The External Performance of Socially-Responsible Mutual Funds

Mario G. Reyes and Terrance Grieb

INTRODUCTION

Socially responsible investing is fast becoming an acceptable investment style, with approximately $10 billion of mutual fund assets and upwards of $200 billion of union pension funds invested in socially responsible portfolios [Middleton (1995)]. With socially responsible investing only securities of companies with desirable social attributes qualify for inclusion into the portfolio; "socially undesirable" securities are screened out.[1] However, social screening is incompatible with modern portfolio theory, which posits that utility-maximizing investors must hold some combination of the market portfolio of all risky assets and the risk-free asset. By imposing position constraints on security holdings, social screens restrict the feasible investment opportunities set, and could result in dominated portfolios which exhibit lower return for a given level of risk, or higher risk for the same return level [Levy (1978)]. Whether socially responsible investing entails an opportunity cost is an empirical question which several researchers have attempted to investigate. The empirical evidence to date is mixed.

Wood (1992) and Tepper (1993) report that socially responsible investing had underperformed market averages. Wood reports that in the UK socially responsible portfolios underperformed the Financial Times-All Share Stock Index during the period 1988-1991. Tepper, testing various social screens on US stocks, finds that socially screened portfolios would have underperformed the S&P 500 stock index by 1% on an annualized basis during the period 1984-1989. On the other hand, Litvak (1992) reports that the rate of return from a socially-screened portfolio compared favorably with the S&P 500 stock index during the period 1979-1989. Luck and Pilotte (1993) find the Domini Social Stock Index outperforming the S&P 500 by a margin of 2% per year during the subsequent period 1990-1992. Litvak and Luck and Pilotte did note that the socially screened portfolio exhibited higher volatility than the market portfolio. Recently, Diltz (1995) finds from his analysis of common stock portfolios that ". . . the market appears to reward good environmental performance, charitable giving, and an absence of nuclear and defense work, and it appears to penalize firms that provide family-related benefits such as parental leave, job sharing, and dependent care assistance" (p. 69).

For individual investors interested in socially responsible investing the most convenient medium is a socially responsible mutual fund (SRF hereafter). But do SRFs provide individual investors a viable investment alternative? The only systematic analysis of this issue is conducted by Hamilton, Jo, and Statman (1993). They employ Jensen's (1968) risk-adjusted evaluation technique to test the investment performance of seventeen SRFs during the period 1981-1990. They report that SRFs do not earn statistically significant excess returns and their performance is not statistically different from the performance of conventional mutual funds.

This paper extends the Hamilton, Jo, and Statman study.[2] We conduct cointegration tests and Jobson-Korkie (1983) significance tests of SRF performance during the period 1986-1995. Cointegration analysis is conducted in order to examine the temporal behavior of fifteen SRFs relative to their respective peer group (i.e., non-socially-

Dr. Mario G. Reyes is Associate Professor of Finance, University of Idaho. He has published in numerous journals such as *Journal of Business Ethics* and *Journal of Economics and Finance*.

Dr. Terrance Grieb is Assistant Professor of Finance, University of Idaho. His research interests include international capital markets and market efficiency.

screened and screened funds with similar investment objectives). Cointegration provides a test of whether a SRF and its peer group drift together, even though individually each return series may tend to trend up or down in a non-stationary fashion. If a SRF and its peer group are cointegrated, then they share a common underlying trend. The absence of cointegration would be indicative of the impact of social screens on the behavior of SRF portfolios; that is, the temporal behavior of a SRF is different from its peer group.

