

SCREENING FOR CORPORATE SOCIAL RESPONSIBILITY

John Dravenstott, Natalie Chieffe¹

Ohio University, Athens, Ohio 45701

1. Introduction and Summary

After more than a decade of studying the benefits of socially responsible investing (SRI) or ethical investing (EI), some uncertainty of the financial viability of the practice still exists.² SRI has attracted upwards of \$2 trillion of invested assets, and the amount is set to grow as the popularity of the practice increases. SRI can have many meanings³; in general it refers to an investment practice taking into account a company's social responsibility before considering it for investment. Figure 1 illustrates the types of investments that SRI generally targets. Companies are arranged by a responsibility rating with very responsible companies falling in section C, irresponsible companies in section A, and most commonly, companies with a mixture of positive and negative traits in Section B. SRI portfolios attempt to invest in more companies on the responsible side of the spectrum, in sections B and C, with traits such as good corporate governance, good diversity, and regard for the environment. A portfolio without SRI-constraints would make no such distinction, investing in companies with no regard for social rating and thus possibly being comprised of relatively more companies from the irresponsible side of the spectrum.

[Figure 1 about here]

¹ Contact author: Tel (740) 593-9320; Fax (740) 593-9539; chieffe@ohio.edu

² See, for example, Statman (2000), Kurtz (2005), Shank, Manullang, and Hill (2005), or de Graaf (2009).

³ See, for example, Henningsen (2002), Hill, Stephens, and Smith (2003), Hawken (2004), Gay and Klaassen (2005), or Kurtz (2005).

The performance of SRI portfolios could be measured by comparing the returns of companies from the B and C sections of the spectrum to the returns of the entire investment universe, A, B, and C, since this would directly compare an SRI strategy with a standard strategy. A more precise assessment of financial performance differences due to SRI traits, however, would come from comparing portfolios A and C, since these are groups of companies that are expected to be relatively more heavily weighted in standard and SRI portfolios, respectively.

This comparison of portfolios comprised of companies from sections A and C cuts to the core of the argument: does a portfolio with relatively more companies from the responsible side of the social spectrum outperform a portfolio with relatively more companies from the irresponsible side of the spectrum? This study finds, on the contrary, that socially responsible companies perform worse than irresponsible companies. The next question is why there is a performance disparity. The study finds reasons behind this performance disparity through analyzing the components of a company's social responsibility rating – individual social screens – determining that certain screens have different impacts on returns than others, some positive and some negative. Finally, investors are interested in the characteristics of these responsible companies and whether building such a portfolio exposes investors to biases in company size and sector. This study gives insight into the companies most likely to receive positive and negative social ratings to better inform investors considering such a practice.

After we test the tendency of irresponsible companies to outperform responsible companies, our study takes a close look into the components of responsibility ratings to determine why. As shown in Table 1, strengths and weaknesses in individual social screens

affect company financial returns differently. The Corporate Governance screen is the only screen for which a positive relationship with returns is observed: a positive rating helps returns while a negative rating hurts returns. All other screens (with the possible exceptions of the Product screen, which gave conflicting results, and the Community screen, where results are not statistically significant at 10 percent) show an inverse relationship; poor social responsibility causes higher returns, or good social responsibility causes lower returns. These results offer insights into why socially responsible companies underperform in the study.

[Table 1 about here]

While analyzing the returns of socially responsible and irresponsible companies in individual social screens, we assess the demographics of these companies to profile responsible and irresponsible companies. Larger companies are more likely to be rated and as shown in Table 2, certain sectors are more likely to be considered responsible (e.g. financials and IT) or irresponsible (e.g. energy, materials, and utilities). Sectors are somewhat important to a company's likelihood of being rated, but they are especially relevant in the analysis of individual social screens.

[Table 2 about here]

Though this study suggests that socially responsible companies underperform irresponsible companies, this does not necessarily condemn socially responsible investors to lower returns. Our tests sort companies into portfolios solely on the basis of responsibility ratings; no serious SRI fund manager would form a portfolio this way. SRI also requires a degree of financial consideration. Some SRI mutual fund and index studies have shown outperformance for SRI, suggesting that some SRI managers have been effective in supplementing a socially responsible strategy with financial criteria. Indeed, the analysis of

individual social screens shows that certain SRI criteria actually lead to positive portfolio returns, suggesting that a careful use of social ratings can benefit a portfolio, even if social ratings themselves are not an effective way to build a portfolio.

2. Background

Lusk et al. (2006) attempt to find whether social ratings data from Kinder, Lydenburg and Domini (KLD) affect a firm's stock price. Concentrating on seven KLD issues, the authors added the positive and negative scores for each issue. A positive score would mean that the company had a "good" ranking for the issue, while a negative score would suggest a "bad" ranking. They then put the "pure-relative-strength organizations (PRSO)" group (at least one net positive issue and no negative issues) and the "pure-relative- concerns organizations (PRCO)" group (at least one net negative issue and no positive issues) into two portfolios. They found that over the short period studied, the PRSO group significantly outperformed the PRCO group based on Jensen's alpha and other measures.

Though the study did not directly examine the significance of each KLD issue, the groundwork that the authors laid in their study was influential to developing the methodology to our report's analysis of aggregate social responsibility.

Building on the research discussed above, we compare the returns and characteristics of socially responsible firms to those of socially irresponsible firms. Our approach involves a large sample of approximately 3,000 total companies as well as Lusk et al.'s determination of social responsibility, which should add color to the observations of the field.⁴ The results of such tests show would-be SRI practitioners the risks in orienting investment towards responsible

⁴ See the description of Lusk et al. (2006) methodology. The PRSO portfolio designates responsible companies, while the PRCO portfolio designates irresponsible companies.

companies, answering the question of whether social ratings alone can be used to build a portfolio.

