Journal of Banking & Finance 47 (2014) 54–62

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Does it pay to be ethical? Evidence from the FTSE4Good

Yacine Belghitar [a], Ephraim Clark [b], Nitin Deshmukh [b,*]

[a] Cranfield School of Management, Bedford MK43 0AL, United Kingdom
[b] Middlesex University Business School, Department of Accounting & Finance, London NW4 4BT, United Kingdom

A R T I C L E I N F O

Article history:
Received 24 September 2013
Accepted 29 June 2014
Available online 8 July 2014

JEL classification:
G1
G11

Keywords:
Socially Responsible Investments
SRI
Marginal Conditional Stochastic Dominance

A B S T R A C T

The empirical mean–variance evidence comparing the performance of Socially Responsible Investments (SRI) and conventional investments suggests that there is no significant difference between the two. This paper re-examines the problem in the context of Marginal Conditional Stochastic Dominance (MCSD), which can accommodate any return distribution or concave utility function. Our results provide strong evidence that there is a financial price to be paid for socially responsible investing. Indices composed of socially responsible firms are MCSD dominated by trademarked indices composed of conventional firms as well as by indices carefully matched by size and industry with the firms in the SRI indices. Zero cost portfolios created by shorting the SRI index and using the proceeds to invest in the conventional index generate higher average returns, lower variance and higher skewness than either of the two indices standing alone. They also MCSD dominate the SRI and conventional indices standing alone.

1. Introduction

Socially Responsible Investments (SRI) have a long and noble history, but have only grown to prominence in the recent past. In the beginning, SRI was treated as a fad by the wider finance community, a fad that would either disappear or confine itself to the fringes (Renneboog et al., 2008a). As of 2014, however, SRI investments account for 11% ($3.74 trillion out of $33.7 trillion) of assets under management in the US,[1] and 27% (£1.235 trillion[2] out of £4.5 trillion[3]) of assets under management in the UK. Such widespread prominence puts it in a position that warrants closer scrutiny.

Given the importance of the sector and its implications for resource allocation, the question we ask in this paper is whether there is a price to be paid for restricting investment opportunities to the SRI subset of the overall investment opportunity universe. Mean–variance theory suggests that reduced diversification opportunities should be reflected in inferior investment performance. But, a close look at the literature, which is reviewed in the following section, shows that there is no conclusive evidence that this is the case. These studies, however, are subject to serious shortcomings in how performance has been measured and tested.

Some studies have compared the performance of SRI funds with conventional funds (for example: Hamilton et al., 1993 and Bauer et al., 2005). This approach ignores the fact that the difference in performance may arise due to other factors like fund size, age, investment universe, etc. To overcome these problems, others, such as Mallin et al. (1995), Gregory et al. (1997), and Kreander et al. (2005), used a matched pair approach, i.e., they first matched the SRI funds with similar conventional funds using the criteria of size, age, investment universe and country. This approach, although an improvement, ignores the fact that differences in performance may be due to differences in the ability of fund managers rather than the nature of the investments (SRI vs conventional). Statman (2000, 2006) and Schroder (2007) provide a solution to this issue by comparing the performance of SRI indices with conventional indices based on the argument that indices are immune to biases associated with specific funds, such as management quality, operating costs, size, age, etc. and hence serve to isolate the impact of the SRI factor on performance.

All of the foregoing studies suffer from a common weakness. Performance measurement has been limited to the first two moments of equity return distributions and testing has concentrated on differences in first moments (equity returns) or some form of the mean–variance (MV) framework, often, but not always, based on the capital asset pricing model (CAPM). Although it is intuitively attractive and widely accepted throughout the financial profession, the MV framework is only a special case of expected utility maximization, which lies at the heart of modern investment theory and practice and, in its most general form, considers all

* Corresponding author. Tel.: +44 208 411 4175.
E-mail addresses: yacine.belghitar@cranfield.ac.uk (Y. Belghitar), e.clark@mdx.ac.uk (E. Clark), n.deshmukh@mdx.ac.uk (N. Deshmukh).

[1] The Forum for Sustainable and Responsible Investment is the US (2014).
[2] The UK Sustainable Investment and Finance Association (2014).
[3] Investment Management Association (2014).

http://dx.doi.org/10.1016/j.jbankfin.2014.06.027
0378-4266/© 2014 The Authors. Published by Elsevier B.V.

moments of return distributions. By neglecting higher moments of return distributions, potentially pertinent information on performance is being eliminated, while tests of performance based on specific asset pricing models may have more to say about the models themselves than about performance. There are no studies that generalize the measurement of performance by considering all the moments of the distributions of equity returns and test the effects directly, that is, outside the context of a specific asset pricing model. This paper is a first step to fill this gap.

We argue that besides mean and variance, performance measures should reflect the third and higher moments of equity return distributions. There are strong reasons to believe that third moments and higher are important determinants of performance. First of all, it is well known that the first and second moments are only appropriate for quadratic utility maximisers or normally distributed returns. It is also well known that quadratic utility functions have many shortcomings[4] and it is a well documented fact since Mandelbrot (1963) that asset returns are generally not normally distributed. More importantly, it has been shown that the third and the fourth moments of return distributions – skewness and kurtosis, respectively – do matter to investors, who show a preference for positive skewness and an aversion to kurtosis (see, Kraus and Litzenberger, 1976; Fang and Lai, 1997; Dittmar, 2002; Post et al., 2008). Clark and Kassimatis (2013) show that diversification opportunities increase significantly when all moments of return distributions are considered.

With this in mind, we use indices and an innovative performance measure to compare SRI and conventional investments. More specifically, we use the FTSE4Good Index Series as the SRI investment universe and the concept of Marginal Conditional Stochastic Dominance (MCSD) developed by Shalit and Yitzhaki (1994) to estimate investment performance. Under the general assumption that investors are risk averse, MCSD provides the probabilistic conditions under which all risk-averse investors prefer one risky asset over another. In the terminology of stochastic dominance, MCSD provides the tools to assess the "dominance" or superiority of one asset over another. Dominance means that the utility of all risk averse investors can be improved by increasing the share of the dominant asset at the expense of the dominated asset.[5]

There are no assumptions regarding the efficiency of the global market portfolio or the distributions of equity returns. The only assumption is that investors are risk averse and that part of their investment decision process is to improve the return distribution of their portfolios, i.e., they diversify but do not necessarily aim to create efficient portfolios in the sense of Markowitz portfolio optimization. MCSD tells us if investors will prefer an index because it can improve their portfolio's characteristics, or if they avoid it because it affects their portfolio negatively.