In addition, we conduct significance tests using the Sharpe (1966) measure to evaluate the external performance[3] of SRFs in relation to its peer funds. Sharpe's performance measure is more appropriate in the case of SRFs in light of empirical results reported by Martin, Keown, and Farrell (1982). Martin, Keown, and Farrell note that Sharpe's performance measure contains a lower bias, because it uses portfolio variance as the measure of risk thereby accounting for any extra-market covariation in returns.[4] Extra-market covariation could be a problem with SRFs because their portfolio holdings are concentrated as a result of their stated investment objective and social screens. Martin, Keown and Farrell report that concentrated holdings dictated by a fund's investment policies induce positive extra-market covariation in returns, which, in turn, cause beta to understate portfolio risk.[5]

The empirical results discussed in the next section show that all of the fifteen SRFs examined are not cointegrated with their peer funds. However, the Jobson-Korkie significance tests also show that the risk-adjusted performance of SRFs are not statistically significantly different from their peer groups.

EMPIRICAL ANALYSIS

Data

The data consist of monthly prices and rates of return on fifteen SRFs taken from the *Wilson Associates Capital Asset Management System (CAMS) Database* for the period January 1986 through December 1995. Price data are used for the cointegration tests because the variables observed must be non-stationary. The returns data are used in the Jobson-Korkie tests.

By beginning the sample period in January 1986, we ensure that all SRFs have the same number of observations and are evaluated during the same time period.[6] The data used in this study also include monthly prices and rates of return on four peer group indexes, namely: aggressive growth funds, balanced funds, growth funds, and growth and income funds, over the same time period. This is necessary for performance evaluation, because SRFs vary in their stated investment objectives. For example, the Pioneer Fund is listed as a growth and income fund, whereas Evergreen Fund has aggressive growth as its stated objective. Hence, comparing SRFs across fund categories would be misleading. Instead, SRFs should be evaluated relative to funds with identical investment objectives; that is, based on their external performance [Korkie (1983)]. Finally, from the *Dimensional Fund Advisors Returns Database* we collected monthly rates of return on 30-day Treasury bills.

COINTEGRATION ANALYSIS

Cointegration tests are conducted as a first step in the evaluation of SRF performance. Cointegration analysis provides a methodology for analyzing the behavior of SRFs vis-a-vis their peer group. For example, Parnassus is classified by the Investment Company Institute as a growth fund, but it may not behave like one because of its social screens. Earlier analysis has shown that capitalization-weighted socially-screened portfolios tend to have greater volatility, smaller size, less foreign income, and lower dividend yield (Hopkins (1992)). Hence, it is plausible that a SRF's temporal behavior could be different from its peer funds. Cointegration analysis provides a methodology for the examination of this question. If there is a linear relationship between two sets of non-stationary time series variables (in this case month-end prices of a SRF and its peer group index) which renders them stationary, then the variables are said to be cointegrated. This suggests that, even though a SRF or its peer group may tend to trend up or down in a non-stationary fashion, they may have a long-run equilibrium relationship.[7] This implies that social screens do not cause a SRF's temporal behavior to diverge from its peers.

Prior to conducting cointegration analysis, we tested for a unit root process in the price series of each mutual fund. Failure to reject the null hypothesis indicates a unit root process. The results, which are reported in Table 1, show that none of the fifteen tests are significant at the 1% level. This then indicates the presence of a unit root in each of the fifteen mutual funds and the four mutual fund indexes.[8] This implies that cointegration tests can be conducted between a SRF and its peer group index.

We apply the Engle-Granger (1987) two-step procedure to test for cointegration. First, we run the following cointegrating regression with a constant:

$$P_{\mathrm{Fund}\,i,t} = \gamma_0 + \gamma_1 P_{\mathrm{Peer},t} + u_{it} \qquad \textbf{(1)}$$

where $P_{Fund\ i,t}$ represents the end-of-month price level on SRF i and $P_{Peer,t}$ represents the corresponding end-of-month price level on SRF i's peer group. The error term $u_{i,t}$ is then a linear combination of $P_{Fund\ i,t}$ and $P_{Peer,t}$.