3. Data

We obtain several years of company financial data from Research Insight, published by Standard & Poor's (formerly COMPUSTAT PC Plus) for approximately 11,000 active publicly traded companies in the United States and Canada. We also obtain corporate social ratings data from KLD Research & Analytics, Inc's (KLD) KLD STATS (Statistical Tool for the Analysis of Trends in Social and Environmental Performance) database for each year from 2001 through 2005.

KLD created "the first benchmark for equity portfolios subject to multiple social screens." It measures corporate social responsibility across multiple dimensions. The screens are grouped into seven qualitative areas – Community, Corporate Governance, Diversity, Employment, Environment, Human Rights, and Product – and six exclusionary screens – Alcohol, Gambling, Tobacco, Firearms, Military, and Nuclear – discussed as a group as the "Exclusionary" screen. Within each screen are several criteria measuring positive and negative aspects of a company's compliance with the given screen. Sharfman (1996) comments that, "no specific theory was used to develop" KLD's ratings criteria, but that "they have become common social performance 'screens' in the investment community." In each of the social screens, companies are given strengths, weaknesses, both, or no scores. Brief descriptions of the screens can be found in the Appendix.

4. Methodology

The tests in this study focus on whether companies benefit or suffer financially from strong social policies relative to companies with weak social performance. We assume that

socially strong companies are included as often in SRI portfolios as in standard portfolios and socially weak companies are likely screened out of socially responsible portfolios but are more often found in standard portfolios. The lessons of these tests can also be applied to the field arguments of whether an SRI strategy can outperform other investment strategies.

4.1 Developing the Portfolios

Though KLD's database offers vast information about company social ratings, it does not attempt to give a determination of the most responsible and irresponsible companies overall. Even given guidance by the company and users of the data, it is difficult to determine the most appropriate way to divide companies into "responsible" and "irresponsible" portfolios in order to compare the characteristics of responsible and irresponsible companies. Several researchers have noted the problems associated with finding an aggregate measure of responsibility through KLD's and other firms' single-issue screen scores. Derwall et al. (2005) separate responsible companies from irresponsible companies by putting the highest rated 30 percent of market capitalization in a responsible portfolio and the lowest rated 30 percent of market capitalization in an irresponsible portfolio. Waddock (2000) separates companies in the S&P 500, putting the approximately 250 companies in the Domini 400 Social Index into a "responsible" portfolio and the approximately 250 companies not included in the Domini Index into an "irresponsible" portfolio.

Brammer et al. (2006) have no problem simply normalizing and adding ratings from the Ethical Investment Research Service (EIRS) across three equally-weighted disaggregated categories of social performance measurements. Sharfman (1996), however, notes that a simple addition of positive ratings and subtraction of negative ratings across screens using KLD data does not give an accurate picture of social responsibility since some screens are more important

to socially responsible investors, and suggests that weightings be used to determine a company's social responsibility. Few, however, have attempted this.⁵ KLD itself "urges caution in drawing inferences by using an aggregate KLD score," but Lusk et al. (2006) found an innovative way to use KLD measures in the assessment of aggregate social responsibility that seems to eliminate some of the ambiguity in determining aggregate social ratings. We use this methodology to divide companies for this study's purposes as follows.

Using KLD Stats 2005 social ratings of 3,015 companies, we group companies into four portfolios, classifying them as responsible, irresponsible, somewhere in the middle, or with no data available. A cursory historical analysis of KLD ratings finds little deviation in social ratings over the period 2001 to 2005; the number of companies screened by KLD changed drastically over the period (only 1,000 companies are screened in 2001, and of those, hundreds lack financial data in 2006 because of mergers, bankruptcies and other events causing de-listings); and precedence exists for studies using only the final year of data, since it has been shown that concurrent relationships between social ratings and financial performance are just as strong as lagging and leading relationships.⁶

We examine the seven qualitative screens and the exclusionary screens to find whether companies are, according to Lusk et al., "pure-relative-strength organizations (PRSO)" or "pure-relative-concerns organizations (PRCO)." A company with at least one net positive rating in a social screen and no net negative ratings is considered PRSO, or socially "Responsible," while a company with at least one net negative rating and no net positive ratings is considered PRCO, or socially "Irresponsible." A net positive is recorded if a company has more strengths than

⁵ See Sharfman (1996) for two examples of social screen weightings from the early 1990s. Little other evidence of such weightings being used in SRI studies exists.

⁶ See Orlitzky et al. (2003) for this finding of concurrent relationships and Waddock (2000) and Mattingly and Berman (2006) for precedence in using social ratings in this way.

concerns in a given social screen. For example, Exxon-Mobil has two diversity strengths (+2) and one diversity concern (-1), giving it a net positive rating for Diversity (+1). On the other hand, Credence Systems has one Employment strength (+1) and two concerns (-2), giving it a net negative rating in Employment (-1). Dell has one Governance strength (+1) and one Governance concern (-1), resulting in neither a net positive nor a net negative rating in Governance (0).

Companies that fail to merit a net rating in any social screen are considered Zero scores. This does not necessarily mean that the companies have no ratings, though this is the case for many Zero-score companies, which tend to be much smaller than rated companies.⁷ It can also mean that any positives in a screen are canceled out by negatives in the same screen. Because size tends to be the main issue in a company's achievement of a Zero rating rather than its social performance, in accordance with Lusk et al. (2006) and Brammer et al. (2006), the Zero portfolio is considered to be separate from the Middle portfolio. Companies in the Middle portfolio have at least two ratings contradicting each other, as in the Dell example above.