In this paper we compare the four socially responsible FTSE4-Good indices with other trademarked conventional indices as well as with indices composed of conventional firms carefully matched to the firms in the FTSE4Good indices. In the major contribution of this paper, our results show that although there is nothing to be gained or lost from socially responsible investing in terms of mean and variance, there is a high price to be paid in investor utility when the higher moments of the return distributions are taken into consideration. In four of six comparisons with trademarked conventional indices, the FTSE4Good indices are MCSD dominated by the trademarked conventional indices. Importantly, they are also dominated in all four of the comparisons with indices composed of conventional firms carefully matched to the firms in the FTSE4Good indices. These results are evidence that risk averse investors can improve their utility by reducing holdings of FTSE4Good indices and purchasing conventional ones. We test this proposition by constructing zero cost portfolios created by shorting the SRI index and using the proceeds to invest in the conventional index. In all cases, these zero cost portfolios generate higher average returns, lower variance and higher skewness than either of the two indices standing alone. Most importantly, they also MCSD dominate the SRI and conventional indices standing alone, which confirms the proposition that risk averse investors can improve their expected utility by reducing holdings of SR firms and purchasing conventional stocks.

We proceed as follows. Section 2 discusses the relevant literature. Section 3 describes the data. Section 4 presents the MCSD and MV methodology used in the testing. Section 5 presents the results and Section 6 concludes.

2. Previous related work

There is a stream of research that compares the performance of SRI stocks with conventional stocks. Some of this research provides evidence that investors in stocks with high SRI ratings are at an advantage. For example, Kempf and Osthoff (2007) using the Carhart (1997) four factor model find that a strategy of buying stocks with high SRI scores and selling those with low SRI scores produces an abnormal return of up to 8.7% per year. Chan and Walter (2014), who also use the Carhart (1997) four factor model, report a "green premium" of 7% for environmentally "friendly" firms. However, Statman and Glushkov (2009) using a similar approach find that the performance of SRI stocks is not statistically different from conventional ones. On the other hand, Hong and Kacperczyk (2009), also using the Carhart four factor model find that a portfolio comprised of "sin stocks", i.e., alcohol, tobacco and gaming, significantly outperforms similar comparable stocks, thus implying that investors in SRI stocks seem to be at a disadvantage. However, after controlling for managerial skills, transaction costs and fees, Humphrey and Tan (2013) find no outperformance of portfolios that include "sin" stocks.[6]

A number of other studies have compared the performance of SRI funds with conventional funds. Some of these used only the CAPM to compare performance (for example: Hamilton et al., 1993; Mallin et al., 1995; Goldreyer et al., 1999) while others have used multifactor models (for example: Amenc and Sourd, 2008; Bauer et al., 2005; Fernandez and Matallin, 2008; Geczy et al., 2005; Gregory et al., 1997; Kreander et al., 2005; Renneboog et al., 2008b; Munoz et al., 2013). However, none of these studies found any statistically significant difference in performance between the SRI and conventional funds.

Recognizing that differences in performance may be due to differences in the portfolio construction process or in the ability of fund managers rather than the nature of the investments themselves, some authors have compared the performance of indices. Two of the first studies by Sauer (1997) and Statman (2000) compared the performance of the Domini Social Index (an SRI or screened version of the conventional S&P 500) with the S&P 500. They used the Sharpe ratio and the CAPM to estimate Jensen's alpha for the comparison and found no significant difference in the performance of the two indices. Statman (2006) extended his earlier (2000) study and compared the performance of four popular SRI indices with the S&P 500 index. The four SRI indices used were: Domini Social Index, Calvert's Social Index, Citizen's Index and Dow

[4] For example, third derivatives and higher are equal to zero or do not exist, which rules out prudent and temperant behaviour. For a discussion of prudence and temperance see Eeckhoudt and Schlesinger (2006).

[5] The size of the diversification adjustment can also be calculated (see: Clark and Jokung, 1999). Shalit and Yitzhaki (2010) show how MCSD rules can be easily applied for portfolio choices. In this paper we are only interested in identifying dominance.

[6] See Derwall et al. (2011) for an excellent and detailed review of SRI stock performance studies. See Malik (2014) for a review of the literature on corporate social responsibility and firm value.

Table 1
List of indices used in this study.

Index type	Index names	Country	Currency
SRI (S)	FTSE4GOOD-UK-50	UK	£
Conventional (C)	FTSE-100	UK	£
Market (M)	FTSE-ALL SHARE	UK	£
SRI (S)	FTSE4GOOD-UK-50	UK	£
Conventional (C)	FTSE-250	UK	£
Market (M)	FTSE-ALL SHARE	UK	£
SRI (S)	FTSE4GOOD-US-100	US	$
Conventional (C)	S&P-100	US	$
Market (M)	D J Total Stock Market Index US	US	$
SRI (S)	FTSE4GOOD-US-100	US	$
Conventional (C)	DJ Ind. Average	US	$
Market (M)	D J Total Stock Market Index US	US	$
SRI (S)	FTSE4GOOD-EU-50	EU	€
Conventional (C)	STOXX-50	EU	€
Market (M)	STOXX-TM	EU	€
SRI (S)	FTSE4GOOD-GLOBAL-100	GLOBAL	$
Conventional (C)	S&P-GLOBAL-100	GLOBAL	$
Market (M)	FTSE-ALL WORLD	GLOBAL	$

The indices are grouped as follows: 1 SRI index, 1 conventional index and 1 market index within each set. The market index is used as the market portfolio in the MV approach and for wealth ranking in the MCSD approach. Within each group we ensure that weekly values for all the 3 indices included are collected in the same currency.