Next, we use the augmented Dickey-Fuller procedure to test for a unit root in the error term:

$$\Delta\hat{u}_{Fund\ i,t} = \delta_{Fund\ i}\hat{u}_{Fund\ i,t-1} + \sum_{k=1}^{q} \phi_k \Delta\hat{u}_{Fund\ i,t-k} + w_{Fund\ i,t} \qquad \textbf{(2)}$$

Specification tests on $\Delta\hat{u}$ in (2) are conducted using the Akaike Information Criterion (AIC) to ensure that the

regression residual is approximately white noise.[9] The null hypothesis is $\delta_i=0$. Failure to reject the null hypothesis implies that the SRF and its peer group index do not share a common underlying trend, which would suggest that the SRF does not behave like other funds in its category. This, in turn, would provide evidence of the impact of social screens on the temporal behavior of the SRF returns. On the other hand, rejection of the null hypothesis would imply a similarity in the temporal behavior of the SRF and its peer group. The augmented Dickey-Fuller (ADF) test statistics are reported in Table 1.

The cointegration tests fail to reject, at the 1% level, the null hypothesis in all of the fifteen SRFs. This suggests that for all of the SRFs in the sample, social screens cause its portfolio to behave quite differently from their respective peer groups. Whether this influences SRF performance is an issue addressed in the next section.

PERFORMANCE EVALUATION

The viability of any socially responsible mutual fund stems from its performance vis-a-vis competing conventional and other socially responsible mutual funds. To test whether a SRF has total risk-return performance equivalent to other funds, we use the Sharpe measure,

$$SH_{Fund\ i} = \frac{\overline{R}_i - \overline{r}_f}{s_i} \tag{3}$$

where $\overline{R}_i$ is the average monthly rate of return on fund i over the sample period, $\overline{r}_f$ is the contemporaneous average monthly rate of return on 30-day Treasury bill, and s_i is fund i's standard deviation of monthly return. Summary statistics and Sharpe performance measures (SH_{Fund}) for the fifteen SRFs in the sample and the peer groups (SH_{Peer}) are reported in Table 2.

The average rate of return for all fifteen SRFs was 1.086% per month, which is equivalent to a 13.8% compounded annual rate of return. The best performing SRF was Putnam OTC Emerging Growth (POEGX), an aggressive growth fund which earned an average rate of return of 1.659% per month (21.83% compounded annual return). The second best was another aggressive growth fund, Putnam Health Sciences (PHSTX), which earned a monthly average of 1.488%, or 19.39% annually. Both funds also exhibited standard deviation of over 18% annually.

Of the fifteen SRFs studied, only four funds are found to have Sharpe performance measures greater than the Sharpe measures of their respective peer groups. The outperforming SRFs are New Economy (ANEFX), Scudder Growth & Income (SCDGX), Putnam Health Sciences (PHSTX), and Putnam OTC Emerging Growth (POEGX). To test whether the investment performance of SRFs vis-a-vis their peer groups are indeed significantly different, we use the Jobson-Korkie (1983) test statistic:

$$Z - stat = \frac{\overline{R}_i s_I - \overline{R}_I s_i}{\sqrt{\theta}} \tag{4}$$

where the estimated variance is:

$$\theta = \frac{1}{T}\left[2s_i^2 s_I^2 - 2s_i s_I s_{i,I} + \frac{1}{2}\overline{R}_i^2 s_I^2 + \frac{1}{2}\overline{R}_I^2 s_i^2 - \frac{\overline{R}_i \overline{R}_I}{2s_i s_I}(s_{i,I}^2 + s_i^2 s_I^2)\right] \tag{5}$$

$\overline{R}_i$ is fund i's mean monthly return, $\overline{R}_I$ is the mean monthly return of fund i's peer group, s_i is the standard deviation of fund i, s_I is fund i's peer group standard deviation, and $s_{i,I}$ is the estimated covariance between the returns on fund i and its peer group I. The test statistic is normally distributed in large samples with zero mean and unit standard deviation under the null hypothesis H_0: $SH_{Fund\ i} = SH_{Peer}$, i.e., the Sharpe performance of SRF i and its peer group are equal. A Jobson-Korkie *Z-statistic* exceeding the critical value at 1% will result in the rejection of null hypothesis of equivalent Sharpe performance.