4.2 Validity Test

We test to determine whether the Lusk et al. groupings accurately reflect companies that are considered to be socially responsible. KLD Stats 2005 gives true or false statements regarding a company's 2005 membership in one of three KLD social indexes: the Domini 400 Social Index (DS 400), which is based on the S&P 500 index; the KLD Large Cap Social Index (LCS), which is based on the Russell 1000; and the KLD Broad Market Social Index (BMS), which attempts to find the best environmental, social, and governance performers of the Russell 3000 index.⁸ Companies that are part of these indexes are viewed as being more socially

⁷ This observation is consistent with the findings of Brammer et al. (2006) who find that only 296 of 451 companies had scores, with firm size rather than neutral firm social performance appearing to be the main reason for companies' not having ratings.

⁸ See index descriptions on the KLD website, <<http://www.kld.com>>.

responsible. Responsible companies consistently have higher social index membership than Irresponsible companies, and are more likely to be included in social portfolios, though these companies do not always represent the most socially responsible portfolio. Notably, more Middle companies are included in the DS 400 and the LCS. These portfolios, however, include large cap firms. If a company is not involved in the DS 400 or the LCS, it may mean that the company is simply not large enough. A company's non-involvement in the BMS, however, should suggest that KLD does not consider it among the highest social performers. When the all-encompassing BMS is considered, the results are as expected. Responsible companies have by far the highest percentage of participation and Irresponsible companies have the lowest. Thus, the validity of the Lusk et al. criteria of determining aggregate social responsibility using KLD single-issue screens is supported.

4.3 Demographics of Aggregate Portfolios

We use S&P Research Insight to find market capitalization data for 2,738 of the 3,015 companies rated by KLD in 2005. The remaining companies do not have Market Capitalization data available for a variety of reasons. Unavailable market capitalization values are similar for the four portfolios, with close to 91 percent of market capitalization values available for each.

The Middle portfolio has the highest average market capitalization of any portfolio by far. Responsible companies seem slightly larger than irresponsible companies, but the difference is not statistically significant at the five percent level. The results suggest that market capitalizations are not smaller for socially responsible companies and may in fact be larger.

4.4 Likelihood for Being Tested

Certain results beg the question of whether company size and sector designation affect a company's likelihood of being rated by KLD. It appears that a company's having strength and

concern ratings is largely a product of its size, as companies in the Zero portfolio, those with no net ratings in social screens, are much smaller in terms of both market capitalization and sales than all other companies with ratings. Companies in the Middle portfolio, those with at least two ratings contradicting each other, are much larger. It also appears that some sectors, such as Health Care, Consumer Staples, and Materials, receive more ratings scrutiny.

We use a regression to test the hypothesis that size and sector affect the likelihood of being tested by KLD. We record the number of “ticks,” the sum of positive and negative social ratings scores, for each company. If a company is given either a positive or negative rating in any aspect of a screen, a tick with a value of one (1) is recorded for that aspect. General Electric and General Motors each collect 28 ticks, the highest amount recorded. The average number of ticks is 2.97 per company for the 3,015 companies rated in 2005, with a median of two ticks per company and a standard deviation of 3.13, shown in Table 4.

[TABLE 4 ABOUT HERE]

The number of ticks is the dependent variable in the regression shown in Table 5, while market capitalization and sector designations are independent variables. As is apparent in the table, market capitalization is a strong indicator of whether a company will be tested by KLD. The larger a company is in terms of market capitalization, the more likely the company will be subjected to KLD ratings scrutiny, positive or negative. An additional tick is associated with a \$10 billion increase in market capitalization. No sector seems to be tested excessively at a statistically significant five percent level, though it appears that Materials and Utilities companies have a rather higher likelihood of receiving ratings relative to other sectors at low levels of significance ($p < .10$).

[TABLE 5 ABOUT HERE]

4.5 Aggregate Responsible/Irresponsible Portfolio Returns

An independent T-test determines whether the portfolio of socially responsible companies outperforms the portfolio of irresponsible companies over the five year period studied. Companies with less than five years of monthly pricing data available are eliminated from the analysis. Prices are adjusted for dividends and splits. Consistent data is available for all five years for 2,418 of the 3,015 companies in the sample. Though returns are available for all four portfolios, in accordance with Lusk et al. (2006), only the responsible and irresponsible portfolios are compared to view a responsibility effect on returns “because if we expect to see a market effect, this partition is where it should be observed” (p. 109).

The Irresponsible and Responsible portfolios consist of 999 and 344 companies respectively. As is apparent from Table 6, the Irresponsible portfolio has monthly returns of .018 to the Responsible portfolio’s .014. The difference is significant at five percent.

[TABLE 6 ABOUT HERE]

A multi-factor regression analysis is used to assess whether differences in firm size or sector cause the returns difference. For the 1,287 companies we use a dummy variable to indicate the company’s responsibility. A one (1) indicates that the company is responsible, while a zero (0) represents a company considered to be irresponsible. Table 7 suggests that even after taking market capitalization and economic sector into account, Irresponsible companies outperform Responsible companies over the five year period with statistically significant results. A negative beta is observed for the social responsibility variable, suggesting that for an increase in responsibility rating, one would expect a decrease in returns.

The results are inconsistent with those of Waddock (2000). That study predicts “marginally higher financial performance may be consistent with better social performance”

however the differences are not statistically significant. Results are consistent, however, with several other researchers who predict lower returns for socially responsible portfolios⁹.

[TABLE 7 ABOUT HERE]

We use social ratings for the year 2005 for each of the seven qualitative social screens as well as for the composite category of exclusionary screens. Due, however, to the extremely small sample size of the companies rated in Human Rights, we omit data for this screen, thus limiting the trial to the study of six qualitative screens and one exclusionary screen.