Jones Sustainability US Index. This study also had a larger time horizon extending up to 2004, but, as with his previous paper, was limited to the US. He found evidence that the returns of the SRI indices exceeded the returns of the S&P 500, but the results were not statistically significant. Schroder (2007) was the first study on this topic to look outside the US. He studied the performance of 29 SRI indices worldwide. Using the simple CAPM to estimate alpha as the performance parameter, he found no significant evidence of under/over performance.[7]

The upshot of all this is that there is no conclusive evidence that there is anything to be gained or lost from socially responsible investing.

3. Data and sample description

The FTSE4Good series covers four geographical regions: US, UK, Europe and Global. It has one tradable index for each region. The FTSE4Good advisory committee decides whether a company is "responsible" enough to be included in the index series. Broadly speaking they look at the following issues: corporate social responsibility, non-discriminatory labour policies, fair stakeholder practices, environmental sustainability and transparent management. These SRI indices contain the largest 50 or 100 companies in the region and are thus basically "SRI screened" versions of the more popular conventional indices like the FTSE100.

Following Sauer (1997) and Statman (2000, 2006) we compare the performance of the FTSE4Good series with a similar conventional index and use relevant benchmarks to represent the parent market portfolio for both the SRI as well as conventional indices.[8]

All the trademarked indices used in this study are listed in Table 1. Since we will be making pair wise comparisons, the indices are grouped together with each group consisting of one market index, one SRI index and one conventional index. The FTSE4Good index series has 4 tradable indices. We compare these SRI indices with similar conventional ones. This gives us 6 groups and in all 14 individual indices. The market index is used as the market portfolio in the MV approach and for wealth ranking in the MCSD approach.[9]

We collect weekly data from DataStream for all the indices. We also collect weekly data for the risk-free rates in the currency that matches the currency of the indices in each of the 6 groups. Within each group we ensure that weekly values for all 3 indices are collected in the same currency. The study period starts from July 2001, i.e., when the FTSE4Good index series was launched, and ends at November 2010. This gives us almost 10 years of weekly data amounting to 488 observations. We then calculate the weekly returns for each index using the following formula:

$$R_{i,t} = \left(\frac{P_{i,t}}{P_{i,t-1}}\right) - 1 \tag{1}$$

where $R_{i,t}$ = Return for index i in week t; $P_{i,t}$ = Closing value for index i in week t; $P_{i,t-1}$ = Closing value for index i in week $t-1$.

Table 2 shows descriptive statistics of the weekly return series for all the indices included in this study.

Looking at the raw mean returns in Table 2, we find that the FTSE4Good indices underperform their conventional counterparts in all six cases. In five out of those six cases the FTSE4Good indices also have higher risk as estimated using standard deviation. All the indices in the sample have negative skewness and excess kurtosis that are significant at the 5% level. We note again that skewness and kurtosis, respectively, do matter to risk averse investors, who show a preference for positive skewness and an aversion to kurtosis (see, Kraus and Litzenberger, 1976; Fang and Lai, 1997; Dittmar, 2002; Post et al., 2008). With this in mind, we perform the Shapiro–Wilk test on the return series of all the indices. We find that none of the returns are normally distributed. This provides further evidence that MV analysis is ill-suited for this data set. It has also been argued that stock return data is more likely to be log-normally distributed than normally distributed because stock prices cannot be negative. Hence, we test to see if the data is log-normally distributed using the Shapiro–Wilk test. We reject normality in all cases.[10]

It is important to note that none of the previous studies have discussed the issue of normality nor presented any tests to show that the returns were normally distributed. Given the importance of normality for the MV paradigm they employ, this is an oversight that casts a shadow of doubt on their results.

As a first comparison, we test the homogeneity of means and variances of the two indices in each set. The t-test and Levene's F test are commonly used to compare means and variances respectively. Both of these tests assume normally distributed data but are robust to minor deviations from normality. Therefore, as a first step we use these tests. However, since our data is significantly non-normal we also compare means and variances using tests that are specially designed to be robust for non-normally distributed data. Thus, in the next step, for comparing means we use the

[7] Schroder (2007) argued against the need to use multi-factor models since indices do not follow specific investment styles and are closely related to the market index.

[8] It is worthwhile noting that some prior studies have compared directly SRI indices with their relevant benchmarks (e.g., Schroder, 2007). We believe that this approach is somehow flawed because it compares two indices, one of which is far more diversified than the other, thereby violating the canon of likewise comparisons. For example, comparing the performance of the FTSE4Good-UK-50 which is an SRI index comprised of 50 stocks with that of the FTSE-Allshare which is composed of almost all listed stocks in the UK. It would make for a fairer comparison if one were to use the FTSE-Allshare to represent the market index/portfolio while comparing the performance of the SRI FTSE4Good-UK-50 with the conventional FTSE-100 or the FTSE-250.

[9] Following the original empirical implementation of MCSD by Shalit and Yitzhaki (1994), we use the index as a proxy for daily changes in individual wealth. Since there is no need to specify utility functions, any monotone transformation of individual wealth is appropriate.

[10] We list here only the results for the arithmetic returns series since that is the one used in our study.

Table 2
Descriptive statistics of the index return series.

Type	Index name	Min	Max	Mean	SD	Skew.	E. Kurt.	SWT
S	FTSE4GOOD-UK-50	−0.12978	0.16798	−0.00015	0.02709	−0.23246	6.24343	0.909**
C	FTSE-100	−0.12532	0.16689	0.00000	0.02665	−0.25559	5.96409	0.916**
M	FTSE-ALL SHARE	−0.11853	0.16581	0.00015	0.02602	−0.25487	5.76553	0.921**
S	FTSE4GOOD-UK-50	−0.12978	0.16798	−0.00015	0.02709	−0.23246	6.24343	0.909**
C	FTSE-250	−0.12268	0.17345	0.00114	0.02794	−0.14800	4.29900	0.947**
M	FTSE-ALL SHARE	−0.11853	0.16581	0.00015	0.02602	−0.25487	5.76553	0.921**
S	FTSE4GOOD-US-100	−0.15863	0.11603	−0.00039	0.02709	−0.62937	5.35700	0.931**
C	S&P-100	−0.13991	0.13236	−0.00033	0.02599	−0.44257	5.01291	0.936**
M	DJ-TSMI-US	−0.16620	0.11907	0.00022	0.02663	−0.72910	5.45943	0.937*
S	FTSE4GOOD-US-100	−0.15863	0.11603	−0.00039	0.02709	−0.62937	5.35700	0.931**
C	DJIA	−0.13852	0.11950	0.00009	0.02479	−0.47110	4.54216	0.941**
M	DJ-TSMI-US	−0.16620	0.11907	0.00022	0.02663	−0.72910	5.45943	0.937**
S	FTSE4GOOD-EU-50	−0.15164	0.13536	−0.00110	0.03027	−0.46256	4.68550	0.926**
C	STOXX-50	−0.14877	0.14565	−0.00089	0.03340	−0.44893	3.45514	0.942**
M	STOXX-TM	−0.14273	0.16196	−0.00055	0.03152	−0.38910	3.60632	0.942**
S	FTSE4GOOD-GLOBAL-100	−0.11813	0.11368	−0.00044	0.02775	−0.33170	2.97959	0.954**
C	S&P-GLOBAL-100	−0.10980	0.11253	−0.00014	0.02665	−0.27752	2.77667	0.959**
M	FTSE-ALL WORLD ($)	−0.13127	0.13044	0.00045	0.02674	−0.46875	3.37907	0.951**

