McDonald (1974) found that aggressive-growth funds had the highest performance measured against an equally weighted index in the 1960–69 period.

TABLE 5 **Summary Statistics on 155 Funds of the 274 Sample That Had Data Available on Turnover, Management Fee, Expense Ratio (Excluding Fee), and Net Asset Value: Average Values of Fund Characteristics by Investment Objectives as Calculated on December 31, 1974**

	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
Initial no. of funds	29	20	61	35	8	1	1
Net asset value (millions)	68.21	195.94	93.25	199.74	161.24	35.80	42.20
Expense ratio (%)[a]	.44	.29	.40	.37	.34	1.06	.54
Management fee (%)[b]	.60	.52	.56	.51	.43	.63	.75
Fund turnover (%)[c]	68.0	31.5	42.3	35.1	44.3	35.5	15.0

Correlation Matrix: $N = 155$

	Expense Ratio	Management	Log Turnover	Jensen Measures		
				Hypothetical	Actual	Difference[d]
Net asset value	−.35**	−.38**	−.22**	−.18*	−.06	−.16*
Expense ratio		.17*	.16	.16*	.05	.15
Management fee			.34**	.07	−.07	.19*
Log turnover				.22**	.24**	−.04
Jensen Measure—hypothetical					.73**	.32**
Jensen Measure—actual						−.40**

[a] Expenses less management fees as a percentage of net asset value.
[b] Stated management fees as a percentage of net asset value.
[c] Dollar purchases plus sales as a percentage of net asset value.
[d] The difference between the Jensen Measure of the hypothetical return of the fund and its actual return, which is an estimate of transaction costs.
* Significant at .05 level.
** Significant at .01 level.

evidence of the existence of superior portfolio managers.[12] More conclusive support for this interpretation is found in Section VII.

VI. Net Asset Value and Performance

A second way to categorize the sample of mutual funds is by net asset value. One might conjecture that small mutual funds have an advantage over large funds in that they can more easily purchase and sell securities without altering securities prices. However, it is also possible that small funds are more susceptible to survivorship bias and may experience higher transaction costs than larger funds because they cannot take advantage of certain economies of scale.

To examine the effect of fund size, we formed five portfolios of the mutual funds on the basis of their net asset value rankings as of the end of 1974. Table 6 reports tests on the Jensen Measures of these five portfolios, each of which equally weights funds in a quintile ranking. The F-statistics, $F1$ and $F2$, respectively, test whether all five Jensen Measures in a given row are equal to zero and whether they are equal to each other. The $F2$ statistic indicates that measured transaction costs differ across the size portfolios, while survivorship bias does not differ significantly. Note, however, that the portfolio of mutual funds with the smallest net asset values displays the highest transaction costs and the largest survivorship bias.

The mutual funds with the smallest net asset values also realized the largest performance. Using the P8 benchmark, the average performance of these funds was about 2.5% per year, which is highly significant. However, their actual returns did not offer investors an opportunity to achieve abnormal performance, most likely because of the high transaction costs of these funds.

These results could be related to those in the previous section in that aggressive-growth and growth funds constitute a large fraction of the smallest net asset value portfolio. Of the 55 funds in the smallest quintile, 21 are aggressive-growth funds, and 17 are growth funds. To further explore the possibility that the size effect may have been generated by the abnormal performance of the growth and aggressive-growth funds, we divided a sample consisting only of growth and aggressive-growth funds into five quintiles based on the size criteria. Once again, the

12. The t- and F-statistics test different things. The t-statistic tests whether the average performance of a group equals zero, while the F-statistic measures whether each of the funds has zero performance. Although the tests are nested in the sense that, if average performance is positive, then some fund has a nonzero risk-adjusted return, either of the tests could be more powerful against certain alternative hypotheses. For example, if all of the funds were nearly identical and had slightly positive performance, the t-test would be more likely to reject than the F-test. Alternatively, if all but one fund had zero performance and the other had very large performance, the F-test would be more likely to reject.

TABLE 6 **Jensen Measures (with *t*-Statistics in Parentheses) of Hypothetical and Actual Returns of Equally Weighted Portfolios of Funds Ranked by Net Asset Value in December 1974**