For the qualitative screens, companies that have at least one positive rating for an individual screen in 2005 are put into the “Strength” portfolio for that screen, while companies with at least one negative rating are put into the “Concern” portfolio for the screen. Following the guidance of Mattingly and Berman (2006), a company can be included in each of these portfolios, as companies sometimes have both a positive and negative rating for certain screens. In the case that a company does not have a positive or a negative rating for a screen, that company is placed in the “Zero” portfolio, not implying any strength or weakness. For the Exclusionary composite screen, companies are simply sorted into Strength and Concern portfolios, since a company can only have a concern or not have a concern in the exclusionary screens. KLD does not give any strength scores in these screens; thus not having a concern is enough to merit a “strength” rating for the Exclusionary screen.

Once the portfolios are developed, we determine the demographics of each social screen to see if the assessments of the aggregate tests – that company size and sales do not have a statistically significant effect on a company’s rating and that certain sectors are more likely to be rated positively than others – are supported within each screen.

⁹ See, for example, Bauer et al (2005), or Entine (2003).

4.6 Effects of Individual Social Screens on Financial Performance

After examining the different size and sector attributes of companies across screens, we compare performance. We use independent T-tests to assess differences in returns for companies with strength and concern scores in individual social screens. Then we use a multifactor regression and consider the possible effects of industry and market capitalization on returns.

We collect the average monthly returns for each of the three portfolios as shown in Table 9. We conduct independent T-tests to determine whether the difference between the Strength and Concern portfolios is statistically significant.

The results show that Concern companies outperform Strength companies in the screens for Community, Product, Diversity, Employment, and Environment. Strength companies only outperform in the Corporate Governance screen, and the portfolios perform equally in Exclusions. Results are statistically significant at the five percent level, with the exception of the comparison of Strength and Concern companies in the Product screen. Because of the observable differences in the size and sector of the companies screened in these portfolios we develop a multifactor regression, presented in Table 10, to account for some of the disparities.

[TABLE 9 ABOUT HERE]

The dependent variable is average monthly holding period returns. Again, because of the small sample of companies with Human Rights ratings, the screen was omitted. Thus, concerns and strengths for the six qualitative screens and the exclusionary concern category make up 13 factors of social responsibility for the regression.

[TABLE 10 ABOUT HERE]

The results of the regression are rich in that the findings are mixed. As hypothesized, ratings in certain screens affect portfolio returns differently than ratings in others. Also

consistent with expectations, many observations suggested by the independent T-tests have proven to be inaccurate in evaluating returns when other factors (in this case market capitalization and economic sector) are taken into consideration. This is a consistent finding of researchers in SRI – among them Kurtz (1997) – that evaluations of SRI must take into account the composition biases of SRI portfolios. Of the screens studied, the Corporate Governance screen helps returns most unambiguously. A strength in Corporate Governance helps monthly returns, while a weakness in Corporate Governance conversely hinders returns. A concern in Product hinders returns, but the statistical significance of this finding is tempered by the finding that a strength in Product also hinders returns.

An inverse relationship to returns is observed for the Employment screen, as companies with strengths in Employment achieve lower returns over the period. Diversity ratings have a similar inverse relationship with returns; a concern in Diversity correlates strongly to higher returns. Companies with good Environment ratings also underperform over the period.

Other screens have lower levels of significance, but their results are interesting as well. Though the Exclusion screen does not show a performance difference for strength and concern companies in the independent T-test, the regression finds a correlation, significant at 10 percent, between higher returns and exclusionary ratings. The Community screen does not show significant correlations between ratings and returns either way, though the independent T-test indicates outperformance for irresponsible companies.

One should note that, as suggested by Mattingly and Berman (2006), a positive relationship to returns with a favorable rating is not always mirrored by a negative relationship with a poor rating, or vice versa. Indeed, Corporate Governance is the only screen for which a strength causes higher returns and a concern lower returns. This is likely due to the fact that

KLD ratings criteria for Corporate Governance are polar: the criteria for strengths are often the inverse of the criteria for weaknesses. For example, “Limited compensation” and “Ownership strength” earn a positive point, while “High Compensation” and “Ownership concern” cost a negative point. For other screens, such as Product, the Strength and Concern ratings do not directly contradict. For example “Quality,” “Innovation,” and “Benefits to economically disadvantaged,” earn positive ratings, contrasted by “Product safety,” “Marketing/contracting concern,” and “Antitrust” receiving negative ratings. This difference in the corporate responsibility aspects measured could explain why Product ratings hinder returns whether they are positive or negative. Two completely different ideas are in fact being measured.

It appears, by comparing results for independent T-tests and the regression, that taking into account industry and size bias causes many of the effects on returns due to social ratings to diminish into statistical insignificance. Still, many results remain statistically significant and lead to interesting discussions on the impacts of SRI.

5. Implications and Conclusions

There do not seem to be major differences in terms of size between socially responsible and irresponsible companies. The lack of statistically significant size differences for companies with positive and negative social ratings contradicts inferences and conclusions of many studies and surveys of mutual funds and indexes, among them Camejo (2002) and Gay and Klaassen (2005). Specifically, this finding calls attention to the fact that larger companies are not necessarily less socially responsible than smaller companies in terms of social ratings. However, larger companies are more often rated than smaller companies and are likely to receive ratings in more categories since larger companies have more operations, subjecting them to scrutiny in more measurements.