The indices are grouped as follows: 1 SRI index, 1 conventional index and 1 market index within each set. All the skewness and excess kurtosis values are significant at the 5% level. SD is the standard deviation. Skew is the Skewness. E. Kurt is Excess Kurtosis. SWT is The Shapiro–Wilk test for normality. S = SRI, C = Conventional, M = Market and SD = Standard Deviation.
** Significant at the 1% level.

Table 3
Results of the homogeneity of means and variances tests.

Type	Index name	*t*-test	Mann–Whitney *U*	Levene's *F* test	Brown–Forsythe
S	FTSE4GOOD-UK-50	−0.078	117,789	0.001	0.003
C	FTSE-100	(0.937)	(0.860)	(0.981)	(0.955)
S	FTSE4GOOD-UK-50	−0.747	113,143	2.298	2.189
C	FTSE-250	(0.455)	(0.220)	(0.130)	(0.139)
S	FTSE4GOOD-US-100	−0.021	118,314	0.357	0.295
C	S&P-100	(0.983)	(0.950)	(0.550)	(0.587)
S	FTSE4GOOD-US-100	−0.259	117,326	1.343	1.303
C	DJIA	(0.796)	(0.770)	(0.247)	(0.254)
S	FTSE4GOOD-EU-50	0.154	116,537	3.270	2.754
C	STOXX-50	(0.878)	(0.640)	(0.071)	(0.097)
S	FTSE4GOOD-GLOBAL-100	−0.159	117,917	0.169	0.170
C	S&P-GLOBAL-100	(0.874)	(0.880)	(0.681)	(0.680)

To compare means we use an independent sample *t*-test and the Mann–Whitney *U* test. To compare variances we use the Levene's *F* test and the Brown & Forsythe test. The Mann–Whitney *U* and Brown & Forsythe tests are robust for non-normally distributed data. For more details on these tests please see pages 9 and 10. S = SRI, C = Conventional. *P*-Values are in parentheses.

Mann–Whitney *U* test and for comparing variances we use the Brown and Forsythe (1974) test.[11]

Table 3 lists results of the comparison of means and variances tests. We find no statistically significant difference at the 5% level between the means and variances of the SRI vs conventional indices return series. Thus, based on MV analysis our results are insignificant and we proceed to the next stage of comparison using Marginal Conditional Stochastic Dominance (MCSD).

[11] The Mann–Whitney *U* test is defined as follows: $U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - \sum_{i=n_1+1}^{n_2} R_i$. Where: n_1 = sample size of the first sample; n_2 = sample size of the second sample; R_i = pooled ranks. *U* can be thought of as the number of times observations in one sample precede observations in the other sample within the pooled ranks. The Brown and Forsythe (1974) test is run in two steps. In the first step a new time series $Z_i = |y_i - m_i|$ is calculated, where m_i is the median for group *i*. The new time series is the distance of each observation in the original times series from the median of the original series. Step 1 is then repeated for the other index as well. In Step 2 ANOVA is used to test if the means of these two new series are equal. As per ANOVA, if the two means are equal then we can say that variances of the original series are equal as well.

4. Methodology

4.1. MCSD and Absolute Concentration Curves

Under the general assumption that investors are risk averse, MCSD provides the probabilistic conditions under which all risk-averse investors prefer one risky asset over another. In the terminology of stochastic dominance, MCSD provides the tools to assess the "dominance" or superiority of one asset over another. Dominance means that the utility of all risk averse investors can be improved by increasing the share of the dominant asset at the expense of the dominated asset. Thus, in our case with respect to the matched pair of SRI and conventional indices, we use MCSD for performance evaluation. If dominance exists, the dominating index outperforms the dominated index. In the absence of dominance, performance is deemed equivalent.

According to the MCSD theorem, given a portfolio α, asset k dominates asset j for all concave utility functions if and only if:

$$\text{ACC (k)} \geqslant \text{ACC (j) with at least one strong inequality} \quad (2)$$

Table 4
Results of the performance tests.

Type	Index name	Sharpe Ratio	Treynor Ratio	Jensen's Alpha	4 Factor Alpha	MCSD Test
S	FTSE4GOOD-UK-50	−0.032	−0.0008	−0.0003	−0.0026	No dominance
C	FTSE-100	−0.027	−0.0007	−0.0001	−0.0005*	
M	FTSE-ALL SHARE	−0.022	−0.0006	NA	NA	
S	FTSE4GOOD-UK-50	−0.032	−0.0008	−0.0003	−0.0026	Conventional dominates SRI
C	FTSE-250	0.015	0.0005	0.0010	0.0043*	
M	FTSE-ALL SHARE	−0.022	−0.0006	NA	NA	
S	FTSE4GOOD-US-100	−0.030	−0.00083	−0.00061*	−0.0068*	Conventional dominates SRI
C	S&P-100	−0.029	−0.00079	−0.00056*	−0.0013	
M	DJ-TSMI-US	−0.008	−0.0002	NA	NA	
S	FTSE4GOOD-US-100	−0.030	−0.0008	−0.0006*	−0.0068*	Conventional dominates SRI
C	DJIA	−0.014	−0.0004	−0.0001	−0.0002	
M	DJ-TSMI-US	−0.008	−0.0002	NA	NA	
S	FTSE4GOOD-EU-50	−0.052	−0.0017	−0.0006	−0.0022	No dominance
C	STOXX-50	−0.041	−0.0013	−0.0003	−0.0017*	
M	STOXX-TM	−0.033	−0.0009	NA	NA	
S	FTSE4GOOD-GLOBAL-100	−0.031	−0.0009	−0.0009*	−0.00172*	Conventional dominates SRI
C	S&P-GLOBAL-100	−0.021	−0.0006	−0.0006	−0.00173*	
M	FTSE-ALL WORLD ($)	0.001	0.0000	NA	NA	