The Jobson-Korkie significance test statistics reported in Table 2 show that null hypothesis of equivalent performance cannot be rejected for all the fifteen SRFs in the sample. This suggests that the risk-adjusted performance of SRFs is not statistically different from their peers. This result provides additional support to Hamilton, Jo, and Statman (1993), who conclude that ". . . investors can expect to lose nothing by investing in socially responsible mutual funds . . . " (p. 66).

CONCLUSION

This study examines the investment performance of socially responsible mutual funds. Analysis of monthly returns over the period January 1986 - December 1995 shows that all funds were not cointegrated with their respective peer funds. This is indicative of the impact of social screens on the temporal behavior of SRFs; screening causes the time series behavior of a SRF to diverge from that of its peer group. Performance evaluation using the Sharpe measure shows that only four SRFs outperformed their peer funds on a risk-adjusted basis, however, Jobson-Korkie significance tests show that the risk-adjusted performance between SRFs and their peer funds are not statistically significant different. Our results provide additional indication that social responsibility characteristics are not priced in the market (see, e.g., Hamilton, Jo, and Statman (1993)).

ACKNOWLEDGEMENTS

The authors would like to thank two anonymous referees and Randy Byers for their insightful comments, Wilson Associates for providing the Capital Asset Management Systems and Database, and Dimensional Funds Advisors, Inc. for providing the Returns Database. Mario Reyes thanks the UI CBE Summer Research Grant Program for financial assistance.

NOTES

[1] Funds and investors use different social screens. Examples of qualitative social screens include environmental performance, corporate citizenship, and employee relations. Examples of exclusionary screens include military weapons, nuclear weapons, alcohol, tobacco, and gambling.

[2] We tried to collect data on the same funds in the Hamilton, Jo, and Statman study, but the Wilson Associates CAMS database did not have the return data on SFT Environmental Awareness. Moreover, the database did not have monthly return on Greenspring Fund after November 1992 and on Transamerica Capital Appreciation after September 1994. Both funds were dropped from the sample. Only fourteen of the SRFs are common to both samples. We added Pioneer Fund to increase our sample to fifteen SRFs. See Appendix A for the list of SRFs in the sample.

[3] We do not make any attempt to examine specific social criteria. The focus of this study is to determine whether these social criteria influence the temporal behavior of a SRF and its investment performance. See Korkie (1983) for a discussion on the distinction between external and internal performance.

[4] We did not adjust the estimated Sharpe measures for sample size bias shown by Miller and Gehr (1978), who proposed the following approximately unbiased Sharpe performance estimator:

$$\text{Adjusted Sharpe ratio} = \frac{\overline{R}_i - r_f}{s_i}\left(\frac{N}{N+0.75}\right)$$

where N is the number of monthly observations. With a sample size of 120 monthly observations used in this study, the sample bias is minor.

[5] For instance, in the Martin, Keown, and Farrell study, the extra-market covariation for the Pioneer Fund, one of the first SRFs to be established, was found to account for 18 percent of the fund's portfolio variance (see their Table 2, p. 25). With this much extra-market risk, beta understates Pioneer Fund's portfolio risk.

[6] January 1986 is a convenient starting date, since some of the funds were not established until 1983, 1984, or 1985.

[7] Within the cointegration literature, the term equilibrium is used to indicate an observed relationship which has, on average, been maintained by the two series for a long period (Cuthbertson, Hall, and Taylor (1992), p.132).

[8] We use the 1% level in all our analyses in order to keep the probability of a type I error (i.e., false rejection of the null hypothesis) to a reasonable level. When there are 15 pairs to compare, the probability of a type I error is 0.139 ($= 1 - (1-0.01)^{15}$). This is dramatically lower than 0.537, which is the probability of a type I error at the 5% level.