Though company size does not seem to be much of a factor in determining social ratings, economic sector plays a large part. Many assumptions in the field are supported regarding sectors expected to have lower social ratings. Energy and Utilities companies are more likely to be considered Irresponsible in the aggregate test, while Financial and IT companies are more likely to be Responsible. These findings are backed by SRI literature¹⁰. It appears that Consumer Discretionary companies receive positive ratings for Diversity but these are overshadowed by weak ratings in Corporate Governance and Employment. Whether the differences in social ratings for companies in different economic sectors are positive or negative, the tests confirm that sectors are an important indicator of social performance. Studies of SRI should thus continue to take into account sector and industry bias, and investors should know that SRI necessarily exposes them to this risk of not being properly diversified, overexposed to certain sectors and underexposed to others.

5.1 Irresponsible Companies Outperform

The findings of this study suggest that developers of socially responsible portfolios must take into account other factors besides social performance when composing SRI portfolios. This in large part describes the current SRI industry, wherein mutual funds often list several factors influencing decisions in addition to choices based on responsibility. Previous studies alluding to SRI enhanced or equal performance typically study these mutual funds or indexes, and it seems that appropriate use of social data in tandem with financial criteria helps these portfolios to enjoy equal or better performance. This study, though, measures only a strict implementation of social ratings in investment style and finds that such a bent necessarily hinders returns.

¹⁰ For example, see Camejo (2002); Kurtz (1997).

Interestingly, the findings of underperformance for responsible companies contradict Lusk et al. (2006) who created portfolios in the same way, though with hundreds fewer companies, fewer years of data, and no consideration of industry and market capitalization factors. Lusk et al find the opposite result that “organizations with a strong KLD profile seem for the most part to be reaping market rewards, compared to those organizations which have policies that have raised concerns.” This contradiction confirms that results for SRI studies are largely dependent on their design and that creative studies should continue to be done to give a broad perspective of the impacts of SRI on returns. Though consistency and assessment standards could be useful, the “robustness” of results – from authors from a variety of fields studying SRI in a variety of ways with a variety of data – gives the field insights into whether, how, and why SRI returns differ from standard returns and how investors can implement the practice effectively.

6. Conclusion

It is apparent from this research that socially responsible companies underperformed irresponsible companies over the study period. This was true both for aggregate social ratings and for the majority of the individual social screens. This suggests that SRI portfolios (with more socially responsible companies and less irresponsible ones) should underperform standard portfolios (with relatively less socially responsible companies and more irresponsible companies). Yet this is not a clear indictment of SRI weakness.

The portfolios for this study are constructed solely on the basis of social ratings with no regard for financial criteria or sector weightings, while actual practitioners of SRI use other criteria to varying extents. Indeed, the definition of SRI used in Section One calls SRI a “process that considers the social and environmental consequences of investments, both positive

and negative, within the context of rigorous financial analysis.” Perhaps partial implementation of SRI principles in tandem with “rigorous financial analysis” can lead to equal performance or outperformance with the practice, particularly if individual SRI screens are implemented effectively. The individual screen test shows that certain social screens may be supportive of higher returns, while other screens hinder statistically significant higher returns. Regardless, a less-than rigorous use of social ratings in portfolio composition could easily lead to sector biases, exposing investors to additional risk than that of standard, more diversified portfolios.

Having confirmed in this study that social responsibility assessments alone cannot create a financially viable portfolio, future studies should further analyze actual SRI portfolios, those indexes and mutual funds composed using both social and financial criteria, to determine whether an effective use of SRI can in fact aid portfolio performance and the reasons why, especially in regard to the screens used. KLD data can be used to test whether each company held by SRI funds is in fact responsible enough for investors in these funds and whether funds have appropriately followed social investors’ screening criteria. The future of SRI depends on the ability of social investments to satisfy the dual goals of effectively implementing principled investment strategy and earning appropriate levels of financial returns, and all future SRI studies should address both of these goals.

7. Appendix.

Brief Descriptions of KLD Individual Social Screens¹¹

Community.

KLD rates companies in the community screen based on their records in charitable giving, support for education and housing, economic impact on communities, and other notable achievements or concerns in regard to community impact. Seven specific criteria are used to evaluate community strengths in 2005 and four specific criteria are used to evaluate community concerns. In 2005, 15 percent of companies have positive or negative ratings in the category.

Corporate Governance.

Corporate governance ratings are assigned based on how companies compensate executives and directors, transparency and accountability in social, political, or financial records, and other positive initiatives or controversies. Five specific criteria are used to evaluate corporate governance strengths in 2005 and six specific criteria are used to evaluate concerns. In 2005, 40 percent of companies have positive or negative ratings in the category.

Diversity.

KLD gives diversity ratings in eight specific criteria for strengths and three for concerns in 2005. Ratings are based on the inclusion of women and minorities in top management, directorships, and promotions, policies for gay and lesbian employees and disabled employees, and other strengths or concerns in diversity. In 2005, 70 percent of companies receive strength or concern ratings in the category.

Employment.

Employment ratings are given to companies based on relationships with unions and employees; profit-sharing, retirement, and health benefits provided; workplace health and safety; and layoff policies. In 2005, six specific positive criteria are used and five specific negative criteria are used. Of KLD's 3,015 companies, 48 percent receive strength or concern scores in employment.

Environment.

KLD assigns environmental ratings to companies based on their environmental impact through five specific positive and seven specific negative criteria, including recycling policies, pollution, energy derivation, and climate change impact among other criteria. In 2005, 15 percent of companies have positive or negative ratings in the category.

Human Rights.