S = SRI, C = Conventional and M = Market. The market index is used as the market portfolio in the MV approach and for wealth ranking in the MCSD approach.
* Indicates significance at the 5% level.

where:

ACC = Absolute Concentration Curves

More simply, asset k dominates asset j if the ACC of asset k lies above the ACC of asset j. We follow Shalit and Yitzhaki (1994) to calculate the said ACCs as follows.

In each set we have 3 indices: one SRI, one conventional and one market. We take the weekly returns for the 3 indices where the number of observations (N) = 487 in each series. We use the market index as the wealth index and sort (or rank) the returns on this index from lowest to highest.[12] The returns of each index are then matched to the return on the wealth index. For example, if the lowest return on the wealth index was for the 10th week of observations, we match the returns of each index for the 10th week of observations. Next, each of the terms in the two index return series (SRI and conventional) is multiplied by $1/N$ to obtain equally weighted returns. We now take the cumulative sum of this weighted return series for each index, i.e., each term in the cumulative sum series is the sum of all previous terms of the weighted return series. For example, the 3rd term of the cumulative return series of index A is the sum of the 1st, 2nd and 3rd terms from the weighted return series for index A. This cumulative return series for index A is known as the ACC for index A. Similarly we calculate the ACC for the other index. Next we compare the two ACCs calculated above at each of the 487 points. According to the MCSD criteria, one index dominates the other if its ACC is either equal to or greater than the ACC of the other at all the points. The results of the MCSD tests for the six sets of indices are reported in Table 4.

4.2. Mean–variance testing

To allow for comparison with previous studies and to check the robustness of our results, we also perform the MV analysis implemented by earlier studies. We calculate and compare the Sharpe Ratios, Treynor Ratios, Jensen's Alphas and four factor alphas for each pair of indices.

The Sharpe Ratio (Sharpe, 1966) is defined as the excess return of a portfolio per unit of risk, which is measured as the standard deviation of the return.

$$\text{Sharpe Ratio} = \frac{r_i - r_f}{\sigma_i} \tag{3}$$

where: r_i = mean return of index i; r_f = risk free rate for the given period in the respective currency; σ_i = standard deviation of the index i returns.

The Treynor Ratio (Treynor and Mazuy, 1966) is similar to the Sharpe Ratio. It calculates the excess return of a portfolio per unit of risk which is measured as the Beta of the portfolio.

$$\text{Treynor Ratio} = \frac{r_i - r_f}{\beta_i} \tag{4}$$

where: r_i = mean return of index; r_f = risk free rate for the given period in the respective currency; β_i = beta of the index relative to the market portfolio.

Jensen's alpha (Jensen, 1968) is used to calculate the excess return of a portfolio. Simply speaking, this is the constant in the CAPM regression:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + \epsilon_{i,t} \tag{5}$$

where: $r_{i,t}$ = return of index at time t; $r_{f,t}$ = risk free rate at time t; α_i = excess return or Jensen's alpha for index i; β_i = beta for index i; $r_{m,t}$ = return of the market at time t; $\varepsilon_{i,t}$ = random error term at time t.

A similar approach, proposed by Carhart (1997), is used for the four factor model. In addition to the market index, it includes three other risk factors: value, size and momentum. Thus the alpha of this model "risk adjusts" the excess returns for 4 factors in the following model.[13]

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_{i,mkt}(r_{m,t} - r_{f,t}) + \beta_{i,val}VAL_t + \beta_{i,size}SIZE_t + \beta_{i,mom}MOM_t +_{i,t} \tag{6}$$

where: $r_{i,t}$ = return of index at time t; $r_{f,t}$ = risk free rate at time t; $r_{m,t}$ = return of the market at time t; α_i = excess return or 4 factor alpha for index i; $\beta_{i,mkt}$ = market beta for index i; $\beta_{i,val}$ = value factor

[12] As mentioned above, since there is no need to specify utility functions, any monotone transformation of individual wealth is an appropriate wealth proxy.

[13] In our case, the four factors for US, Europe and Global portfolios were obtained from the Kenneth French website (2013). The four factors for UK were obtained from Gregory et al. (2013).

Table 5
Performance improvements in skewness and kurtosis.

Type	Index Name	Skew.	Kurt.	Δ Skew.* (%)	Δ Kurt. (%)
S	FTSE4GOOD-UK-50	−0.2325	6.2434	36.33	31.14
C	FTSE-250	−0.1480	4.2990		
S	FTSE4GOOD-US-100	−0.6294	5.3570	29.68	6.42
C	S&P-100	−0.4426	5.0129		
S	FTSE4GOOD-US-100	−0.6294	5.3570	25.15	15.21
C	DJIA	−0.4711	4.5422		
S	FTSE4GOOD-GLOBAL-100	−0.3317	2.9796	16.33	6.81
C	S&P-GLOBAL-100	−0.2775	2.7767		
**Average Δ =>				26.87	14.90

Table 5 shows performance improvements in skewness (Skew.) and kurtosis (Kurt.) for MCSD dominant indices, when investors move their investment from the SRI index to the conventional one.

$\Delta\text{Skew} = \left(\frac{\text{Skew(E)}-\text{Skew(C)}}{\text{Skew(E)}}\right) \times 100\%;$; $\Delta\text{Kurtosis} = \left(\frac{\text{Kurt(E)}-\text{Kurt(C)}}{\text{Kurt(E)}}\right) \times 100\%.$

* Since we are dealing here with negative skewness in all cases, improvement/increase in skewness implies that the conventional index has lesser negative skewness than the SRI one.

