

Skill and style in socially responsible portfolios: A holdings-based analysis

This version: May 2013

Please do not quote or cite without permission of authors

Junyan Boon

Jacquelyn E. Humphrey*

Geoffrey J. Warren

Research School of Finance, Actuarial Studies and Applied Statistics, Australian National University

Acknowledgements: the authors thank Cessare Fracassi, Clemens Sialm, Laura Starks, Sheridan Titman and workshop participants at the University of Texas at Austin for helpful comments.

*Corresponding Author: Dr Jacquelyn Humphrey
Research School of Finance, Actuarial Studies and Applied Statistics
Level 4, College of Business and Economics Building 26C
Australian National University, ACT 0200 Australia
Telephone +61 2 6125 7292
Fax+61 2 6120 087
Email jacquelyn.humphrey@anu.edu.au

Abstract

We investigate the components of active returns in constrained and unconstrained portfolios. Using a sample of socially responsible (SRI) funds, we find that constraining portfolios neither hinders nor improves managers' ability to generate returns. SRI funds outperformed a matched sample of conventional funds over the period 2000 to 2010. The managers exhibited no difference in security selection abilities, but SRI managers generated superior style timing performance. SRI funds' returns were hampered by being over-exposed to large capitalization stocks. We also provide an analysis of methods used to match SRI and non-SRI funds, and find inappropriate matching can affect results.

1. Introduction

There is a large literature dating back almost 50 years which suggests that, after fees and trading costs, U.S. fund managers are not able to outperform passive benchmarks (see for example Sharpe, 1966; Jensen, 1968; Gruber, 1996; Carhart, 1997; Barras et al., 2010). More recently, however, some have argued that traditional fund performance models do not do justice to fund managers, as returns-based analysis includes the effect of expenses and trading costs, while benchmarks are free of these costs (Cremers et al, 2012). Further, if these costs are not homogeneous across all funds, the true picture of a particular fund manager's ability could be distorted. There also may be misspecification of the appropriate benchmark, which again will lead to erroneous conclusions about a manager's skill. To address some of these weaknesses, holdings-based models have been developed, which examine the performance of the actual stocks held by the fund, as opposed focusing on overall fund return, net of fees. Results from these studies demonstrate that managers may have better ability than previously suggested by returns-based studies (see Grinblatt and Titman, 1993; Daniel, Grinblatt, Titman and Wermers, 1997; Wermers 2000; Chen et al., 2000).

In this paper we investigate the generation of active returns by fund managers whose mandates require them to constrain their investible universe, but still provide a competitive return. Specifically, we investigate whether managers of socially responsible funds (hereafter SRI funds) exhibit differing abilities from managers of conventional (i.e. not SRI) funds. We also examine the extent to which SRI funds' returns are affected by holding stocks of differing style than conventional funds, holding managerial skill constant. Our study employs the characteristic-based approach of Daniel, Grinblatt, Titman and Wermers, (1997 – hereafter 'DGTW'), and is the first to use this method to examine the performance of SRI funds. The DGTW approach is ideal for our study, as it is able to disentangle managerial skill and fund style from other fund characteristics that may impact return.

SRI funds form their investment portfolios by screening for particular social, ethical, environmental or other preferences. This is usually operationalized by applying negative and/or positive screens to the investible universe. Negative screening is where undesirable firms are excluded from the portfolio. Popular negative screens are tobacco, alcohol and gambling, but firms could also be excluded because of issues such as having poor human rights records or environmental degradation. In contrast, positive (or best-of-sector) screening includes firms with desirable characteristics into the portfolio. Examples of positive screens are including firms with good human rights records, diversity policies, or firms involved in clean technology. Most SRI funds apply multiple screens, both positive and negative.¹

The SRI market is no longer a tiny niche market: one out of every nine dollars under professional management in the U.S. today follows some form of SRI strategy (U.S. SIF 2012). Since 1995, US SRI assets under management have increased by 486 percent. In comparison, the universe of professionally managed assets has increased 376 percent (U.S. SIF 2012).

While the extant literature on SRI funds' performance has primarily focussed on funds' after-fees performance (see, for example, Hamilton et al. 1993; Sauer 1997; Statman 2000; Bauer et al. 2005; Kreander et al. 2005; Benson et al, 2006), we are interested in whether constraining the portfolio has an impact on the underlying sources of active returns, including managerial ability and style biases. The DGTW approach is well suited to this purpose, as it disaggregates performance into three components: manager stock selection ability; manager characteristic timing ability and fund style. We also use an improved method of matching SRI and conventional funds to the techniques typically used in prior studies. Our analysis incorporates an investigation of whether the matching process impacts upon the results of studies comparing two groups of funds.

