Chapter 2

The stocks at stake: return and risk in socially responsible investment

2.1 Introduction

Socially responsible investing (SRI) attracts a lot of investor attention. According to the Social Investment Forum (www.socialinvest.org), approximately 10% of investments are managed according to some screening process related to SRI. In addition to financial performance, social investors are concerned about non-financial dimensions of corporate performance, such as the impact on the environment, social relations, and corporate governance. This multi-dimensional nature of corporate performance raises the question whether a trade-off exists between the financial dimensions of performance and the non-financial dimensions. Most empirical studies suggest no trade-off since there is little difference between the (risk-adjusted) returns of stocks satisfying SRI criteria (for an overview, see Renneboog et al., 2008). However, Hong and Kacperzcyk (2009) report higher expected returns for stocks that are excluded from a portfolio because of negative ethical issues (companies producing alcohol, tobacco, and gaming).

Theoretical work on the relationship between SRI and expected returns focuses on the discrepancies in prices that are the result of demand differences for different types of stocks. The demand differences can either be due to incomplete in-

This chapter is based on Galema, Plantinga and Scholtens (2008)

formation (Merton, 1987), (green) investor preferences (Heinkel et al., 2001) or the internalization of externalities (Dam, 2008). These explanations all share a basic feature, namely that excess demand for socially responsible stocks and a shortage of demand for irresponsible stocks will lead to overpricing of the first and underpricing of the latter. The shortage of demand for irresponsible firms' stock implies that the risk sharing opportunities for people investing in these stocks are limited and therefore command a return premium (Merton, 1987). In a neo-classical equilibrium model, Dam (2008) shows that even when we assume equal risk levels for responsible and irresponsible firms, socially responsible firms will have both lower returns and lower book-to-market ratios than their irresponsible industry peers. In contrast, the empirical literature on SRI does not conclude that responsibility is priced by the capital markets (Renneboog et al., 2008).

We study the contradiction between the results found in the empirical literature and the predictions from the theoretical models. In our view, it is the result of a misinterpretation of the risk-adjusted performance measures used in most empirical studies. This misinterpretation arises from two possible errors. The first error is that financial performance is calculated by controlling for systematic risk, whereas the empirical measure of systematic risk captures part of the trade-off. The second error is that aggregate measures of SRI may confound existing relationships between individual dimensions of SRI and returns.

The first error is related to the use of regression models including risk factors such as those suggested by Fama and French (1992). In these, the intercept is usually interpreted as a measure of risk-adjusted performance. An intercept not significantly different from zero is usually interpreted as evidence that the assets are priced according to a rational asset pricing model. However, SRI and non-SRI firms with equal risk levels may have different book-to-market ratios due to an excess demand for SRI stocks. In the context of the model of Fama and French (1992), this implies that the exposure to the factor reflecting the influence of the book-to-market factor is independent of the risk profile of the underlying cash flows. As a result, the trade-off between financial and SRI performance is at least partly captured by the book-to-market ratio.

The second error relates to the use of aggregate measures of SRI. One of the reasons why the empirical literature yields few significant relations between SRI and expected returns may be due to the aggregation over different dimensions that have confounding effects. For example, it is possible that positive news on environmental friendly production is positively related to expected returns, whereas news

pertaining to good employee relations is negatively related. Therefore, we investigate different dimensions of responsibility.

We will first analyze the risk-adjusted performance in a way that is consistent with the existing literature. To test the return implications of SRI, we form portfolios based on positive scores on the strength and concern screens of six SRI dimensions on the universe of stocks tracked by KLD Research & Analytics, Inc. during the period 1992 to 2006. We test whether each of these portfolios can deliver excess returns by estimating the Fama-French (1993) asset pricing model augmented with the Carhart (1997) momentum factor in a system GMM framework. Our results confirm earlier findings that the risk-adjusted performance of SRI stocks is not significantly different from zero. We take a closer look at the impact of SRI scores on excess returns by regressing these returns on lagged SRI scores for our six dimensions using Fama-MacBeth (1973) regressions. Next, we analyze the book-to-market ratio and find it is affected by some individual SRI scores. In particular, Diversity and Environment have a negative relationship with the book-to-market ratio whereas Governance has a positive relationship. Therefore, we conclude that SRI impacts on stock returns by lowering the book-to-market ratio and not by generating positive alphas in a linear regression model. This result is consistent with the theoretical work suggesting that SRI is reflected in demand differences between SRI and non-SRI stock. It also explains why so few studies do find a relationship between alphas and SRI investing.

This chapter is structured as follows. In section 2 we briefly discuss the literature on the relation between stock returns and SRI. In section 3, we discuss the data and the methodology. In section 4 we discuss the empirical results. Finally, section 5 concludes.

2.2 Background

Kurtz (1997) and Bauer et al. (2005) establish that socially responsible stocks do not appear to underperform the market as a whole (see Renneboog et al., 2008, for an overview of the literature). However, Kempf and Osthoff (2007) report positive and significant risk-adjusted returns during 1992-2004 for a US portfolio based on a sample of SRI stocks from the KLD database. Their portfolio is based on a long-short strategy by investing in the 10% best SRI stocks within each industry and shorting the 10% worst SRI stocks within each industry. In contrast, Hong and Kacperzcyk (2009) report higher expected returns for stocks that are usually

excluded from a portfolio because of negative ethical issues. For example, they exclude companies involved in alcohol, tobacco, and gaming. They find that these so-called sin stocks are underpriced, have higher book-to-market values, and higher excess returns than other stocks. In addition, they find that sin stocks have less analyst coverage and are less held by pension funds.