	Size						
	Smallest	2	3	4	Largest	$F1^a$	$F2^b$
Initial no. of funds for 274	55	55	55	55	54	...	...
Initial and ending no. of funds for 157	29	25	31	34	38	...	...
	Panel A: P8 Index						
274 (Hypothetical)	.0021	.0009	.0014	.0010	.0007	3.02*	2.09
	(3.30)**	(1.15)	(2.14)*	(1.49)	(1.04)	(.014)	(.087)
157 (Hypothetical)	.0026	.0010	.0018	.0013	.0010	4.62**	2.39
	(4.03)**	(1.19)	(2.65)**	(1.76)	(1.49)	(.000)	(.055)
157 (Actual)	.0000	−.0009	−.0006	−.0006	−.0003	.61	.75
	(.01)	(−1.07)	(−.88)	(−.81)	(−.42)	(.692)	(.563)
Survivorship bias: 157 (hypothetical) minus 274 (hypothetical)	.0006	.0002	.0004	.0003	.0003	2.60*	.187
	(2.14)*	(.40)	(1.30)*	(1.59)	(2.13)*	(.029)	(.945)
Transaction costs: 157 (hypothetical) minus 157 (actual)	.0028	.0021	.0021	.0019	.0011	6.12**	5.32**
	(4.88)**	(3.70)**	(4.39)**	(3.50)**	(2.99)**	(.000)	(.001)

	Panel B: F10 Index						
274 (Hypothetical)	−.0018	−.0029	−.0014	−.0021	−.0021	2.24	2.18
	(−1.91)	(−2.41)**	(−1.62)	(−2.40)*	(−2.57)*	(.056)	(.076)
157 (Hypothetical)	−.0015	−.0033	−.0011	−.0019	−.0019	2.58*	3.25*
	(−1.61)	(−2.47)*	(−1.26)	(−2.09)*	(−2.19)*	(.031)	(.015)
157 (Actual)	−.0034	−.0044	−.0026	−.0036	−.0029	2.87*	2.56*
	(−3.42)**	(−3.68)**	(−2.95)**	(−3.55)**	(−3.18)**	(.018)	(.043)
Survivorship bias:							
157 (hypothetical) minus							
274 (hypothetical)	.0003	−.0004	.0004	.0002	.0002	1.54	.92
	(1.41)	(−1.27)	(1.22)	(1.07)	(1.25)	(.183)	(.454)
Transaction costs:							
157 (hypothetical) minus							
157 (actual)	.0021	.0013	.0015	.0016	.0008	4.87	3.55**
	(4.23)**	(2.61)**	(3.78)**	(3.95)**	(2.89)**	(.000)	(.009)

[a] $F1$-statistic tests whether all five portfolios of funds have Jensen Measures of zero. Significance level is given in parentheses.
[b] $F2$-statistic tests whether all five portfolios of funds have Jensen Measures equal to each other. Significance level is given in parentheses.
* Significant at .05 level.
** Significant at .01 level.

smallest funds had the best performance. In the sample without survivorship bias, the average Jensen Measure of the smallest quintile of aggressive-growth and growth funds was over 3.6% with the P8 benchmark. The t-statistic (3.86) for this estimate was highly significant. The corresponding Jensen Measures for the other size-ranked portfolios of the aggressive-growth and growth funds were somewhat lower but also positive and statistically significant, suggesting that both net asset value and investment objective are determinants of abnormal performance.

VII. Benchmark Inefficiency, Passive Investment, and Performance

Our interpretation of the results in the previous two sections is likely to be controversial. One reason is that other empiricists have looked at similar measures of fund performance with capital asset pricing model (CAPM) and Arbitrage Pricing Theory (APT) benchmarks and found no evidence of superior investment talent among mutual fund managers. Moreover, as Roll (1978) has stressed, it is difficult to distinguish between investment performance and benchmark inefficiency. Hence, one could argue that the positive performance with the P8 benchmark, documented in the previous sections, is indicative of the inefficiency of that benchmark rather than true performance.

This section presents evidence that the observed performance was at least partly generated by superior active portfolio management. The analysis contrasts the hypothetical returns of the sample of mutual funds with two sets of corresponding returns. The first was generated by using the December 31, 1974, portfolio weights throughout the sample period. These can be considered passive portfolios and should thus realize zero abnormal performance with benchmarks that are mean-variance efficient. The second set of returns updates the portfolio weights annually (rather than quarterly). These returns should exhibit lower performance than the portfolios that are updated in a more timely fashion if the performance is indeed generated by superior information, but they may still realize abnormal performance with an efficient benchmark.

Tables 7 and 8 report the performance measures of hypothetical portfolios that employ the initial portfolio weights without updating. The performance of the aggressive-growth and growth funds are not statistically significant with the P8 benchmark. In tables 9 and 10, the more actively managed portfolios that are updated annually do achieve significant abnormal performance, but they are outperformed by the corresponding portfolios in tables 3 and 4 that are updated quarterly. The reported betas in these tables suggest that the risk associated with these strategies are all similar. This lends credence to our assertion that

TABLE 7 Jensen Measures—Only First Quarter Weights Used for Hypothetical Returns[a]