¹ See – <http://charts.ussif.org/mfpc/> - date accessed 01/17/2013.

We would expect to find differences in the sources of active returns across the two types of funds for several reasons. First, the screening mechanism implemented by SRI funds may hinder managers. Recent research has found that stocks typically excluded by SRI funds - such as stocks in alcohol, tobacco, gambling, firearms, military and nuclear industries - outperform (Fabozzi et al., 2008; Statman and Glushkov, 2008; Hong and Kacperczyk 2009). Excluding these stocks from the portfolio could harm portfolio performance, and precludes fund managers from identifying any profitable opportunities in these segments of the market.

In addition, SRI portfolios may potentially suffer from a loss of diversification, resulting in inefficient portfolios with lower risk-adjusted returns. Specifically, since most SRI funds utilize negative screens,² by default they are subject to a restricted investment universe and might be expected to underperform relative to conventional funds.

On the other hand, it is possible that the investment screening process could convey value-relevant information that manifests in SRI managers outperforming conventional funds. It has been argued that financial markets may undervalue the impact of positive corporate responsibility information on firms' future cash flows (Derwall et al., 2011; Borgers et al., 2013). For example, there is recent evidence to suggest abnormal returns can be earned by investing in firms that have high employee satisfaction (Edmans, 2011; Statman and Glushkov, 2009). This would mean that SRI managers who positively screen their portfolios overweight firms with good behavior along particular responsibility metrics (e.g. employee relations) could therefore earn higher returns. Whether financial markets will systematically undervalue responsibility in the long run, and therefore constantly provide SRI managers with opportunities to earn abnormal returns, seems questionable (Derwall et al., 2011; Borgers et al., 2013).

² Ibid.

Girard et al. (2007) also argue the very fact that SRI fund managers have a smaller investable universe means they have a more intimate knowledge of firms in their portfolios. Consequently, SRI managers could exhibit better skills than conventional managers. Indeed, Kacperczyk et al. (2005) find evidence that more concentrated portfolios perform better than broad portfolios, suggesting that fund managers may benefit from concentrating their holdings in stocks where they have informational advantages.

We might also expect a performance differential between SRI and conventional funds that is purely attributable to the type of stocks the funds have a tendency to hold. Prior literature has documented that SRI and conventional funds have significantly different loadings onto the size, book-to-market and momentum factors, which indicates that they may hold stocks with differing styles to conventional funds (Bauer et al., 2005). Consequently, if small, value, momentum stocks deliver higher return, and SRI funds are over- or underweight across these dimensions relative to conventional funds, this will result in a difference in the returns of the two types of funds.

Our results indicate that SRI funds outperform their conventional counterparts by 0.35% per quarter over our sample period. When we disaggregate performance into its constituent elements, we discover that both SRI and matched conventional fund managers exhibit positive and significant stock selection abilities, consistent with DGTW (1997). However, there is *no significant difference* between the two groups on this metric. SRI fund managers' relative outperformance arises from a 0.52% per quarter difference in characteristic timing. This predominantly reflects conventional managers' poor timing skills (-0.39% per quarter), coupled with some success by SRI managers (0.13% per quarter), which is largely due to timing the book-to-market characteristic well. Finally, we discover a significant difference in the style of SRI and matched conventional funds. SRI funds tend to hold stocks that are larger in size, with higher book-to-market ratios and higher prior-year returns

(momentum). The Average Style (AS) measure shows that these style differences translate into SRI funds underperforming conventional funds by 0.15% per quarter. This is largely due to SRI funds' lower exposure to small capitalization stocks, which outperformed during the sample period.

Our examination of the techniques typically used by prior literature to match SRI and conventional funds indicates that matching can impact results. Over a typical sample period, funds may grow or shrink substantially in size, and may change their style. These changes result in a degradation of the quality of the match between the funds over time, and may translate into incorrect inferences on performance differentials. We conclude that matching needs to be performed multiple times over the sample period to ensure that the characteristics of the two groups of funds do not substantially diverge over time.

Overall, we conclude that constraining a fund manager, at least using responsible criteria, neither hinders nor improves managerial skill. However, SRI funds may generate different performance due to style biases. In particular, they may suffer from their over-exposure to large capitalization stocks in time periods where these stocks underperform. Our results also suggest that SRI managers may possess better style timing ability. However, the difference is largely attributable to poor timing skill by conventional funds, and evidence of positive timing skill for SRI managers is modest. Further, the reason for the positive timing ability is not readily apparent, so we are wary about drawing any strong conclusions based on this finding.