Most formal theoretical work on the relationship between SRI and expected returns is reminiscent of Merton (1987). Merton (1987) focuses on the discrepancies in prices that are the result of incomplete information in the sense that individual investors do not have information about all securities regarding expected return, variance, and covariance with other securities. Investors do not include securities in their portfolio for which they lack such information. Merton (1987) shows that in the presence of this type of incomplete information, firm size, firm-specific return variance and the fraction of investors that know about a security all impact on risk-adjusted returns. This model can also be used to analyze the impact of self-imposed restrictions on investment portfolios by considering the stocks that should be excluded from the portfolio based on some a priori criterion as stocks with lacking information (see also Białłkowski et al., 2008). Thus, in the context of SRI, the increased (or decreased) risk of a firm acting in a socially responsible manner could increase (or decrease) its expected return. Other theoretical models are based on differences in investor preferences regarding non-financial performance characteristics. For example, in Heinkel et al. (2001), investors differ with respect to their preference for green companies. Fama and French (2007) refer to differences in preference for any non-financial performance characteristic. Dam (2008) models investors as differing in tastes and firms as differing in their policies of internalizing externalities. These models all share one basic feature, namely that excess demand for socially responsible stocks and a shortage of demand for irresponsible stocks will lead to overpricing of the first and underpricing of the latter. The shortage of demand for irresponsible firms' stock implies that the risk sharing opportunities for people investing in these stocks are limited and therefore command a return premium.

2.3 Data & Method

2.3.1 Data

We obtain data on social responsibility from KLD Research & Analytics, Inc. and financial performance measures from Datastream. Other researchers have used

these databases too when investigating the relationship between financial performance and SRI (e.g. Hillman and Keim, 2001; Kempf and Osthoff, 2007). KLD uses screens to monitor SRI and it has expanded its universe of coverage over the last couple of years. In the 1990s, it covered the S&P500 Index and the Domini 400 Social Index. In 2001 the database was extended to include all constituents of the Russell 1000. In 2003 the database was further extended to include all stocks from the Russell 2000 as well. KLD does not have historical ratings data for non-US companies, unless it is a member of the S&P500. In our study, we will include all stocks covered by KLD.

KLD uses multiple criteria on which firms are evaluated, using both positive and negative screens. Positive screens indicate strengths and negative screens indicate weaknesses of the firm. Each screen is a binary variable that reflects whether the firm meets the particular criterion. These screens are renewed at the end of each calendar year. The screens are summarized in groups of corresponding items referring to a general theme. In the period under review, KLD identifies six themes: Community involvement, corporate governance, diversity, employee relations, environment, and product (see www.kld.com for detailed information about the themes and their strengths and weaknesses). The first relates to how the firm interacts with its social environment. Corporate governance relates to how the firm is governed and directed. Diversity is about the composition of the workforce, especially senior management and the board. Employee relations are about the relationship between the company and its employees and in particular concerns issues related to employee compensation. Environment is about environmental management and policies. Finally, product is about strengths and weaknesses in relation to the quality of the products and production processes of the firm. Apart from these six themes, KLD also investigates companies' behavior with respect to human rights. However, as this is undertaken since 2000 only, we do not include this item in the analysis because it substantially reduces the number of observations in our analysis. Furthermore, KLD has exclusionary screens for alcohol, gambling, firearms, military, nuclear power, and tobacco. Given the nature of these screens, namely focusing only on concerns, it is excluded from our analysis.

Return and accounting data were obtained from Datastream. This includes monthly data on returns, market values, company age, R&D expenditures, net sales, book equity, number of shares outstanding, and net income. The time period covered for these data is June 1992 to July 2006 for our monthly portfolio and Fama-MacBeth (1973) regressions and December 1991 to December 2004 for our yearly

market-to-book regressions. Monthly data is measured at the end of each calendar month and yearly data at the end of each calendar year. The return and accounting data were linked with the KLD data based on ticker and name for the oldest data and on CUSIP code for the more recent data. Finally, the independent variables used in our portfolio regressions, the value-weighted market proxy, the SMB, HML and MOM factors and the risk free rate, were obtained from Kenneth French's website.

To assess the effect of KLD scores about corporate social responsibility on financial performance, we perform three types of analysis. First, we use these scores to form portfolios and assess their performance. Second, we perform a series of Fama-Macbeth (1973) regressions to test the direct effect of KLD scores on excess returns. Finally, we look into the impact of the scores on firm value by means of pooled book-to-market regressions.

2.3.2 SRI Portfolios

Twelve portfolios are formed for the six SRI dimensions based on whether stocks score on the particular strength or concern screen. KLD assigns CSR ratings at the end of each year. To be certain that the social responsibility and financial information for year $t-1$ is known, we calculate returns on monthly equally weighted portfolios beginning in July of year t to June of year $t+1$. Equally weighted portfolios are rebalanced at the beginning of July each year. Portfolios that capture the positive and negative dimension of a certain screen are not mutually exclusive as a portfolio can obtain a positive score on both (see also Driessen and Laeven, 2007, and Chua et al., 2008). The monthly excess returns of the portfolios are assessed using the Fama and French (1993) three-factor model expanded with the Carhart (1997) momentum factor:

$$R_{i,t} - RF_t = \alpha_i + \beta_i(RM_t - RF_t) + s_i SMB_t + h_i HML_t + m_i MOM_t \tag{2.1}$$

where $R_{i,t}$ is the return on portfolio i, constructed as explained above, RM_t is the return in month t on a value-weighted market proxy, RF_t is the return in month t of a one-month treasury bill, SMB_t is the difference in monthly return between a small and large-cap portfolio, HML_t is the difference in return between a value and a growth portfolio, and MOM_t is the monthly return on a portfolio long on past one-year winners and short on past one-year losers. The momentum factor is designed to capture the risk due to the momentum found in stock returns by