** On average, investors can improve both skewness and kurtosis of their portfolios by choosing not to invest responsibly, they can increase their skewness by 27% and reduce their kurtosis by 15%.

beta for index i; $\beta_{i,size}$ = size factor beta for index i; $\beta_{i,mom}$ = beta for the momentum factor; $\varepsilon_{i,t}$ = random error term at time t.

If these alphas are positive and significant then the asset is said to outperform. Negative and significant alphas indicate poor performance. We test both types of alphas (Jensen's and 4 factor) to see if they are statistically significant using the t-test and the White and Newey–West standard errors which are robust to heteroskedasticity and serial correlation.[14]

5. Empirical results

5.1. Performance comparison and analysis

Table 4 presents the results of both MCSD and MV tests. We find that the Jensen alpha pairs are significant in only 1 out of 6 cases. Similarly, 4 factor alphas provide conclusive results in only 1 case. The Sharpe and Treynor Ratios are negative, which renders them meaningless. This is because the Sharpe and Treynor Ratios calculate the excess return over the risk free rate per unit of risk. Other things being equal, when excess returns are positive, a higher level of risk will render a smaller value of the Sharpe/Treynor Ratio. Thus, if two investments have identical excess returns, the investment with the lower risk will have a higher Sharpe/Treynor ratio. However, if two investments have identical negative excess returns, a higher level of risk produces a smaller negative number and thus the investment with the higher risk comes out on top. This is antithetical to the concept of performance for a risk averse investor. Thus, negative excess return ratios are misleading. These results are consistent with Statman (2000, 2006) and Schroder (2007).

The MCSD approach paints a different picture. The FTSE4Good-US-100 SRI index is dominated by the similar conventional S&P 100 index as well as by the Dow Jones Industrial Average. The FTSE4Good-Global-100 SRI index is dominated by the conventional S&P-Global-100 index and the FTSE4Good-UK-50 is dominated by the FTSE-250 but not the FTSE-100. Over all we find that conventional indices outperformed the SRI indices in the UK, the US and the Global context. In the European context the conventional and SRI indices performed equally. These results suggest a clear pattern of inferior performance of SRI indices with respect to the conventional indices.

As outlined above, dominance signifies outperformance. The insignificant MV results suggest that the mean and the variance alone cannot explain the outperformance. Thus, we look at skewness and kurtosis, moments three and four, to examine whether SRI investors pay a price in the higher moments by way of lesser skewness and higher kurtosis. We find that this is indeed the case. Table 5 shows the cost of SRI in terms of skewness and kurtosis for the 4 cases where MCSD dominance has been established.[15] We find that SRI investors can improve both the skewness and kurtosis of their portfolios by choosing to invest in the conventional index as opposed to the SRI index. Investors can increase, on average, their skewness by 27% and reduce their kurtosis by 15% by choosing not to invest responsibly.

As a further robustness test we construct arbitrage portfolios to see whether investor utility can be improved by exploiting the dominance. To this end we create a zero cost portfolio by selling the SRI index short and using the proceeds to invest in the conventional index.[16] We do this for all the cases where the conventional index MCSD dominates the SRI index. We then compare the returns on this portfolio with those of the conventional and SRI indices. The results in Table 6 show that in all four cases the mean of the arbitrage portfolio is higher than the means of both the conventional and SRI indices, the standard deviation is lower and the skewness is higher. In fact, the skewness is positive for all four arbitrage returns while it is negative for the conventional and SRI indices. Finally, kurtosis is lower for all but three cases. This is strong evidence that selling the SRI and purchasing the conventional improves the return distribution. A test for MCSD confirms that this is indeed the case and that the improvement increases investor utility. The arbitrage portfolios MCSD dominate both the SRI and conventional indices in all four cases.

5.2. Further analysis with carefully matched samples of firms

To confirm the foregoing results we control for potential bias due to size. If SRI indices are mainly made up of large sized stocks while conventional indices contain small sized stocks, the difference in performance between the two may arise due to the size factor.[17] There is a hint of this in Table 4 where the FTSE4Good UK-50 index is not dominated by the FTSE-100 index but is

[14] In order to ensure the robustness of our regressions, we also test all the index return series for stationarity using the Augmented Dickey–Fuller (ADF) test. The return series are stationary in all cases. For brevity we do not report the ADF tests, but are available on request.

[15] In the other 2 cases since there is no MCSD dominance we cannot reject the hypothesis of equal performance.

[16] This procedure follows Clark and Kassimatis (2012).

[17] We thank an anonymous referee for this insight.

Table 6
Arbitrage portfolio results.

Type	Index name	Mean	SD	Skew.	E. Kurt.	MCSD test
S	FTSE4GOOD-UK-50	−0.00015	0.02709	−0.23246	6.24343	A dominates S
C	FTSE-250	0.00114	0.02794	−0.14800	4.29900	A dominates C
A	Long C and Short S	0.00132	0.01767	0.80200	4.78900	
S	FTSE4GOOD-US-100	−0.00039	0.02709	−0.62937	5.35700	A dominates S
C	S&P-100	−0.00033	0.02599	−0.44257	5.01291	A dominates C
A	Long C and Short S	0.00003	0.00469	0.17100	1.60200	
S	FTSE4GOOD-US-100	−0.00039	0.02709	−0.62937	5.35700	A dominates S
C	DJIA	0.00009	0.02479	−0.47110	4.54216	A dominates C
A	Long C and Short S	0.00050	0.00758	0.02500	3.31800	
S	FTSE4GOOD-GLOBAL-100	−0.00044	0.02775	−0.33170	2.97959	A dominates S
C	S&P-GLOBAL-100	−0.00014	0.02665	−0.27752	2.77667	A dominates C
A	Long C and Short S	0.00028	0.00476	0.13200	9.38300	

S = SRI, C = Conventional, A = Arbitrage Portfolio, SD = Standard Deviation, Skew is the Skewness and E. Kurt is Excess Kurtosis. Arbitrage portfolios (A) are zero cost portfolios constructed by selling the dominated SRI index (S) short and using the proceeds to buy the dominant conventional index (C).

Table 7
Mean comparison of firm size.