The remainder of the paper proceeds as follows. In section 2 we provide an overview of the benefits of holdings-based analysis and in section 3 we outline the DGTW and matching methodologies. Results are in section 4 and section 5 concludes.

2. Holdings-based analysis

The existing empirical literature on SRI funds has largely focused on the differences between the returns of SRI and conventional funds using returns-based analysis. Most studies find that SRI funds do not produce significantly different risk-adjusted performance than conventional funds (see, for example, Hamilton et al. 1993; Sauer 1997; Statman 2000; Bauer et al. 2005; Bello 2005; Kreander et al. 2005; Benson et al., 2006).

However, the few studies that have examined the relative *skill* of SRI fund managers provide mixed results. Benson et al. (2006), using an industry factor model, find that U.S. SRI fund managers do not exhibit significant stock selection abilities. In contrast, Kreander et al. (2005) find that European SRI fund managers possess significantly positive selection abilities.

Similarly, findings on the market timing abilities of SRI fund managers are mixed. Using the conditional Treynor-Mazuy market timing model, Renneboog et al. (2008) find little evidence of SRI fund managers being able to time the market. Schroder (2004) documents insignificant timing coefficients and Kreander et al. (2005) find no difference in the timing abilities of European SRI and conventional fund managers. We are aware of no studies that have examine the style timing abilities of SRI managers.

All studies to date that have examined SRI fund performance have used returns-based analysis, and many have used four-factor (Carhart, 1997) alphas. However, a recent study by Cremers et al. (2012) finds four-factor models can be misleading for performance measurement, and that the alpha estimates from these models are very noisy. In addition, returns-based analysis has been criticised as not being able to adequately detect manager skill: expenses and trading costs may differ between funds and therefore drive a wedge between a manager's true performance and total portfolio performance, thus distorting the real picture of a particular manager's ability. Kothari and Warner (2001) document that the DGTW (1997)

methodology has increased power over returns-based measures since there are more observations using holdings than fund returns.³

The DGTW method also has benefits in terms of characterising a fund's style. Benchmarking on a security-by-security basis, with each security constituting a separate observation, allows for clear characterization of the fund manager's style in choosing stocks (Wermers, 2000). Since the analysis is performed quarterly, a fund's style is captured each quarter and therefore the DGTW method is superior to factor models where the style coefficients are fixed over the sample period.

These shortcomings of returns-based analysis provide motivation for comparing SRI and conventional funds using the DGTW methodology. The approach offers cleaner measures of stock selection skill, and better handles variation in factor exposures over time. The approach also allows us to disentangle returns due to manager ability from returns due to fund style.

3. Data

SRI funds are identified from lists provided by the U.S. Social Investment Forum's 1999, 2007 and 2010 Trends Reports. Data on fund size, objective and inception dates are from the CRSP Survivor-Bias-Free U.S. Mutual Fund Database. We limit our analysis to funds that hold portfolios of U.S. equities. First, we exclude mutual funds with asset codes other than 'EQ' and, where this is missing, Lipper investment objective. We exclude all international funds, balanced funds, utility funds and fixed income funds. As we are interested in analysing active manager skill, we remove index funds. Our final sample comprises 96 SRI and 1090 conventional funds. Quarterly fund holdings data are from Thomson Reuters Mutual

³ Of course, the DGTW (1997) methodology is not without its shortcomings. Managers may trade for numerous other reasons, such as portfolio rebalancing, liquidity needs, and in response to fund flows. Further, fund managers may not be able to attain their target portfolios: perhaps because insufficient stock was available at an acceptable price, or the stock price moved before the fund manager was able to trade.

Fund Holding Database (henceforth Thomson Reuters). Monthly stock returns are from the CRSP U.S. Stock and Index Database and are compounded to quarterly returns. DGTW (1997) benchmark portfolio returns and stock assignments data are from Russ Wermers' website.^{4 5} The sample period is 2000 to 2010.