Jegadeesh and Titman (1993). Summary statistics on the portfolios and factors are reported in panel A of Table 2.1. In addition to testing the returns on the individual portfolios, we also test the return on a differenced portfolio (see Derwall et al., 2005):

$$R_{t,i,s} - R_{t,i,c} = \alpha_i + \beta_i(RM_t - RF_t) + s_i SMB_t + h_i HML_t + m_i MOM_t \qquad (2.2)$$

where $R_{t,i,s}$ is the return on one of the six strength portfolios and $R_{t,i,c}$ is the return on its accompanying concern portfolio. The independent variables are similar to those in (2.2), except for α_i, which is now the differential excess performance.

The set of portfolio equations is tested in a GMM system as in MacKinlay and Richardson (1991) and Clare et al. (1997). This estimation allows the errors of equations to be correlated. Given the fact that our portfolios are all correlated with a correlation coefficient of .80 or higher, estimating a system is more efficient. Estimating the portfolios in a GMM system as opposed to OLS or SUR also has the advantage of being able to rely on weaker assumptions. In particular, GMM does not rely on the assumption of homoskedasticity and normality of returns. The estimation procedure is to construct a series of errors from equation system 1 so that they are orthogonal to a vector of instruments for all equations. Given the fact that we do not include any conditional information, these instruments consist of a constant and our independent variables. The GMM procedure then chooses parameters $B_0 = [\alpha, \beta, s, h, m]$ that set particular linear combinations of the moment conditions to zero. We define $x_t = [1, (MKT_t - RF_t)', SMB_t', HML_t', MOM_t']$ and $\epsilon_t = R_t - Bx_t$. The moment conditions used by the GMM estimation of B are:

$$E[g_t] = E[x_t \otimes \epsilon] = 0 \qquad (2.3)$$

Since the system is exactly identified, the estimate of B does not depend on the weighting matrix and remains the same as the OLS estimate of the system (Hansen, 1982). However, when there are more restrictions than parameters, the system is overidentified and the moment conditions can no longer be zero for any choice of parameters. In that case, the GMM estimator is identified by minimizing such that is as close to zero as possible. As in Clare et al. (1997), we use a two-stage GMM approach for which we report adjusted t-statistics to take account of small sample bias (Ferson & Foerster, 1994). The adjustment factor with which we multiply the asymptotic variance is $T/(T-P)$, where T is the number of time series observations and P is the number of model parameters. To test whether coefficients

Table 2.1. Summary Statistics

Panel A: Summary statistics equally weighted SRI portfolios

Variable	Mean return (in %)	Standard deviation (in %)
SRI Portfolios		
Community strengths	1.29	3.76
Community concerns	1.24	4.52
Diversity strengths	1.27	4.30
Diversity concerns	1.18	4.91
Employee relations strengths	1.25	4.48
Employee relations concerns	1.26	5.00
Environment strengths	1.16	4.12
Environment concerns	1.16	4.24
Product strengths		4.47
Product concerns		4.05
Governance strengths	1.27	4.49
Governance concerns	1.27	4.70
SRI Difference Portfolios		
Community	0.05	1.85
Diversity	0.09	1.74
Employee relations	0.00	1.24
Environment	0.00	1.15
Product	0.04	1.97
Governance	0.00	2.25
Portfolios used in Fama and French regressions		
Market portfolio	0.64	4.13
SMB portfolio	0.26	3.83
HML portfolio	0.46	3.53
MOM portfolio	0.91	4.99

Notes: This table presents the summary statistics for the returns used in each of the three different analyses. It consists of panel A, B and C. Panel A presents the statistics for the data used in the performance assessment of the portfolios based on individual KLD criteria. First we present the mean and standard deviation of time series returns for portfolios constructed on the basis of different SRI criteria. The SRI difference portfolios represent the time series of differences in return of a portfolio consisting of stocks that score positively on a particular screen and the return of a portfolio consisting of stocks that score negatively. Finally, we present the mean and standard deviation of the time series of returns of the Fama and French factors used in the analysis.

are zero across the system, we use a J-test of overidentifying restrictions (Hansen, 1982). Since we have six instruments for each equation, (2.1) and (2.2) generate 72 and 36 orthogonality conditions, respectively. However, restricting one of the coef-

Panel B: Summary statistics of variables used in Fama-MacBeth Regressions

Variable	Time-Series Average of Means	Time-Series Average of Standard Deviations
Excess return (%)	0.959	10.154
Beta	0.965	0.400
Log(Size) ($\times$ 1,000)	7.910	1.583
Past 12 month returns (%)	1.260	10.156
Turnover (%)	0.566	0.903
Log(B/M)	-0.927	0.704
Log(Age)	3.093	0.504

Notes: This panel reports summary statistics on the variables used in the cross-sectional regressions. The summary statistics of the SRI variables are reported in Panel C. Excess return is a stocks monthly return net of the risk free rate. Beta is a stocks post-ranking Beta calculated following Black et al. (1972). Log(Size) is the logarithm of company market capitalization at the end of month $t-1$. Log (B/M) is the logarithm of the book-to-market ratio at the end of month $t-1$. Return is the simple average of returns during the past 12 months, lagged one month. Turnover is the 1 month lagged monthly average of daily share turnover, which is calculated as average shares traded divided by shares outstanding during month t. Log(Age) is the logarithm of company age measured at the end of the previous year.

ficients to be zero across the system implies that the number of parameters to be estimated is restricted to 60 and 30, respectively. So the system is overidentified. Therefore, we also report J-statistics for iterated GMM estimates since Ferson and Foerster (1994) show that these are more reliable.