Type	Index name	Mean (millions)	Min (millions)	Max (millions)	*T* statistic	Currency
S	FTSE4GOOD-UK-50	20,531	1677	131,853	−1.34	£
C	FTSE-100	13,680	68	131,173	(0.18)	£
S	FTSE4GOOD-UK-50	20,531	1677	131,853	−4.47*	£
C	FTSE-250	1458	97	4441	(0.00)	£
S	FTSE4GOOD-US-100	41,809	9010	354,472	2.42*	$
C	S&P-100	64,620	814	432,166	(0.02)	$
S	FTSE4GOOD-US-100	41,809	9010	354,472	3.98*	$
C	DJIA	124,427	8089	432,166	(0.00)	$
S	FTSE4GOOD-EU-50	59,067	15,208	214,894	−2.72*	€
C	STOXX-50	38,705	2903	120,400	(0.01)	€
S	FTSE4GOOD-GLOBAL-100	62,887	16,786	354,472	2.26*	$
C	S&P-GLOBAL-100	844,586	7054	28,676,640	(0.03)	$

Firm size is measured as the market capitalisation of the stocks which comprise the indices used in this study. *T* test is the test for difference in the means. S = SRI and C = Conventional. *P*-Values in parentheses.
* Indicates significance at the 5% level.

Table 8
Descriptive statistics of carefully matched index return series.

Index match	Min	Max	Mean	SD	Skew.	E. Kurt.	SWT
UK matched	−0.12498	0.14257	0.00026	0.02604	−0.070	4.094	0.947**
US matched	−0.09417	0.09259	0.00018	0.01801	−0.168	3.886	0.957**
EU matched	−0.12761	0.17041	−0.00016	0.02867	−0.041	5.577	0.916**
Global matched	−0.09098	0.11367	0.00026	0.02458	−0.019	3.128	0.951**

Table 8 shows descriptive statistics of the return series for the carefully matched indices. SD is the standard deviation. Skew is the Skewness. E. Kurt is Excess Kurtosis. SWT is The Shapiro–Wilk test for normality.
** Significant at the 1% level.

dominated by the FTSE-250 index, where the latter contains smaller sized stocks.[18] Table 7 shows descriptive statistics on size i.e. market capitalisations of stocks that comprise the indices in this study. We find that in 5 out of 6 cases there is a statistically significant difference in size. In 2 out of those 5 cases the conventional index has smaller sized stocks while in the other 3 cases the SRI index has smaller sized stocks. In the four cases where MCSD was observed, the conventional index has larger firms in three cases and smaller firms in one case. Thus, the size effect is not supported by this preliminary testing.

To pursue the analysis and control for a potential industry effect along with a size effect, we construct carefully matched samples of conventional firms based on industry and size for each SRI index. In so doing, all firms which are not included in SRI indices are classified as conventional firms. Since we have four SRI indices we construct four carefully matched samples. The matching procedure used is as follows. We consider a conventional firm as matching an SRI firm if the former has the same 4-digit industrial classification number and its market capitalisation is between 70% and 130% of the SRI firm. Table 8 reports the descriptive statistics for the matched indices.

Finally, we run the MCSD test on these pairs of matched indices in order to compare performance. The results in Table 9 show that in all four cases the FTSE4Good indices are dominated by a

[18] We ran the MCSD test between the FTSE100 and FTSE250 and find that the latter dominates the former.

Table 9
MCSD test results for FTSE4Good and matched indices.

Index name	# Of firms	Mean	*T*-test	MCSD test
FTSE4GOOD-UK-50 (S)	50	20,531	0.80	CM dominates S
UK Matched (CM)	55	16,117	(0.43)	
FTSE4GOOD-US-100 (S)	100	41,809	1.32	CM dominates S
US Matched (CM)	124	34,130	(0.19)	
FTSE4GOOD-EU-50 (S)	50	59,067	−0.48	CM dominates S
EU Matched (CM)	48	63,417	(0.63)	
FTSE4GOOD-GLOBAL-100 (S)	100	62,887	−1.39	CM dominates S
Global Matched (CM)	447	70,738	(0.17)	

S = SRI and CM = Carefully Matched. *P*-Values in parentheses.

conventional index composed of firms matched to the firms in FTSE4Good indices by size and industry. Thus, neither size nor industry seems to account for the underperformance of the SRI indices and is further evidence that there is a price to be paid by risk averse investors for socially responsible investing.

6. Conclusions

Previous research on the comparative performance of SRI vs conventional investments has used the MV framework, which is a special case of expected utility maximization based on normal returns or quadratic utility functions, and generally found no conclusive evidence of a significant advantage or disadvantage to socially responsible investing. Given that all the return series in our sample are significantly non-normally distributed and that investors do not necessarily have quadratic utility functions, we abandon mean–variance in favour of Marginal Conditional Stochastic Dominance, which can accommodate any return distribution or concave utility function.

Our results show that although there is nothing to be gained or lost from socially responsible investing in terms of mean and variance, there is a price to be paid in the higher moments of the return distributions. For example, on average conventional indices have 27% higher skewness and 15% lower kurtosis than their SRI counterparts. More specifically, we show that indices composed of socially responsible firms are dominated by indices composed of conventional firms in trademarked indices as well as in indices carefully matched with the firms in the SRI indices. This means that risk averse investors can improve their expected utility by reducing their holdings of SR firms and purchasing conventional ones. We test this proposition by constructing zero cost portfolios created by shorting the SRI index and using the proceeds to invest in the conventional index. These zero cost portfolios generate higher average returns, lower variance and higher skewness than either of the two indices standing alone. Most importantly, they also MCSD dominate the SRI and conventional indices standing alone, which confirms the proposition that risk averse investors can improve their expected utility by reducing their holdings of SR firms and purchasing conventional stocks.

Our results provide strong evidence that there is a financial price to be paid for socially responsible investing. The loss in financial utility is presumably compensated by the non-financial utility that SRI investors derive from the responsible nature of their investments. Integrating the role of non-financial utility into the investment paradigm looks like a fruitful prospect for future research.