4. Methodology

4.1 DGTW method

The DGTW (1997) method compares each stock held by a fund at each point in time to a benchmark portfolio. To form the benchmarks, all common stocks listed on the NYSE, AMEX and Nasdaq are first ranked by their market capitalizations and divided into quintiles using NYSE size breakpoints. Next, within each size quintile, stocks are divided into five quintiles on their book-to-market ratios. Finally, within the 25 size/book-to-market portfolios, stocks are further subdivided into quintiles based on their prior-year return. This three-way ranking procedure results in 125 portfolios, with each portfolio having a distinct combination of size, book-to-market and momentum characteristics. Next, each stock is assigned to a particular benchmark portfolio depending upon its size, book-to-market and momentum characteristics.⁶ These data are used to estimate the returns to selectivity, timing and style, which are in turn aggregated into total portfolio returns.

⁴ The DGTW benchmarks are available at <http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm> - date accessed 01/16/2013. We thank Russ Wermers for making these data available.

⁵ We note that this database has been extensively used for holdings-based analyses but, unfortunately only includes information on US equity holdings. For our sample, this means approximately 8.85% of all holdings are not accounted for. We refer interested readers to Hunter (2012) for a discussion on unmapped portfolio holdings.

⁶ The returns on the benchmark portfolios and on the stock assignments to each portfolio are available at <http://www.rhsmith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm> - Date accessed 01/16/2013. We are grateful to Russ Wermers for making these data available.

4.1.1 Characteristic Selectivity (CS)

The CS measure analyses a fund manager's stock selection ability, controlling for the characteristics of the stocks in the fund. First, each stock in the fund is assigned to one of the 125 benchmark portfolios. Each stock's return is then taken in excess of the appropriate benchmark portfolio's return. This excess return is then multiplied by the weight of that stock in the portfolio,⁷ and the weighted returns of all stock are summed to obtain the benchmark-adjusted return for the fund. The process is repeated every quarter. We follow Wermers (2000) in calculating the quarter t constituent of the CS measure as:

$$CS_t = \sum_{j=1}^N \tilde{w}_{j,t-1} (\tilde{R}_{j,t} - \tilde{R}_t^{b_{j,t-1}})$$

Where:

$\tilde{w}_{j,t-1}$ = Portfolio weight on stock j at the end of quarter $t-1$;

$\tilde{R}_{j,t}$ = Quarter t return of stock j ; and,

$\tilde{R}_t^{b_{j,t-1}}$ = Quarter t return of the benchmark portfolio (matched to stock j at quarter $t-1$).

The fund's overall CS measure is the average CS measure over all quarters that the fund exists. A significant positive (negative) CS measure indicates superior (inferior) stock selection ability: this manager has been able to earn a return above (below) that of stocks with the same size, book-to-market and prior-year return characteristics.

4.1.2 Characteristic Timing (CT)

The CT measure is used to analyse a fund manager's success at timing investment characteristics. If the returns to size, book-to-market and momentum strategies are time varying, a fund manager with timing ability may be able to earn higher returns by "tilting"

⁷ The portfolio and stock weights are defined relative to holdings for which data are available.

portfolio weights on stocks with certain characteristics when the payoff is highest. The quarter t component of the CT measure is calculated as follows:

$$CT_t = \sum_{j=1}^N (\tilde{w}_{j,t-1} \tilde{R}_t^{b_{j,t-1}} - \tilde{w}_{j,t-5} \tilde{R}_t^{b_{j,t-5}})$$

Where:

$\tilde{w}_{j,t-i}$ = Portfolio weight on stock j at quarter $t-i$;

$\tilde{R}_t^{b_{j,t-i}}$ = Quarter t return of the benchmark portfolio (matched to stock j at quarter $t-i$).

A manager with timing ability would change the portfolio composition at quarter $t-1$ to hold (not hold) stocks with particular characteristics when those characteristics' payoffs are high (low) in the following period. Removing the quarter $t-5$ buy-and-hold benchmark return eliminates fund returns that are attributable to fund style (captured by the AS measure), and allows us to isolate a fund manager's characteristic timing ability. The overall CT measure for the fund is the average CT over all quarters that the fund exists.

4.1.3 Average Style (AS)

AS measures the returns attributable to the fund's tendency to hold stocks with particular size, book-to-market and prior-year return characteristics. The quarter t component of AS is calculated as follows:

$$AS_t = \sum_{j=1}^N \tilde{w}_{j,t-5} \tilde{R}_t^{b_{j,t-5}}$$

Where:

$\tilde{w}_{j,t-5}$ = Portfolio weight on stock j at the end of quarter $t-5$;

$\tilde{R}_t^{b_{j,t-5}}$ = Quarter t return of the benchmark portfolio (matched to stock j at quarter $t-5$).