2.3.3 Fama-MacBeth regressions

In the previous section, we used portfolios to establish the relationship between risk and returns for groups of stocks satisfying specific SRI criteria. Here, in order to establish the relationship between returns and SRI at the level of individual stocks, we use cross-sectional regressions because we are interested in assessing the direct impact of KLD scores on excess returns. This allows us to identify whether aggregation over different dimensions have confounding effects in identifying the relation between SRI and return. Using the SRI scores defined above along with a host of control variables, we estimate the following regression:

$$R_{i,t} - RF_t = \alpha_i + \beta_1 \boldsymbol{SRI}_{t-1} + \beta_2 Beta_{i,t} + \beta_3 \boldsymbol{X}_{i,t-1} + \epsilon_{i,t} \tag{2.4}$$

where $R_{i,t}$ is the monthly return of stock i in month t and the risk free rate is as defined above. The vector $\boldsymbol{SRI}_{t-1}$ includes the scores of the six SRI variables as

Panel C: Summary statistics of variables used in market-to-book regressions

Variable	Time-Series Average of Means	Time-Series Average of Standard Deviations
Log(B/M)	-0.958	0.693
Community	0.277	0.658
Diversity	0.385	1.083
Employee Relations	0.123	0.822
Environment	-0.139	0.838
Product	-0.127	0.695
Governance	-0.258	0.621
ROE	0.112	0.250
R&D/Sales	0.062	0.846
R&Dmissing	0.607	0.488
Log(Age)	2.730	0.661
Community strengths	0.335	0.634
Community concerns	0.057	0.228
Diversity strengths	0.603	0.954
Diversity concerns	0.218	0.401
Employee relations strengths	0.403	0.641
Employee relations concerns	0.280	0.516
Environment strengths	0.240	0.491
Environment concerns	0.379	0.819
Product strengths	0.150	0.368
Product concerns	0.278	0.590
Governance strengths	0.088	0.284
Governance concerns	0.346	0.524

Notes: All variables are measured at the end of the year. Log(B/M) is the logarithm of book-to-market value. The SRI variables community, diversity, employee relations, environment, product and governance are computed as the sum of all strengths minus the sum of all concern scores within each dimension. R&D Sales is the fraction of sales spent on R&D expenditures in year t. R&D missing is a dummy that is one when the variable R&D/Sales is missing and zero otherwise. Log(Age) is the logarithm of age based on the base date in Datastream and measured at the end of each year. ROE is the return on equity of firm i in year t winsorized to exclude the 0.5% smallest and largest observations.

in the cross-section regressions, but this time measured at the end of year t. $Beta_{i,t}$ is a stock's post-ranking beta estimated using the traditional method of Black et al. (1972). $\boldsymbol{X}_{i,t-1}$ is a vector of control variables similar to those used by Hong and Kacperczyk (2009), including the natural logarithm of firm i's market capitalization at the end of month $t-1$, the logarithm of the book-to-market ratio of stock i at the end of month $t-1$, a firm's simple average of returns during the past 12 months, the

one-month lagged monthly average of daily share turnover in stock i, the natural log of a company's age measured at the end of month t, and a dummy that is equal to one if a stock belongs to the Russell 1000 or 2000 index but not to the S&P 500. This dummy captures any effects that are due to the enlargement of the number of stocks covered by KLD.

2.3.4 Book-to-market regressions

Thirdly, we investigate the impact of KLD scores on the value of the firm by using pooled book-to-market regressions. To this extent, we regress a firm's book-to-market ratio on a set of well-known predictors of book-to-market ratios and the KLD scores we also used in the cross-section regressions. The KLD scores are formed on six dimensions adding a point when a stock scores positively on a dimension's strength screen and subtracting a point when a stock scores negatively on a dimension's strength screen. The most elaborate specification we test is the following:

$$log(B_{i,t}/M_{i,t}) = c_0 + \boldsymbol{c_1 SRI_{i,t}} + \boldsymbol{c_2 X_{i,t}} + \epsilon_{i,t} \tag{2.5}$$

where $log(B_{i,t}/M_{i,t})$ is the logarithm of the book-to-market ratio of stock i at the end of year t. $\boldsymbol{SRI_{i,t}}$ is a vector containing the (lagged) scores of the six SRI variables; Community, corporate governance, diversity, environment, employee relations and product. The vector $\boldsymbol{X_{i,t}}$ includes several control variables known to correlate with the book-to-market. It includes the fraction of sales spent on R&D expenditures in year t, a dummy that is one when the fraction of sales spent on R&D is missing, and a dummy for stocks belonging to the Russell 1000 or 2000, but not the S&P500. Furthermore, we included the logarithm of age as measured at the end of each year. Finally, ROE is the return on equity winsorized to exclude the 0.5% smallest and largest observations. Summary statistics are in panel B of Table 2.1.

In estimating (2.5), we do not use Fama-Macbeth (1973) regressions but pooled OLS with robust standard errors. This is because the Fama-Macbeth approach to estimating panel data is useful in adjusting for correlation in the cross-section, but understates standard errors when the dependent variable is correlated across time (Bali et al., 2008; Petersen, 2009). Return data only suffer from minor autocorrelation. However book-to-market ratios are very much correlated across time and in the cross-section. Therefore, we cluster standard errors both by firm and by time (see Thompson, 2011). However, these standard errors are only asymptotic-

ally correct. For the standard errors clustered by firm this poses no problem, but for standard errors clustered by time it does since we only have 14 years of data. Therefore, we also report a second specification in which we adjust for autocorrelation by including time dummies.