[9] To verify the robustness of the specification suggested by the AIC, we also conducted specification tests using the Schwartz Bayesian Criterion (SBC). For the initial root tests and the ADF cointegration tests, there were five cases in which the SBC suggested a slightly different specification. However, the significance of the test results are identical to those reported here.

REFERENCES

Cuthbertson, K., S.G. Hall, and M.P. Taylor (1992), *Applied Econometric Techniques* (Ann Arbor: The University of Michigan Press).

Diltz, J.D. (1995), "The Private Cost of Socially Responsible Investing," *Applied Financial Economics* 5, 69-77.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica* 55, 251-76.

Grinblatt, M. and S. Titman (1989), "Portfolio Performance Evaluation: Old Issues and New Insights," *Review of Financial Studies* 2, 393-421.

Hamilton, S., H. Jo, and M. Statman (1993), "Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds," *Financial Analysts Journal* (November/December), 62-66.

Hopkins, Jay (1992), "Money Managers No Longer Blind to Social Investments," *Pension World* (April), 10-13.

Jensen, M. (1968), "The Performance of Mutual Funds in the Period 1945-1964," *Journal of Finance* 23, 389-416.

Jobson, J.D. and B.M. Korkie (1981), "Performance Hypothesis Testing with the Sharpe and Treynor Measures," *Journal of Finance* 36, 889-908.

Korkie, B.M. (1983), "External vs. Internal Performance Evaluation," *Journal of Portfolio Management* 9 (Spring), 36-42.

Levy, H. (1978), "Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio," *American Economic Review* (September), 643-658.

Litvak, L. (1992), "Do Social Screens Harm Performance?" *Values*, US Trust Company of Boston (June), 1, 5.

Luck, C. and N. Pilotte (1993), "Domini Social Index Performance," *The Journal of Investing* (Fall), 60-62.

MacKinnon, J.G. (1991), "Critical Values for Cointegration Tests," in R.F. Engle and C.W. Granger (eds.), *Long-Run Economic Relationships: Readings in Cointegration* (Oxford University Press), Chapter 13.

Martin, J.D., A.J. Keown, Jr., and J.L. Farrell (1982), "Do Fund Objectives Affect Diversification Policies?" *Journal of Portfolio Management* 8 (Winter), pp. 19-28.

Middleton, T. (1995), "At the Crossroads," *Mutual Funds* (February), 39-45.

Miller, R.E. and A.K. Gehr (1978), "Sample Size Bias and Sharpe's Performance Measure: A Note," *Journal of Quantitative and Financial Analysis* (December), 943-946.

Sharpe, W.F. (1966), "Mutual Fund Performance," *Journal of Business* (January), 119-138.

Tepper, J.A. (1993), "Evaluating the Cost of Socially Responsible Investing," in P.D. Kinder, S.D. Lydenberg and A.L. Domini (eds.), *The Social Investment Almanac* (New York: Henry Holt and Company), 340-349.

Wood, K. (1992), "Ethical Investments in the UK-The BARRA/EIRIS Study," *BARRA Newsletter* (March/April), 5-8.

APPENDIX A

Socially-Responsible Mutual Funds in Sample: 1986.01-1995.12

Fund Name	Ticker Symbol	Fund Type
Calvert Social-Managed Growth Fund	CSIFX	Balanced
Pax World Fund	PAXWX	Balanced
Dreyfus Third Century Fund	DRTHX	Growth
New Alternatives Fund	NALFX	Growth
The New Economy Fund	ANEFX	Growth
Parnassus Fund	PARNX	Growth
Pioneer Fund	PIODX	Growth & Income
Pioneer II Fund	PIOTX	Growth & Income
Pioneer III Fund	PITHX	Growth & Income
IDS Equity Select A Fund	INVPX	Growth & Income
Scudder Growth & Income Fund	SCDGX	Growth & Income
Evergreen Fund	EVGRX	Aggressive Growth
Putnam Health Sciences A	PHSTX	Aggressive Growth
Putnam OTC Emerging Growth A	POEGX	Aggressive Growth
Royce Value Fund	RYVFX	Aggressive Growth