The average AS measure over all quarters is the AS measure for the fund. Lagging weights and benchmark portfolio returns ensures that returns due to timing stock characteristics are eliminated.

4.1.4 Fund Total Return

We calculate each fund's total return in quarter t as the sum of its Characteristic Selectivity (CS), Characteristic Timing (CT) and Average Style (AS).

4.2 Matching SRI and conventional funds

Once we have calculated CS, CT, AS and total returns for all funds, we wish to determine if there are any differences between our SRI and conventional funds. The literature has typically either compared SRI funds with all conventional funds (see, for example, Bauer et al., 2005, Benson et al., 2006, Renneboog et al., 2008), or matched SRI funds to one or a number of conventional funds (see, for example, Mallin et al., 1995; Gregory et al., 1997). We use a matching approach to obtain our comparison group of conventional funds. The disadvantage of comparing SRI funds with all conventional funds is that SRI funds are known to be smaller and younger (Mallin et al. 1995; Bauer et al. 2005; Bollen, 2007), and fund characteristics such as age, size and investment style have been shown to impact performance (see Otten and Bams, 2002; Bauer et al., 2005; Chen et al., 2004 and Bollen, 2007). Since we desire to examine the performance impact of the SRI investment process, it is important any performance differentials between SRI and conventional funds are attributable to the fund being SRI or not, and not to systematic differences in other fund characteristics.

We initially identify all conventional funds with the same investment objective, based on the categories Capital Appreciation, Equity Income, Growth and Income, Growth, Micro-cap, Mid-cap and Small-cap. We then isolate conventional funds of the same objective with

an equivalent fund age, calculated as the difference between a fund's oldest share class and 31 December 2010. We follow Bollen (2007) in allowing matching funds to have an age within three years of the SRI fund. Finally, we locate the conventional funds from within the sample filtered by objectives and age that have the closest fund size. Following Cremers and Petajisto (2009), we estimate fund size by taking the aggregate of the total net assets (TNA) of all share classes for a fund in a given year.⁸ Similar to Bollen (2007), we perform the matching process each year, which we later show matters to the results.

We match each SRI fund to a portfolio comprising up to three conventional funds. In choosing the number of conventional funds to include in the matching portfolio, there is a trade-off between diversifying any noise or further conventional fund-specific differences, and decreasing the quality of the matches as more conventional funds are added to the portfolio. We examine the impact of increasing the number of conventional funds matched to each SRI fund and find that the quality of the match deteriorates substantially after three funds. Hence our default is to match each SRI fund to three conventional funds. However, in some instances we match one SRI fund to two conventional funds as it is not possible to find three conventional funds which satisfy the matching criteria. We provide a battery of robustness checks to ensure that our results are not driven by the matching criteria. Our analysis of the matching method is itself a contribution to the literature, as much of the prior research has simply adopted one matching technique.

4.3 Statistical tests

We initially investigate differences between SRI and conventional funds using t tests between means. Means and standard errors are estimated by pooling returns for all funds over the period of interest. However, this test requires strong assumptions, including normality and

⁸ In instances where fund size information is unavailable from CRSP, we supplement with fund TNA information from Morningstar. We examine a sample of funds for which data was available from both databases and find that fund size reported by CRSP and Morningstar are compatible.

equality of variances. Quantile-quantile plots (available upon request) suggest that our data may not be normal. We therefore also report test statistics from the Wilcoxon rank sum test between medians, which does not assume normality, and Welch's t -test, which does not require equal variances. Differences are tested each year, and also over our full 2000 to 2010 sample period.

5. Results

Table 1 presents descriptive statistics on the characteristics of the stocks held by our samples of SRI and conventional funds. It is clear from the table that SRI and conventional funds differ in the types of stocks they hold. In particular, SRI fund portfolios comprise stocks that are larger, have higher book-to-market ratios (value stocks); and a higher prior-year returns than conventional funds. Bollen (2007) and Bauer (2005) also find that SRI funds tend to hold larger capitalization stocks than conventional funds. Hong and Kacperczyk (2009) document that sin stocks, which would be excluded from SRI funds' portfolios, tend to have lower book-to-market ratios. However, Bollen (2007) finds SRI funds to have *less* exposure to momentum. We now turn to our main analysis.

5.1 Fund Total Returns

We commence by comparing SRI and conventional funds' total returns, and results are presented in Table 2. SRI funds outperformed conventional funds in seven of the eleven years. The t test between means, Wilcoxon test and Welch test provide similar results in terms of sign and significance each year and overall. Over the sample period, SRI funds outperformed by a statistically and economically significant 0.35% per quarter (approximately 1.40% per annum).