2.4 Results

2.4.1 SRI portfolios

In Table 2.2 we present the results of the Fama and French regressions on the portfolios screened for different SRI characteristics. From panel B we see that none of our SRI portfolios shows significant outperformance, although we should note that the adjustment factor is largely responsible for this result. The necessity of adjusting for a large number of factors and equations can also be seen by comparing the two J-statistics that test whether a coefficient is zero across all equations. J-statistic A was computed using 2-stage GMM, whereas J-statistic B was computed using iterated GMM. Following J-statistic A, we should reject the hypothesis of zero alpha across all equations at the 5% level, whereas following J-statistic B we would not be able to reject this hypothesis. Although we know from Ferson and Foerster (1994) that iterated GMM has somewhat lower power than 2-stage GMM, the difference in p-values is large enough to suspect that 2-stage GMM leads to overrejection.

Looking at the difference between SRI strength and concern portfolios has the advantage of reducing the dimensionality of the system, next to being able to assess the difference in factor exposure and performance between strength and concern portfolios. Concerning the first, the difference in p-values of the J-statistics is much smaller so that using different estimation methods does not lead to qualitatively different results. The results in panel A clearly show that there are significant differences between the strength and the concerns portfolios. We find that for four out of six portfolios, the strength portfolio has a significant lower exposure to the HML factor than its accompanying concern portfolio, suggesting that these strength portfolios are more growth oriented than their concern counterparts. Finally, only the community strength portfolio significantly outperforms its accompanying concern portfolio. The excess return associated with this portfolio is 3.4% (significant at the 10% level). The fact that CSR strength portfolios have lower book-to-market ratios than their concern counterparts is consistent with overpricing. However, overpricing of CSR strength stocks vis-à-vis concern stocks also implies that CSR strength portfolios should have lower performance than concern portfolios, which is clearly

not the case and even opposite for the community difference portfolio. To further investigate the pricing impact of SRI scores, we perform book-to-market regressions.

2.4.2 Fama-MacBeth regressions

Looking at the results in Table 2.3, we observe that beta and turnover are not significant and that book-to-market and age have a negative effect on subsequent returns. Although size and past returns do not have a significant influence, they do have the expected signs. Concerning the SRI scores, we see that only the employee relations score has a significant positive effect on excess returns. Although it is only significant at the 10% level, the result is robust to the inclusion of numerous control variables, including 39 industry dummies. In fact, the effect of most controls is consistent with those in Hong and Kacperczyk (2009). The outperformance of about 0.07% per month per point implies that stocks obtaining the maximum (minimum) score on employee relations of 4 (-4) will outperform (underperform) other stocks in the sample by 3.4% on an annual basis.

However, if we assume that the whole of the employee relations score is no more than the sum of its parts - the employee relations subscores - it can be imagined that some subscores have a significant effect on excess returns, whereas others have no effect at all. To disentangle the effect of the subscores we re-estimate equation (2.5), this time including the subscreens of employee relations as dummy variables. A description of the subscreens and the results are provided in 2.4. The size and significance of the dummies suggests that a few subscreens are responsible for the effect of employee relations on excess returns.

Second, only the concern subscreens have a significant effect on excess returns, but they are not consistent in sign. This implies that simply adding the subscores of strength and concern screens to form overall scores leads to confounding effects. It also suggests that what is a concern from a social perspective is not always a concern from a financial perspective. This is due to the fact that social objectives do not always align with firm financial objectives. For instance, Pension/Benefit Concern (Empcon. D) has a positive influence on future excess returns. A possible explanation is that companies that have substantially underfunded defined benefit pension plan have lower future financial obligations and therefore higher expected future cash flows. However, both objectives can be aligned too. For example, when we look at Poor union relations (Empcon. A) and Workforce reductions

Table 2.2. Carhart regressions, equally weighted KLD portfolios, July 1992 - June 2006

Panel A: KLD Difference Portfolios

	α	t-stat	RM - RF	t-stat	SMB	t-stat	HML	t-stat	MOM	t-stat
Community	3.40%*	1.852	-0.15^{***}	-3.06	-0.04	-0.77	-0.30^{***}	-4.49	0.01	0.46
Diversity	2.36%	0.95	-0.08	-1.47	-0.25^{***}	-4.64	-0.1	-1.66	0.06	0.92
Employee Rel.	0.50%	0.305	-0.07^{**}	-2.08	-0.11^{***}	-3.2	-0.11^{***}	-2.09	0.09^{**}	2.39
Environment	1.13%	0.879	-0.04	-1.52	0.06	1.82	-0.11^{***}	-3.01	-0.03	-1.11
Product	1.48%	0.767	0.02	0.46	0.15^{**}	2.55	-0.25^{***}	-4.01	-0.02	-0.79
Governance	0.92%	0.368	-0.21^{***}	-4.22	0.25^{***}	3.89	0.1	1.29	-0.06	-1.2
J-stat A	9.37		24.96		27.38		23.67		14.44	
	(-0.154)		(0.000)		(0.000)		(0.000)		(-0.025)	
J-stat B	7.32		18.59		19.38		13.22		6.36	
	(-0.293)		(-0.005)		(-0.004)		(-0.040)		(-0.384)	