Human rights ratings are given to companies based on their international employment policies and involvement with troublesome regimes. In the past, negative scores were given for companies with involvement in South Africa's apartheid regime, Northern Ireland, and Mexico. Today, screens include involvement in Burma and treatment of indigenous peoples. In 2005,

¹¹ Descriptions are summarized from those provided by KLD Research & Analytics, Inc. All rights reserved. Copyright 2007. For more information, please visit www.kld.com or contact clientservices@kld.com.

four percent of companies have positive or negative ratings in the category, with the vast majority receiving concern scores, a majority of which relate to international labor concerns.

Product.

Product scores are assigned to companies based on product safety and quality, marketing techniques, and antitrust concerns among other issues. In 2005, four specific positive and four specific negative criteria are used, with positive or negative ratings going to 18 percent of the 3,015 companies.

Exclusionary.

Exclusionary screens (referred to by KLD as “Controversial Business Issues”) include company involvement in alcohol, tobacco, gambling, firearms, military contracting, and nuclear power. The screens are grouped together because of the low number of companies receiving scores in any one area and to test a traditional definition of SRI discussed in Section One, which typically lumps many of these concerns together when building an SRI portfolio.¹² Each of the screens test negative criteria based on the percentage of revenues coming from the “vice,” licensing in a negative area, or major association with the concern area. There are no positives in the areas, as a company can only have a negative score or no score. In this regard, seven percent of companies have negative scores, showing an association with a vice.

¹² Some authors omit such screens from studies completely, such as Lusk et al. (2006) who see possible benefits to gambling (lotteries supporting schools), alcohol (health benefits from moderate consumption), and defense (justified and desired protection for citizens) that cloud the issue over whether such issues should even be considered an element of irresponsibility.

8. References

- Bauer, R., Koedijk, K., & Otten, R. (2005). International Evidence on Ethical Mutual Fund Performance and Investment Style. *Journal of Banking and Finance*, 29, 1751-1767.
- Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate Social Performance and Stock Returns: UK Evidence from Disaggregate Measures. *Financial Management*, 35(3), 97-116.
- Camejo, P. (2002). *The SRI Advantage*. Gabriela Island, B.C., Canada: New Society Publishers.
- deGraaf, F. (2009). Guidelines for Integrating Socially Responsible Investment in the Investment Process. *Journal of Investing*, 18(3), 70-78.
- Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal*, 61(2), 51-63.
- Domini, A. L. (2001). *Socially Responsible Investing*. Chicago, IL: Dearborn Financial Publishing, Inc.
- Entine, J. (2003). The Myth of Social Investing: A Critique of its Practice and Consequences for Corporate Social Performance Research." *Organization & Environment*, 16 (3), 352-368.
- Gay, G. R., & Klaassen, J. A. (2005). Retirement Investment, Fiduciary Obligations, and Socially Responsible Investing. *Journal of Deferred Compensation*, 10(4), 34-49.
- Hawken, P. (2004, October). Socially Responsible Investing. *Natural Capital Institute*, Sausalito, CA. Available: http://www.naturalcapital.org/docs/SRI%20Report%2010-04_word.pdf [2007, May 13].
- Henningsen, C. (2002). Investing as if the World Really Mattered. *Corporate Environmental Strategy*. 9(2), 163-171.
- Hill, R.P., Stephens, D. & Smith, I. (2003) Corporate Social Responsibility: An Examination of Individual Firm Behavior. *Business and Society Review*. 108(3), 339-364.
- Kurtz, L. (1997). No Effect, or No Net Effect? Studies on Socially Responsible Investing. *Journal of Investing*, 6(4), 37-49.
- Kurtz, L. (2005). Answers to Four Questions. *The Journal of Investing*, 14(3), 125-139.
- Kurtz, L., & diBartolomeo, D. (2005). The KLD Catholic Values 400 Index. *The Journal of Investing*, 14(3), 101-104.

- Lusk, E. J., Halperin, M., & Zhang, B. (2006). The Balanced Scorecard: Suggestions for Rebalancing. *Problems and perspectives in management*, 4(2), 100-114.
- Mattingly, J. E., & Berman, S. L. (2006). Measurement of Corporate Social Action: Discovering Taxonomy in the KLD Ratings Data. *Business and Society*, 45(1), 20-46.
- Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate Social and Financial Performance: A Meta-Analysis. *Organization Studies*, 24, 403-441.
- Shank, T.M., Manullang, D.K. & Hill, R.P. (2005). Is it Better to be Naughty or Nice? *Journal of Investing*, 14 (3), 82-87.
- Sharfman, M. (1996). The Construct Validity of the Kinder, Lydenberg, and Domini Social Performance Rating Ddata. *Journal of Business Ethics*, 15, 287-296.
- Statman, M. (2000). Socially Responsible Mutual Funds. *Financial Analysts Journal*, 56(3), 30-39.
- Waddock, S. (2000). Performance Characteristics of Social and Traditional Investments. *Journal of Investing*, 9(2), 27-38.