These results are very different than prior literature which has not found any difference in the after fees and expenses performance of SRI and conventional funds (see for example, Hamilton et al. 1993; Sauer 1997; Statman 2000; Bauer et al. 2005; Kreander et al. 2005; Benson et al., 2006). We are, however, primarily interested in the source of this performance differential, so we now turn to an analysis of the DGTW components.

5.2 Characteristic Selectivity

Table 3 documents that both SRI and conventional funds exhibit significantly positive CS measures, overall and in almost all individual years. The outperformance amounts to 0.84% per quarter for SRI funds and 0.85% per quarter for conventional funds; or an annual return of about 3.4%. This finding is consistent with the existing literature that finds evidence of fund managers possessing positive selection abilities (see, for example, Chen et al. 1992; Bello and Janjigian 1997; Kreander et al. 2005). More importantly, our finding is consistent with that of DGTW, who also find managers possess ability according to the CS measure.

In terms of *differences* between the two groups, we do not find evidence of SRI funds and conventional funds exhibiting differing security selection ability over the full period. In terms of individual years, SRI managers exhibited superior skills during 2008, but inferior skills in 2010. However, the majority of CS measures in each individual year are not statistically significant across any of our tests.

We therefore conclude that there is no evidence of any difference in the stock selection ability of SRI and conventional managers. The investment screening process utilized by SRI fund managers appears to neither generate any value relevant information, nor hinder the stock selection process.

5.3 Characteristic Timing (CT)

Characteristic timing results are presented in Table 4. SRI funds outperform conventional funds in characteristic timing by a sizable 0.52% per quarter (approximately 2.08% per annum), with differences significant at the 1% level across all of our difference tests. This result more than accounts for the outperformance of SRI funds at the total fund level. However, the result seems to be driven more by the failure of matched conventional fund managers than the success of SRI managers. While SRI managers exhibit significantly positive timing ability overall, the magnitude of the CT coefficient is relatively small at 0.13% per quarter (approximately 0.52% per annum). In comparison, conventional managers demonstrate significant negative timing ability, with CT measures that average -0.39% per quarter. SRI managers' CT measures in each of the individual years are mostly significantly positive, but small in magnitude compared to the predominantly significantly negative conventional CT measures.

We are surprised by the differences between the two groups, and the natural question is whether we have truly found a group of managers who are better at timing style. Prior literature suggests that fund managers as a group do not possess significant timing abilities (see, for example, Treynor and Mazuy 1966; Henriksson 1984; Jagannathan and Koraczuk 1986), although Jiang et al. (2007) show that timing results are dependent on the methodology used. Those authors document insignificant negative timing abilities with returns-based timing measures, but significantly positive timing abilities using holdings-based measures. Further, it must be remembered that we are examining characteristics timing (switching between stocks of particular styles), and not market timing (switching between stocks and bonds), as prior literature has typically investigated.

We decide to perform further analysis on CT and investigate the *source* of timing ability to ascertain which of the three characteristics (size, book-to-market and momentum) is

driving the results. We perform two tests. The first involves correlating changes in the average characteristic exposure scores reported in Table 1 with returns on the size, book-to-market and momentum factors (SMB, HML and UMD).⁹ We examine the correlations of the Fama French factors with changes in: SRI fund scores; conventional fund scores and the difference between the SRI and conventional scores.¹⁰

The correlation with SRI and conventional funds' size characteristic scores and SMB is positive and similar in magnitude, at 0.20 and 0.19 respectively. Given that a *negative* correlation would indicate timing ability in this instance, we conclude that neither group of managers is able to time the size characteristic.¹¹ Further, the correlation between the *difference* in SRI and conventional size score and SMB is -0.03. This indicates little difference in the ability of SRI and conventional funds to time size, although SRI funds fared slightly better on this aspect.

Both SRI and conventional fund managers exhibit positive book-to-market timing ability, with correlations of 0.78 and 0.20 respectively between changes in the book-to-market characteristic scores and HML. Further, the correlation for the difference between the two scores and HML is 0.28. This indicates that SRI funds were better at timing book-to-market stocks than conventional funds. In fact, this difference correlation is the largest of the three characteristics. This suggests that the positive timing result for SRI funds is largely attributable to timing book-to-market.

We find that neither group is able to time momentum with correlations between the momentum characteristic score and UMD of -0.08 and -0.12 for SRI and conventional funds

⁹ We acknowledge that breakpoints for the size, book-to-market and momentum factors are not the same as those of the DGTW (1997) analysis.