Panel B: KLD Strength and Concern Portfolios

	α	t-stat	RM - RF	t-stat	SMB	t-stat	HML	t-stat	MOM	t-stat
Community Strengths	3.01%	1.45	0.95^{***}	17.66	0.01	0.17	0.49^{***}	4.41	-0.11	-1.77
Community Concerns	-0.38%	-0.13	1.10^{***}	13.4	0.05	0.92	0.79^{***}	5.89	-0.13	-1.48
Diversity Strengths	3.37%	1.34	1.03^{***}	24.25	0.1	1.58	0.39^{***}	3.53	-0.20^{**}	-2.5
Diversity Concerns	0.99%	0.28	1.11^{***}	15.41	0.36^{***}	3.66	0.49^{***}	3.96	-0.26^{***}	-4.87
Employee Strengths	2.68%	1.22	1.07^{***}	24.46	0.20^{***}	3.14	0.43^{***}	4.43	-0.22^{***}	-4.12
Employee Concerns	2.18%	0.73	1.13^{***}	19.79	0.31^{***}	4.03	0.54^{***}	4.05	-0.31^{***}	-3.33
Environment Strengths	0.63%	0.25	0.99^{***}	14.9	0.21^{***}	3.18	0.64^{***}	6.43	$-0.19**$	-2.56
Environment Concerns	-0.50%	-0.19	1.03^{***}	15.41	0.16^{***}	2.5	0.75^{***}	5.95	-0.16	-1.81
Product Strengths	3.42%	1.19	1.03^{***}	18.27	$0.18*$	1.86	0.33^{***}	3.05	-0.19^{***}	-3.5
Product Concerns	1.91%	0.95	1.01^{***}	20.34	0.04	0.63	0.58^{***}	4.67	-0.17^{**}	-2.64
Governance Strengths	4.21%	1.14	0.91^{***}	12.13	0.38^{***}	3.11	0.46^{***}	2.69	-0.29^{***}	-3.87
Governance Concerns	3.27%	1.38	1.11^{***}	28.98	0.13^{**}	2.14	0.36^{***}	3.57	-0.23^{***}	-4.39
J-stat A	24.32		32.12		31.84		30.96		29.8	
	-0.018		-0.001		-0.001		-0.002		-0.003	
J-stat B	12.91		NA		22.56		19.67		14.07	
	-0.376				-0.032		-0.074		-0.296	

Notes: In Panel A we estimated for all portfolios in a system GMM framework the regression: $R(t) - RF(t) = \alpha + \beta[RM(t) - RF(t)] + sSMB(t) + hHML(t) + mMOM(t) + e(t)$. T-statistics are adjusted following Ferson and Foerster (1994) by a factor T/(T-P), where T is the number of time periods and P the number of parameters. We also test the overidentifying restriction generated by the remaining orthogonality conditions when we set one of the parameters to zero for all of the portfolios. For panel A Hansen (1982) shows this test is distributed as $\chi^2(6)$ under the null since in the restricted case the number of instruments is 36 and the number of parameters is 30 so that the number of over identifying restrictions is 6. In panel B this test is distributed as $\chi^2(12)$ under the null since in the restricted case the number of instruments is 72 and the number of parameters is 60 so that the number of over identifying restrictions is 12. J-stat A and B denote J-stats from 2-Stage and Iterated GMM, respectively. P-values are reported between brackets. The covariance matrix is computed following Newey and West (1994). The Bartlett spectral density kernel estimator is used to make sure that the covariance matrix is positive semi-definite. For a description of the variables we refer to Table 1. * 10% significance ** 5% significance *** 1% significance. Not available due to the fact that the iterated estimation procedure did not converge.

Table 2.3. Cross-section regressions of excess stock returns, July 1992 - June 2006

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Community	-0.003	0.011	0.003	0.006	0.058	0.068	0.042
	-0.070	0.231	0.062	0.136	1.26	1.469	0.870
Diversity	-0.049	-0.022	-0.012	-0.009	-0.012	-0.017	-0.046
	-1.018	-0.482	-0.268	-0.202	-0.276	-0.384	-1.104
Employee Rel.	0.045	0.060	0.077*	0.079*	0.071*	0.074*	0.073*
	1.028	1.439	1.897	1.951	1.826	1.908	1.720
Environment	0.020	-0.005	0.034	0.037	0.049	0.038	-0.046
	0.377	-0.09	0.644	0.698	0.942	0.732	-1.105
Product	0.042	0.010	0.062	0.053	0.049	0.045	0.058
	0.615	0.145	0.954	0.850	0.811	0.749	1.146
Governance	0.014	-0.095	-0.093	-0.094	-0.110	-0.104	-0.098
	0.173	-1.183	-1.200	-1.221	-1.479	-1.390	-1.312
Beta	0.163	0.037	0.032	0.019	-0.108	-0.124	-0.327
	0.304	0.070	0.061	0.038	-0.237	-0.271	-0.776
Log(Size)		-0.117	-0.015	-0.019	-0.031	-0.015	-0.041
		-1.913**	-0.240	-0.314	-0.507	-0.251	-0.713

Table 2.3. Cross-section regressions of excess stock returns, July 1992 - June 2006 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(B/M)			-0.477***	-0.482***	-0.591***	-0.608***	-0.637***
			-4.109	-4.19	-5.570	-5.756	-6.481
Returns				0.025	0.014	0.011	0.016
				0.777	0.441	0.351	0.613
Turnover					0.330*	0.291	0.107
					1.803	1.591	0.687
Log(Age)						-0.512***	-0.374**
						-3.779	-2.658
Dummy Russell						-0.002	-0.024
						-0.044	-0.573
Dummy Industry	N	N	N	N	N	N	Y