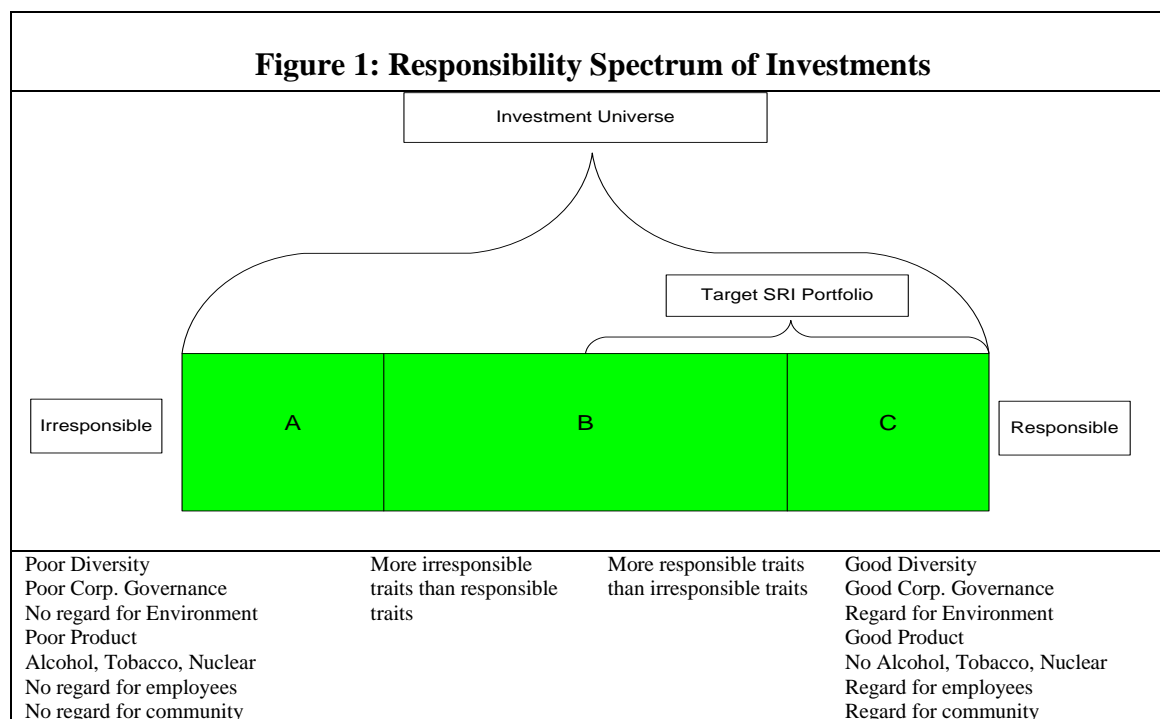


Table 1: Abbreviated Results of Multifactor Regression for Individual Social Screens

	Strength	Concern
Community	Insignificant	Insignificant
Corporate governance	Helped	Hindered
Diversity	Insignificant	Helped
Employment	Hindered	Insignificant
Environment	Hindered	Insignificant
Product	Hindered	Hindered
Exclusion	N/A	Helped

This table displays the social screens measured in the study and the effects of strength (positive ratings in a screen) and concern (negative ratings in a screen) ratings in each screen on portfolio returns. “Insignificant” refers to screens that do not affect returns at a statistically significant level ($p > .10$).

Table 2: Abbreviated Findings of Social Ratings and Responsibility of Companies Based on Sector and Size

		Likelihood of receiving ratings						
	Social orientation	Overall	Corporate Governance	Community	Diversity	Employee	Environment	Product
Energy	Irresponsible	-	+	+	-	-	+	-
Materials	Irresponsible	+	+	+	-	+	+	+
Industrials	Irresponsible	-	-	-	-	+	+	+
Consumer Discretionary	Irresponsible	-	+	-	+	+	-	+
Consumer Staples	Unclear	+	+	+	+	+	+	+
Health Care	Unclear	-	-	-	+	-	-	+
Financials	Responsible	-	-	+	-	-	-	-
Information Technology	Irresponsible	-	+	-	+	+	-	-
Telecommunications	Unclear	-	-	-	+	-	-	-
Utilities	Irresponsible	+	+	+	-	+	+	+
Large Market Capitalization	Unclear	+	+	+	+	+	+	+
Small Market Capitalization	Unclear	-	-	-	-	-	-	-

+ High likelihood of being rated in category (rated more often than average)
- Low likelihood of being rated (rated less often than average)

Table 3: Correlations among Socially Rated Portfolios and Sectors

	Pearson Correlations			
	Responsible	Irresponsible	Middle	Zero
Energy	-0.060**	0.064**	-0.025	0.006
Materials	-0.040	0.049*	-0.007	-0.025
Industrials	-0.023	0.036	-0.026	0.011
ConsDisc	-0.061**	-0.022	0.075**	-0.014
ConsStaples	-0.018	-0.003	0.037	-0.036
Health	0.018	0.023	-0.039	0.003
Financial	0.103**	-0.054**	-0.054**	0.058**
IT	0.052*	-0.022	0.003	-0.030
Telecom	-0.007	-0.019	0.023	0.003
Utilities	-0.061**	-0.010	0.057**	-0.005

A positive correlation suggests that the sector is likely to be included in the portfolio, while a negative correlation suggests that the sector is excluded from the portfolio.

Companies in the Middle Portfolio have at least two ratings contradicting each other.

Companies in the Zero Portfolio have no ratings.

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

N=2,418

Table 4: Likelihood of Receiving Social Ratings - Number of Ticks

	Mean	Standard Deviation	Median	Min	Max
Energy	3.25	3.59	2	0	24
Materials	4.64	4.11	4	0	25
Industrials	2.96	4.00	2	0	28
ConsDisc	3.10	2.88	2	0	28
ConsStpls	4.68	4.10	3	0	21
HealthCare	2.42	2.83	2	0	24
Financials	2.34	2.78	2	0	27
IT	2.88	2.79	2	0	27
Telecom	2.64	3.78	1	0	18
Utilities	5.15	3.74	5	0	18
Overall	2.97	3.13	2	0	28

A tick was recorded for each social rating, positive or negative, in any of KLD's qualitative or exclusionary screens in 2005.