¹⁰ The difference is computed as: (Change in SRI score) – (Change in conventional score).

¹¹ The DGTW scores are such that a larger score would indicate higher exposures to a factor hence larger size, higher book-to-market ratio and higher prior year returns. On the other hand, the Fama-French factors are designed such that a positive return would mean that the following stock characteristics outperformed: smaller stocks (SMB), value stocks (HML) and momentum stocks (UMD). Hence, timing ability would be indicated by 1) negative correlation with the size factor, 2) positive correlation with the book-to-market factor and 3) positive correlation with the momentum factor.

respectively. Both groups of managers seem to have decreased their exposure to momentum stocks when those stocks performed well. The correlation between the difference in scores and UMD is a paltry 0.04, suggesting little differentiation in their ability to time momentum.

In a second set of tests we examine the “success rate” of the timing strategies used by SRI and conventional funds relative to the SMB, HML and UMD factor performance. We deem a strategy successful if the fund is able to increase (decrease) its factor exposure when the respective factor has a positive (negative) return. This allows us to establish the extent to which timing skills are consistent over time. For the size characteristic, SRI funds were successful at timing in four out of ten years, while conventional funds were successful in five out of ten years. In terms of book-to-market, SRI funds made successful changes in a staggering nine out of ten years, while conventional funds were successful in only six years. For momentum, conventional (SRI funds) were successful in five (four) out of ten years. These results corroborate our correlation tests which suggested little difference in the two groups’ abilities to time size and momentum. However, SRI managers may exhibit ability to time book-to-market. Why this may be the case is not readily apparent: we can offer no prior reason for expecting that SRI managers might possess skill in timing exposure to value stocks.

5.4 Average Style (AS)

Average style measures the component of a fund’s return attributable to its tendency to hold stocks with particular size, book-to-market and momentum characteristics. Results appear in Table 5. Both groups of funds delivered significantly positive AS measures over the full period, and in seven of the eleven years. Our primary interest is in the differences between SRI and conventional funds. The difference in means over the full period is negative at -0.15% per quarter. This difference is not significant using the two *t*-tests. However, the

statistic on the Wilcoxon rank-sum test is highly significant. Given that our AS measures are not normally distributed, the Wilcoxon test may be more appropriate.¹² SRI funds deliver significantly lower AS measures than conventional funds in five or six years (depending on test), and outperform in one year. We therefore conclude that the types of stocks held by SRI and conventional funds can result in significant differences in the returns of the two groups of funds.

Again, we are interested in what is driving the AS results. In Table 6 we document the average annual return of stocks in each of the size, book-to-market and momentum quintiles. It is clear from the table that over our sample period, higher returns could be obtained by holding smaller stocks and value stocks, but a clear momentum strategy is not identifiable. We also refer the reader back to Table 1, which documents the characteristics of the stocks held by the two groups of funds. From Table 1 it can be seen that SRI funds held stocks with higher scores than matched conventional funds on all factors in every year, i.e. SRI funds were tilted towards stocks of larger size, higher book-to-market ratios and higher prior-year returns than conventional funds. SRI funds' higher exposure to value stocks would increase their returns relative to conventional funds, but returns would be hindered by their exposure to large stocks. We note that the difference in size score between SRI and conventional funds is nearly double that of book-to-market difference scores. We therefore suggest that SRI funds' underperformance on the AS measure is most likely attributable to these funds holding larger stocks in a time when larger stocks gave a lower payoff than smaller stocks.

5.5 An examination of matching

In our initial tests, we matched each of our SRI funds to two or three conventional funds of the same investment objective, and with similar size and age characteristics, and re-matched

¹² Quantile-Quantile plots (available upon request) indicate the AS measures do not follow a normal distribution.

every year. In this section we investigate whether performing the matching in different ways has an impact on the results. This is important because we of course wish to attribute our findings to the SRI characteristic, and ensure results are not caused by using a poorly specified control group of conventional funds.

We recalculate the total return, CS, CT and AS measures of the two groups using differing matching techniques. We begin by matching each SRI fund to the single best-matched conventional fund (rather than two or three funds) in each year. Next we match only once during the sample period, at the beginning (2000), middle (2004) and end (2010) of our sample period. Matching once has been the approach typically taken in the literature. For instance, Gregory et al. (1997) match one SRI fund to one conventional fund at the start of the sample period, while Kreander et al. (2005) match one SRI fund to one conventional fund in the middle of the sample period. Some authors (for example Mallin et al., 1995 and Statman, 2000) do not document at which point in time size is matched, but presumably only match once. We believe that one-off matching may be problematic. Some funds display tremendous changes in size over our 11-year sample period. Funds may also change their investment objective over the sample period. This will result in the SRI fund being matched to an inappropriate conventional fund for part of the analysis period.