TABLE 1
Unit Root and Cointegration Tests
1986.01 - 1995.12

The critical values at the 1% level were taken from McKinnon (1991). Column 2 presents the unit root statistics. Failure to reject the null hypothesis indicates a unit root process. The results show that none of the fifteen tests are significant at the 1% level. This then indicates presence of a unit root in each of the fifteen mutual funds and the four mutual fund indexes. Column 3 presents the Augmented Dickey-Fuller (ADF) statistic for a bivariate Engle-Granger cointegration test. None of the fifteen tests are significant at the 1% level, implying no cointegration between the SRFs and their peer groups.

Fund Ticker Symbol	Unit Root Test	Cointegration Test
	Panel A: Socially-Responsible Mutual Funds	
CSIFX	-2.319	-3.332
PAXWX	-1.448	-1.370
DRTHX	-2.486	-2.090
NALFX	-3.585	-1.963
ANEFX	-1.872	-1.662
PARNX	-2.070	-1.564
PIODX	-3.088	-1.572
PIOTX	-3.486	-2.554
PITHX	-2.875	-1.020
INVPX	-2.446	-1.075
SCDGX	+1.561	-2.404
EVGRX	-1.822	-1.871
PHSTX	+1.508	-1.510
POEGX	+2.071	+1.440
RYVFX	-2.726	+2.615
	Panel B: Mutual Fund Indexes	
Balanced Funds	-1.835	
Growth Funds	-2.095	
Growth & Income Funds	-1.649	
Aggressive Growth	-2.504	

TABLE 2
The External Performance of Socially-Responsible Mutual Funds
1986.01-1995.12

Fund Ticker	$\overline{R}_{Fund\ i}$	$\sigma_{Fund\ i}$	$SH_{Fund\ i}$	SH_{Peer} [a]	JB Z-Stat [b]
CSIFX	0.859%	2.586%	0.158	0.170	-0.221
PAXWX	0.865	2.712	0.152	0.170	+0.138
DRTHX	0.993	3.905	0.139	0.166	+0.513
NALFX	0.858	3.932	0.103	0.166	+1.635
ANEFX	1.198	4.464	0.167	0.166	+0.174
PARNX	1.303	6.074	0.140	0.166	+1.020
PIODX	0.973	4.095	0.127	0.162	+1.936
PIOTX	0.921	4.306	0.109	0.162	+2.185
PITHX	0.970	4.666	0.111	0.162	+1.947
INVPX	1.129	4.402	0.154	0.162	+1.188
SCDGX	1.164	3.682	0.194	0.162	-1.380
EVGRX	1.041	4.560	0.129	0.134	-0.390
PHSTX	1.488	5.328	0.195	0.134	-1.158
POEGX	1.659	6.425	0.188	0.134	-1.548
RYVFX	0.876	3.632	0.117	0.134	-0.559
Averages	1.086	4.318	0.146	0.157	

[a] SH_{Peer} represents the Sharpe's performance measure of an index of funds with the same stated objective as the socially-responsible fund. Average monthly return and standard deviation of peer group indices are:

Peer Group Index	$\overline{R}_{Peer}$	σ_{Peer}
Balanced Funds	0.944%	2.901%
Growth Funds	1.151	4.208
Growth & Income Funds	1.052	3.700
Aggressive Growth Funds	1.177	5.417

During the same time period, the average monthly rate of return on 30-day T-bills was 0.452%. The S&P 500 averaged 1.256% with a standard deviation of 4.331%. The Sharpe ratio for the S&P500 is 0.186.

[b] For the Jobson-Korkie test, the acceptance regions for a 1% test of equivalent performance are $-2.326<Z<2.326$.