Table 5: Likelihood of Being Tested by KLD Regression

	Unstandardized Coefficients		Standardized	T	Sig.
	B	Std. Error	Coefficient		
(Constant)	-0.024	2.337		-0.010	0.992
MktCap	1xE-04	0.000	0.603	40.975	0.00**
Energy	2.234	2.346	0.161	0.952	0.341
Materials	4.235	2.347	0.288	1.805	0.071*
Industrials	2.508	2.341	0.273	1.071	0.284
ConsDisc	2.799	2.340	0.342	1.196	0.232
ConsStpls	3.393	2.349	0.214	1.445	0.149
HealthCare	1.922	2.341	0.210	0.821	0.412
Financials	1.756	2.339	0.240	0.751	0.453
IT	2.410	2.340	0.297	1.030	0.303
Telecom	1.825	2.367	0.071	0.771	0.441
Utilities	4.388	2.351	0.251	1.867	0.062*

Dependent variable: Ticks (total number of positive and negative ratings).

Adjusted R² =.416

F = 176.87

** Significant at 1%

* Significant at 10%

Table 6: Independent Portfolio T-Test; Responsible Versus Irresponsible Returns

Portfolio	N	Mean	Std. Deviation	Std. Error Mean
Irresponsible	999	0.0176	0.0166	0.0005
Responsible	344	0.0145	0.0143	0.0008

The independent T-test shows statistical significance in the performance difference of the two portfolios (p=.001)

Table 7: Regression of Holding Period Returns Considering Market Capitalization and Economic Sector

	Unstandardized Coefficients		Standardized Coefficient	T	Sig.
	B	Std. Error	Beta		
(Constant)	0.014	0.001		14.376	0.000
MktCap	-1xE-07	0.000	-0.067	-2.466	0.014
Energy	0.013	0.002	0.180	6.041	0.000
Materials	0.006	0.002	0.076	2.594	0.010
Industrials	0.005	0.002	0.102	3.176	0.002
ConsDisc	0.004	0.001	0.081	2.489	0.013
ConsStpls	0.002	0.002	0.026	0.910	0.363
HealthCare	0.001	0.002	0.016	0.506	0.613
IT	0.006	0.001	0.132	4.003	0.000
Telecom	0.006	0.004	0.041	1.475	0.140
Utilities	-0.005	0.003	-0.054	-1.887	0.059
Socially Responsible	-0.002	0.001	-0.061	-2.207	0.028

Dependent variable: Monthly holding period returns.

Adjusted R² = .047

F = 6.80

Table 8: Companies in Corporate Governance Portfolios by Sector

	Companies	% Strength	% Concern	% Zero
Energy	143	14%	36%	53%
Materials	130	15%	37%	55%
Industrials	365	14%	26%	62%
Consumer Discretionary	487	13%	34%	55%
Consumer Staples	116	22%	45%	43%
Health Care	362	12%	28%	63%
Financials	644	16%	17%	69%
Information Technology	483	17%	33%	54%
Telecommunications	47	9%	21%	74%
Utilities	94	20%	28%	59%
Total	2871	15%	28%	60%

Values are presented in bold when the percentage of companies represented in the relevant sector exceeds the average for all companies, presented in the bottom row of the table. Values of each sector add up to over 100 percent, since some companies are sorted into both the Strength and Concern portfolios.

Table 9: Independent T-tests of Portfolio Returns

Social Screen	Strength	Concern	Zero	Significance
Community	0.009	0.013	0.016	**
CGOV	0.017	0.012	0.016	**
Diversity	0.012	0.018	0.015	**
Employment	0.011	0.015	0.017	**
Environment	0.010	0.014	0.016	**
Product	0.009	0.010	0.017	
Exclusion	0.015	0.015		
**	Significant difference between Strength and Concern portfolio at 1%			

Table 10: Regression Run Using All Social Screens Except for Human Rights and Considering Sectors and Market Capitalization as Factors

	Unstandardized Coefficients		Standardized Coefficient	T	Sig.
	B	Std. Error	Beta		
Sectors					
Energy	0.014	0.002	0.190	8.355	0.000
IT	0.005	0.001	0.109	4.267	0.000
Industrials	0.005	0.001	0.101	4.088	0.000
Materials	0.006	0.002	0.084	3.701	0.000
Consumer Discretionary	0.002	0.001	0.051	2.027	0.043
Health Care	0.002	0.001	0.040	1.691	0.091
Utilities	-0.003	0.002	-0.038	-1.645	0.100
Consumer Staples	0.003	0.002	0.035	1.584	0.113
Telecommunications	0.003	0.003	0.021	1.060	0.289
Other Factors					
Market Capitalization	0.000	0.000	-0.037	-1.658	0.097
(Constant)	0.014	0.001		15.157	0.000
Social Screens					
Screens Helping Returns					
Concern: Diversity	0.003	0.001	0.086	3.702	0.000
Strength: Corp. Gov.	0.002	0.001	0.041	2.024	0.043
Concern: Exclusionary	0.002	0.001	0.039	1.858	0.063
Screens Hindering Returns					
Concern: Corp. Gov.	-0.003	0.001	-0.085	-4.019	0.000
Concern: Product	-0.003	0.001	-0.070	-3.195	0.001
Strength: Employment	-0.003	0.001	-0.061	-2.857	0.004
Strength: Environment	-0.005	0.002	-0.058	-2.805	0.005
Strength: Product	-0.004	0.002	-0.041	-1.998	0.046
Strength: Community	-0.001	0.001	-0.025	-1.109	0.267
Concern: Environment	-0.001	0.001	-0.024	-1.029	0.303
Strength: Diversity	-0.001	0.001	-0.023	-0.970	0.332
Concern: Community	0.000	0.001	-0.002	-0.108	0.914
Concern: Employment	0.000	0.001	-0.002	-0.072	0.942

Dependent variable: Monthly holding period returns.

Adjusted R² = .095

F = 11.69