References

Amenc, N., Sourd, V., 2008. Socially Responsible Investment Performance in France. EDHEC Business School.
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.
Brown, M., Forsythe, A., 1974. Robust tests for the equality of variances. Journal of the American Statistical Association 69, 364–367.
Carhart, M., 1997. On Persistence in Mutual Fund Performance. Journal of Finance 52, 57–82.
Chan, P.T., Walter, T., 2014. Investment performance of "environmentally friendly" firms and their initial public offers and seasoned equity offers. Journal of Banking and Finance 44, 177–188.
Clark, E., Jokung, O., 1999. Asset proportions, stochastic dominance and the 50% rule. Management Science 45, 1724–1727.
Clark, E., Kassimatis, K., 2012. An empirical analysis of marginal conditional stochastic dominance. Journal of Banking and Finance 36, 1144–1151.
Clark, E., Kassimatis, K., 2013. International equity flows, marginal conditional stochastic dominance, and diversification. Review of Quantitative Finance and Accounting 40, 251–271.
Derwall, J., Koedijk, K., Ter Horst, J., 2011. A tale of values-driven and profit-seeking social investors. Journal of Banking and Finance 35, 2137–2147.
Dittmar, R., 2002. Nonlinear asset kernels kurtosis preference and evidence from cross section of equity returns. Journal of Finance 57, 369–403.
Eeckhoudt, L., Schlesinger, H., 2006. Putting risk in its proper place. American Economic Review 96, 280–289.
Fang, H., Lai, T., 1997. Co-kurtosis and capital asset pricing. Financial Review 32, 293–307.
Fernandez, A., Matallin, J., 2008. Performance of ethical mutual funds in Spain: sacrifice or premium? Journal of Business Ethics 81, 247–260.
French, 2013. Kenneth French data library. <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html> (15.08.13).
Geczy, C., Stambaugh, F., Levin, D., 2005. Investing in socially responsible mutual funds. Available at SSRN: <http://ssrn.com/abstract=416380>.
Goldreyer, F., Ahmed, P., Diltz, J., 1999. The performance of socially responsible mutual funds: incorporating socio-political information in portfolio selection. Managerial Finance 25, 23–36.
Gregory, A., Matatko, J., Luther, R., 1997. Ethical unit trust financial performance: small company effects and fund size effects. Journal of Business Finance and Accounting 24, 705–725.
Gregory, A., Tharayan, R., Christidis, A., 2013. Constructing and testing alternative versions of the Fama–French and Carhart models in the UK. Journal of Business Finance and Accounting 40, 172–214.
Hamilton, S., Jo, H., Statman, M., 1993. Doing well while doing good? The investment performance of socially responsible mutual funds. Financial Analysts Journal 49, 62–66.
Hong, H., Kacperczyk, M., 2009. The price of sin: the effects of social norms on markets. Journal of Financial Economics 93, 5–36.
Humphrey, J.E., Tan, D.T., 2013. Does it really hurt to be responsible? Journal of Business Ethics 1055. http://dx.doi.org/10.1007/s1-013-1741-z.
Investment Management Association, 2014. Key statistics. <http://www.investmentuk.org/research/ima-annual-industry-survey/key-statistics/> (19.04.14).
Jensen, M., 1968. The performance of mutual funds in the period 1945–1964. Journal of Finance 23, 389–416.
Kempf, A., Osthoff, P., 2007. The effect of socially responsible investing on financial performance. European Financial Management 13, 908–922.
Kraus, A., Litzenberger, R., 1976. Skewness preference and the valuation of risky assets. Journal of Finance 31, 1085–1099.
Kreander, N., Gray, G., Power, D., Sinclair, C., 2005. Evaluating the performance of ethical and non-SRI funds: a matched pair analysis. Journal of Business Finance and Accounting 32, 1465–1493.
Malik, M., 2014. Value enhancing capabilities of CSR: a brief review of contemporary literature. Journal of Business Ethics 1055. http://dx.doi.org/10.1007/s1-014-2051-9.
Mallin, C., Saadouni, B., Briston, R., 1995. The financial performance of ethical investment funds. Journal of Business Finance and Accounting 22, 483–496.
Mandelbrot, B., 1963. The variation of certain speculative prices. Journal of Business 36, 394–419.
Munoz, F., Vargas, M., Marco, I., 2013. Environmental mutual funds: financial performance and managerial abilities. Journal of Business Ethics 1055. http://dx.doi.org/10.1007/s1-013-1893-x.
Post, T., Vlient, P., Levy, H., 2008. Risk aversion and skewness preference: a comment. Journal of Banking and Finance 32, 1178–1187.
Renneboog, L., Horst, J., Zhang, C., 2008a. Socially responsible investments: institutional aspects, performance, and investor behavior. Journal of Banking and Finance 32, 1723–1742.
Renneboog, L., Horst, J., Zhang, C., 2008b. The price of ethics and stakeholder governance: the performance of socially responsible mutual funds. Journal of Corporate Finance 14, 302–322.
Sauer, D., 1997. The impact of social-responsibility screens on investment performance: evidence from the Domini 400 social index and Domini equity fund. Review of Financial Economics 6, 23–35.
Schroder, M., 2007. Is there a difference? The performance characteristics of SRI equity indices. Journal of Business Finance and Accounting 34, 331–348.
Shalit, H., Yitzhaki, S., 1994. Marginal conditional stochastic dominance. Management Science 40, 670–684.
Shalit, H., Yitzhaki, S., 2010. How does beta explain stochastic dominance efficiency? Review of Quantitative Finance and Accounting 35, 431–444.

Sharpe, W., 1966. Mutual fund performance. Journal of Business 39, 119–138.
Statman, M., 2000. Socially responsible mutual funds. Financial Analysts Journal 56, 30–39.
Statman, M., 2006. Socially responsible indexes: composition, performance, and tracking error. Journal of Portfolio Management 32, 100–109.
Statman, M., Glushkov, D., 2009. The wages of social responsibility. Financial Analysts Journal 65, 33–46.
Treynor, J., Mazuy, K., 1966. Can mutual funds outguess the market. Harvard Business Review 44, 131–136.
The UK Sustainable Investment and Finance Association, 2014. About UKSIF. <http://uksif.org/about-uksif/history/> (19.04.14).
The Forum for Sustainable and Responsible Investment is the US, 2014. SRI assets in the United States. <http://www.ussif.org/sribasics> (19.04.14).