	All	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
No. of funds	274	73	19	81	57	31	6	7
EW	−.0023	−.0026	−.0021	−.0030	−.0021	−.0007	−.0009	−.0021
	(−1.15)	(−1.18)	(−1.03)	(−1.34)	(−1.05)	(−.41)	(−.25)	(−.96)
VW	.0014	.0016	.0010	.0006	.0013	.0027	.0028	.0029
	(2.15)*	(1.63)	(1.56)	(.72)	(1.99)*	(2.54)**	(.91)	(2.14)*
F10	−.0004	−.0012	.0001	−.0010	.0002	.0011	.0006	.0001
	(−.67)	(−1.28)	(.18)	(−1.37)	(.27)	(1.46)	(.36)	(.10)
P8	.0007	.0010	.0000	.0008	.0004	−.0001	.0011	.0023
	(.88)	(1.08)	(.05)	(.84)	(.54)	(−.07)	(.28)	(1.72)
Betas:								
EW Index	.65	.73	.58	.65	.62	.58	.63	.81
VW Index	1.02	1.12	.93	1.03	.98	.87	.97	1.20

NOTE.—Tables 7–10 show Jensen Measures of modified passive hypothetical returns in excess of the risk-free rate for samples of 274 and 157 mutual funds categorized by investment objective.

[a] Jensen Measure of an equally weighted portfolio of funds with a given investment objective (with *t*-statistics given in parentheses). The 120 monthly returns were constructed by multiplying the fund's first quarter portfolio weights by CRSP securities returns and adding. The returns based on these weights have no expenses, fees, or transaction costs subtracted from them.

* Significant at .05 level.

** Significant at .01 level.

TABLE 8 **Jensen Measures—Only First Quarter Weights Used for Hypothetical Returns[a]**

	All	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
No. of funds	157	45	10	46	37	13	3	3
EW	−.0024	−.0027	−.0020	−.0033	−.0022	−.0011	−.0001	.0015
	(−1.17)	(−1.17)	(−.98)	(−1.43)	(−1.07)	(−.58)	(−.03)	(−.79)
VW	.0014	.0016	.0012	.0004	.0014	.0022	.0034	.0073
	(2.00)*	(1.48)	(1.70)	(.43)	(1.97)*	(2.13)*	(1.06)	(4.08)**
F10	−.0004	−.0012	.0004	−.0010	.0004	.0012	.0016	.0019
	(−.53)	(−1.13)	(.66)	(−1.39)	(.61)	(1.44)	(.95)	(1.47)
P8	.0009	.0012	.0004	.0008	.0005	.0004	.0009	.0044
	(1.08)	(1.17)	(.47)	(.77)	(.61)	(.30)	(.25)	(3.02)
Betas:								
EW Index	.67	.73	.59	.66	.64	.57	.57	.88
VW Index	1.04	1.13	.95	1.04	1.01	.89	.85	1.24

NOTE.—See note at table 7.

[a] Jensen Measure of an equally weighted portfolio of funds with a given investment objective (with *t*-statistics given in parentheses). The 120 monthly returns were constructed by multiplying the fund's first quarter portfolio weights by CRSP securities returns and adding. The returns based on these weights have no expenses, fees, or transaction costs subtracted from them.

* Significant at .05 level.

** Significant at .01 level.

TABLE 9 Jensen Measures—Annually Updated Weights for Hypothetical Returns[a]

	All	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
No. of funds	274	73	19	81	57	31	6	7
EW	−.0028	−.0028	−.0023	−.0035	−.0023	−.0004	−.0016	−.0030
	(−1.32)	(−1.11)	(−1.08)	(−1.46)	(−1.12)	(−.26)	(−.46)	(−1.16)
VW	.0012	.0019	.0008	.0004	.0013	.0032	.0022	.0028
	(1.80)	(1.50)	(−1.33)	(.51)	(2.28)*	(2.71)**	(.70)	(1.38)
F10	−.0021	−.0030	−.0010	−.0029	−.0011	.0003	.0009	−.0023
	(−2.50)*	(−2.40)*	(−1.29)	(−3.17)**	(−1.60)	(.31)	(.38)	(−1.45)
P8	.0005	.0021	−.0002	.0008	.0001	−.0007	−.0012	.0017
	(.96)	(2.61)**	(−.23)	(1.19)	(.15)	(−.77)	(−.31)	(1.05)
Betas:								
EW Index	.70	.81	.59	.71	.65	.58	.64	.90
VW Index	1.10	1.24	.96	1.14	1.02	.87	.96	1.32

NOTE.—See note at table 7.

[a] Jensen Measure of an equally weighted portfolio of funds with a given investment objective (with *t*-statistics given in parentheses). The 120 monthly returns were constructed by multiplying the fund's (annual) portfolio weights by CRSP securities returns and adding. The returns based on these weights have no expenses, fees, or transaction costs subtracted from them.

* Significant at .05 level.