Notes: This table reports the results on the Fama-MacBeth (1973) regressions. Variables are described in Table 1. Dummy Industry Y (N) indicates whether 39 industry dummies have (have not) been included in the specification. T-statistics are reported below the coefficients. * 10% significance ** 5% significance *** 1% significance***

Table 2.4. Cross-section regressions of excess returns on employee relations screens

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Empstr. A	0.037	0.072	-0.13	-0.154	-0.126	-0.076	-0.025
	0.136	0.258	-0.469	-0.559	-0.458	-0.279	-0.092
Empstr. C	0.088	0.07	0.056	0.077	0.066	0.067	0.083
	0.598	0.468	0.374	0.53	0.489	0.497	0.712
Empstr. D	-0.015	0.015	0.093	0.107	0.15	0.148	0.27
	-0.12	0.123	0.816	0.927	1.349	1.343	2.437**
Empstr. F	-0.083	-0.002	-0.007	-0.007	-0.009	0.018	0.05
	-0.539	-0.016	-0.048	-0.05	-0.066	0.125	0.407
Empstr. X	-0.098	-0.057	-0.032	-0.03	-0.024	-0.03	0.105
	-0.676	-0.397	-0.218	-0.211	-0.169	-0.213	0.74
Empcon. A	-0.544***	-0.485***	-0.394**	-0.355**	-0.325*	-0.322*	-0.158
	-3.102	-2.762***	-2.278**	-2.070**	-1.874	-1.874	-0.982
Empcon. B	-0.119	-0.173	-0.324	-0.314	-0.209	-0.176	-0.038
	-0.553	-0.79	-1.528	-1.478	-1.029	-0.865	-0.199
Empcon. C	-0.425*	-0.494**	-0.506**	-0.492**	-0.464**	-0.412*	-0.402*
	-1.829	-2.155	-2.224	-2.215	-2.144	-1.9	-1.854

Table 2.4. Cross-section regressions of excess returns on employee relations screens (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Empcon. D	0.245	0.254	0.262*	0.282*	0.342**	0.333**	0.355**
	1.539	1.594	1.697	1.824	2.268	2.237	2.321
Empcon. X	-0.005	0.052	-0.288	-0.338	-0.307	-0.383	-0.329
	-0.018	0.199	-1.132	-1.317	-1.204	-1.455	-1.421
Community	0.005	0.017	0.015	0.022	0.071	0.08	0.049
	0.097	0.341	0.307	0.44	1.492	1.666	0.966
Diversity	0.005	0.026	0.044	0.046	0.041	0.035	0.007
	0.096	0.531	0.943	1.021	0.927	0.787	0.179
Environment	-0.032	-0.054	-0.017	-0.011	0.012	0.005	-0.063
	-0.612	-1.022	-0.329	-0.213	0.237	0.095	-1.504
Product	0.055	0.027	0.077	0.067	0.064	0.06	0.067
	0.815	0.395	1.187	1.085	1.057	0.996	1.253
Other	-0.009	-0.126	-0.117	-0.118	-0.128	-0.119	-0.113
	-0.121	-1.597	-1.46	-1.479	-1.664	-1.553	-1.483
Beta	0.068	-0.054	-0.067	-0.079	-0.179	-0.198	-0.349
	0.121	-0.098	-0.122	-0.148	-0.371	-0.41	-0.777

Table 2.4. Cross-section regressions of excess returns on employee relations screens (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(Size)		-0.114	-0.005	-0.006	-0.019	-0.005	-0.034
		-1.723*	-0.076	-0.09	-0.289	-0.073	-0.549
Log(B/M)			-0.471***	-0.471***	-0.585***	-0.599***	-0.663***
			-4.114	-4.074	-5.445	-5.591	-6.498
Returns				0.015	0.003	0.001	0.009
				0.474	0.104	0.041	0.324
Turnover					0.252	0.215	0.045
					1.386	1.185	0.292
Log(Age)						-0.484***	-0.356**
						-3.388	-2.486
Dummy Russell						-0.008	-0.03
						-0.164	-0.68
Dummy Industry	N	N	N	N	N	N	Y

Notes: This table reports the results on the Fama-MacBeth regressions. Variables are described in Table 1. Dummy Industry Y (N) indicates whether 39 industry dummies have been (have not been) included in the specification. Empstr. A: strong union relations. Empstr. C: Cash Profit Sharing. Empstr. D: Employee Involvement. Empstr. X: Other Strenghts. Empcon A: Poor Union Relations. Empcon. B: Safety Controversies. Empcon C: Workforce Reductions. Empcon D: Pension/Benefits Concern. Empcon. X: Other Concern. T-statistics are reported below the coefficients. * 10% significance ** 5% significance *** 1% significance

(Empcon. C), we find that both have a negative effect on future excess returns, possibly because they influence future cash flows negatively. Furthermore, the results in column (7) suggest that the effect of poor union relations is industry specific.

The results for Employee relations are in line with other findings in the literature. For instance, Carroll and Niehaus (1998) show that increases in unfunded pension liabilities worsen firms' debt ratings, implying that investors require a premium for firms with underfunded pension liabilities. In addition, pension obligations (Mittelstaedt and Warshawsky, 1993) and liabilities for retiree health benefits (Feldstein and Seligman, 1981) both are negatively related to share prices. Hallock (1998) shows that layoffs are negatively related to cumulative abnormal returns (CARs). Concerning union relations, evidence suggests that union coverage is negatively related to market value (Hirsch, 1991). Finally, to the extent that poor union relations increase the potential for strikes, we note that Kramer and Vasconcellos (1996) and Persons (1995) find that strike announcements result in negative CARs.