** Significant at .01 level.

TABLE 10 Jensen Measures—Weights Updated Yearly[a]

	All	Aggressive Growth	Balanced	Growth	Growth Income	Income	Special Purpose	Venture Capital/ Special Situation
No. of funds	157	45	10	46	37	13	3	3
EW	−.0026	−.0028	−.0020	−.0034	−.0024	−.0002	−.0004	−.0018
	(−1.18)	(−1.11)	(−.97)	(−1.41)	(−1.13)	(−.13)	(−.14)	(−.78)
VW	.0015	.0019	.0012	.0006	.0013	.0033	.0029	.0051
	(2.20)*	(1.41)	(1.69)	(.76)	(2.27)*	(2.99)**	(.99)	(2.12)*
F10	−.0020	−.0029	−.0004	−.0028	−.0011	.0007	.0014	−.0017
	(−2.18)*	(−2.24)*	(−.55)	(−2.88)**	(−1.57)	(.77)	(.66)	(−1.09)
P8	.0009	.0020	−.0004	.0011	.0002	−.0003	.0002	.0028
	(1.61)	(2.32)*	(−.50)	(1.51)	(.28)	(−.27)	(−.06)	(1.66)
Betas:								
EW Index	.71	.81	.58	.72	.65	.58	.54	1.03
VW Index	1.11	1.24	.94	1.15	1.03	.88	.81	1.44

NOTE.—See note at table 7.

[a] Jensen Measure of an equally weighted portfolio of funds with a given investment objective (with *t*-statistics in parentheses below). The 120 monthly returns were constructed by multiplying the fund's (annual) portfolio weights by CRSP securities returns and adding. The returns based on these weights have no expenses, fees, or transaction costs subtracted from them.

* Significant at .05 level.

** Significant at .01 level.

the active trading of the portfolio manager at least partially generates the observed performance results with the P8 benchmark.[13]

VIII. Conclusion

This research differs from previous studies of mutual fund performance in two important respects. First, samples of fund returns were constructed that approximate the gross returns of mutual funds by employing data on their quarterly portfolio holdings. Second, a benchmark was used that mitigates the possibility that well-known passive strategies could drive our results.

Jensen Measures employing this benchmark indicate that superior performance may in fact exist, particularly among aggressive-growth and growth funds and those funds with the smallest net asset values. It is interesting that these funds also have the highest expenses so that their actual returns, net of all expenses, do not exhibit abnormal performance. This indicates that investors cannot take advantage of the superior abilities of these portfolio managers by purchasing shares in their mutual funds.

References

Admati, Anat, and Pfleiderer, Paul. 1989, in press. Direct and indirect sale of information. *Econometrica*.

Copeland, Thomas, and Mayers, David. 1982. The value line enigma (1965–1978): A case study of performance evaluation issues. *Journal of Financial Economics* 10, no. 3 (November): 289–321.

Dybvig, Philip, and Ross, Stephen. 1985. Differential information and performance measurement using a security market line. *Journal of Finance* 40, no. 2 (June): 383–99.

Grinblatt, Mark, and Titman, Sheridan. 1987. The relation between mean-variance efficiency and arbitrage pricing. *Journal of Business* 60, no. 1 (January): 97–112.

Grinblatt, Mark, and Titman, Sheridan. 1988. The evaluation of mutual fund performance: An analysis of monthly returns. Working paper. Los Angeles: University of California.

Grinblatt, Mark, and Titman, Sheridan. 1989, in press. Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies*.

Jensen, Michael. 1968. The performance of mutual funds in the period 1945–1964. *Journal of Finance* 23, no. 2 (May): 389–416.

Jensen, Michael. 1969. Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business* 42, no. 2 (April): 167–247.

Jensen, Michael. 1972. Optimal utilization of market forecasts and the evaluation of investment portfolio performance. In G. Szego and K. Shell (eds.), *Mathematical Methods in Investment and Finance*. Amsterdam: North Holland.

Lehmann, Bruce, and Modest, David. 1987. Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance* 42, no. 2 (June): 233–65.

Lehmann, Bruce, and Modest, David. 1988. The empirical foundations of the arbitrage pricing theory. *Journal of Financial Economics* 21: (September): 213–54.

13. The positive performance of the passive returns sample with the value-weighted index provides further evidence of the inefficiency of this benchmark.

Mains, Norman. 1977. Risk, the pricing of capital assets, and evaluation of investment portfolios: Comment on Jensen. *Journal of Business* 50, no. 3 (July): 371–84.
McDonald, John. 1974. Objectives and performance of mutual funds, 1960–1969. *Journal of Financial and Quantitative Analysis* 9, no. 3 (June): 311–33.
Miller, Rupert. 1981. *Simultaneous Statistical Inference*. New York: Springer-Verlag.
Roll, Richard. 1978. Ambiguity when performance is measured by the securities market line. *Journal of Finance* 33, no. 1 (September): 1051–69.
Weisenberger, Arthur. 1975–85. *Investment Company Survey*. New York: A. Weisenberger.