2.4.3 Book-to-market regressions

In the book-to-market regressions, we measure both the dependent variable as well as the SRI scores at the end of the year. Thus, the regressions indicate whether a stock that has a low book-to-market ratio scores high on one of the SRI dimensions. In Table 2.5 we present the results for two different specifications of the model. In column (1), we report results for the specification including time dummies, clustering standard errors by firm. In column (2), we report results for the specification where we cluster standard errors by both year and firm. The SRI scores are the sum of scores on strength screens minus the sum of scores on concern screens. Based on the predictions from the theoretical literature, we expect that SRI scores are negatively related to book-to-market ratios.

Consistent with theory, we see in Table 2.5 that diversity, environment, and product have a significant negative impact on the book-to-market ratios. Interestingly, governance scores have a significant positive effect on book-to-market. In panel B we estimated the model by treating the strengths and concerns scores separately. Here, we observe that the concern part of the governance score mainly drives the results, suggesting that the market values excessive remuneration of top executives, although causation might run both ways. Another interesting finding in panel B is that both environment strength and concern scores are positively related to book-to-market ratios. This may be due to the fact that the total score can

Table 2.5. Pooled book-to-market regressions

Panel A: Overall KLD scores

Variable	(1)	T-stat	(2)	T-stat
Community	−0.016	−0.830	−0.007	−0.311
Diversity	−0.047***	−4.440	−0.052***	−4.589
Employee Relations	−0.014	−1.14	−0.013	−0.937
Environment	−0.042***	−3.250	−0.047***	−3.467
Product	−0.046**	−2.21	−0.044**	−2.23
Governance	0.108***	7.08	0.118***	4.187
R&D/Sales	−0.005***	−3.00	−0.005***	−3.595
R&DMissing	0.288***	11.74	0.289***	9.596
Russell3000	−0.001	−0.030	−0.036	−0.915
LogAge	0.098***	7.08	0.093***	7.063
ROE	−0.567***	−7.94	−0.575***	−2.964
Dummy Time	Y		N	

Panel B: Strength and Concern KLD scores

Variable	(1)	T-stat	(2)	T-stat
Community Strengths	−0.004	−0.2	−0.007	−0.28
Community Concerns	0.054	1.64	0.049	1.28
Diversity Strengths	−0.052***	−3.85	−0.066***	−4.18
Diversity Concerns	0.048	2.46	0.045	1.71
Employee Relations Strengths	−0.013	−0.71	−0.017	−0.78
Employee Relations Concerns	0.023	1.5	0.044	1.35
Environment Strengths	0.080***	3.41	0.090***	3.39
Environment Concerns	0.074***	5.52	0.081***	5.7
Product Strengths	−0.097***	−2.63	−0.085**	−2.25
Product Concerns	0.017	0.69	0.014	0.48
Governance Strengths	0.018	0.65	−0.001	−0.03
Governance Concerns	−0.154***	−8.28	−0.174***	−5.44
R&D/Sales	−0.005***	−3.05	−0.003***	−5.16
R&DMissing	0.299***	12.25	0.287***	8.98
Russell3000	−0.008***	−0.23	0.009	0.25
LogAge	0.085***	6.13	0.063***	4.24
ROE	−0.555***	−7.83	0.003***	2.99
Dummy Time	Y		N	

Notes: This table reports the results of the pooled book-to-market regressions. The variables are described in 2.1. In column (1) we report results for the specification including time dummies, clustering standard errors by firm. Dummy Time Y (N) indicates whether 14 year dummies have (have not) been included in the specification. In column (2) we report results for the specification where we cluster by both year and firm. * 10% significance ** 5% significance *** 1% significance

be zero when strength and concern screens are of equal size. Finally, it is interesting to note that the overall community and employee relations scores as well as their strength and concern sub scores do not have a significant effect on book-to-market ratios. Stated otherwise, stocks obtaining high scores on community and employee relations do not appear to be overpriced relative to other stocks.

2.4.4 Robustness

To check the robustness of our results, we performed some additional analyses which are available on request. As a first check, we repeated our portfolio regressions but this time using value-weighted returns instead of equally weighted returns. Results remain qualitatively the same, except that the difference portfolio of employee relations yields a significant positive excess return of 2.81% annually, whereas the excess return on the community difference portfolio becomes insignificant. A second check is splitting the employee relations score into its 10 subscores for the book-to-market regressions. In contrast to the return regressions in 2.4, only some of the strength scores significantly influence book-to-market ratios, while the concern scores do not influence book-to-market ratios. A notable exception to this is Workforce reduction (Empcon C.), which does significantly influence book-to-market ratios.

2.5 Conclusions

In this chapter we set out to investigate the effect of socially responsible investment (SRI) on stock returns. In particular, we want to clarify the existing difference between the theoretical literature suggesting a relation between SRI and stock returns, and most of the empirical literature that does not find a significant relationship. We suggest two potential explanations for this puzzle. First, the aggregate analysis of SRI scores may eliminate a relationship if individual dimensions of SRI have opposite effects on performance. We find little evidence for this explanation. Second, in analyzing stock returns, most researchers control for risk using Fama and French regressions, including the HML factor measuring the sensitivity of a stock to the return difference of stocks with high and low book-to-market ratios. SRI results in lower book-to-market ratios, and as a result, the alphas do not capture SRI effects. We establish that SRI - in particular portfolios that score positive on diversity, environment and product - has a significant impact on stock returns.