For the sake of brevity, we only present the results from the overall sample period (2000-2010) in Table 7.¹³ Overall, the robustness tests are consistent with our main findings. Matching one SRI fund to one conventional fund provides equivalent results, in terms of sign and significance of coefficients, to matching one SRI fund to two or three conventional funds. Our results on CT and AS are the same across all matching techniques. Hence we conclude the findings that SRI fund managers are better than conventional fund managers in

¹³ Full results available upon request.

characteristic timing, but SRI fund returns suffered on fund style, are robust to the matching technique used.

However, the CS results are not homogenous across the different matching techniques. In particular, matching at the beginning or middle of our sample period produces meaningful changes in the CS results. It is likely that matching becomes worse the more distant in time one is from the point at which the match was made (see Kreander et al., 2005). We investigate this proposition by removing the last four years of data points from the sample matched at the beginning of the period, and the first two and last two years of data from the sample matched in the middle. Removing these far-away data points provides results that are consistent with the initial specification: there is no significant difference in SRI and conventional managers' CS. Consequently, we conclude that the differences in results are more likely attributable to shortcomings with the one-off matching technique, than lack of robustness in the results. The results also indicate that the choice of the matched-pairs methodology is vital in performance evaluation; and that matching on a yearly basis is a more robust method. In fact, our findings cast doubt on the reliability of previous studies that use one-off matching.

6. Conclusion

We have examined how constraining a manager's investible universe may influence the generation of active returns. We compare a matched sample of socially responsible (SRI) and conventional funds using Daniel, Grinblatt, Titman and Wermers' (1997) approach, which divides fund returns into three components: stock selection ability; characteristic timing ability, and fund style. Ours is the first study to provide a holdings-based analysis of SRI fund returns.

We find that SRI funds outperformed a matched sample of conventional funds over the period 2000 to 2010. Closer analysis indicates that while both SRI and conventional

managers exhibit stock selection ability, there is no *difference* in the two groups' abilities. We therefore conclude that constraining portfolios, along responsible criteria at least, neither hinders nor enhances managers' abilities.

Rather, the sources of return difference appear related to style exposures. The outperformance by SRI funds largely stems from style timing. While our matched sample of conventional managers generated significant negative returns from style timing, SRI managers generated a moderate positive return from being better able to time characteristics, and in particular the book-to-market characteristic, than conventional fund managers. We are surprised by this finding, as there does not appear to be any *a priori* reason as to why this should be the case. Perhaps our finding is sample specific, and we leave to future research to test the result in other sample periods.

On the other hand, the Average Style exposure of SRI funds made a moderately negative contribution to their performance relative to conventional funds. This seems to be primarily due to exposure to large capitalisation stocks. The nature of SRI funds' screens mean that we would expect the portfolios to be biased towards large capitalisation stocks. Firms that are able to meet the hurdles imposed by positive screening are likely to be large. In the literature, this has been attributable to the fact that it is the large firms which have the resources to spend on corporate responsibility (Freeman, 1984), and, further, larger firms are more in the public's view, so may face pressure from external stakeholders to behave responsibly (Moskowitz, 1972; Ullmann, 1985).

We also demonstrate that the method used for matching SRI to conventional funds can matter. In particular, in a sample that spans a decade or more, one-off matching is not appropriate. The development of a fund over time may be such that the funds' characteristics (for example its size or investment style) may change markedly over a given sample period, resulting in the control fund looking very different than the treatment fund at points in time

distant from the matching date. This can have consequences for the research findings. We recommending matching be performed, if not annually as we have done, at least a number of times over the sample period.

In conclusion, then, managerial ability is not hampered by having an investible universe screened for responsible criteria. Indeed, we even found some evidence of SRI managers exhibiting superior timing ability to conventional fund managers. However, the SRI criteria does appear to result in large stocks being overweight in the portfolio, which is not desirable in periods where large stocks underperformed.

References

- Banz, R., 1981, "The relationship between return and market value of common stock", *Journal of Financial Economics*, Vol. 9, pp. 3-18
- Barnett, M., Salomon, R., 2006, "Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance", *Strategic Management Journal*, Vol. 27, pp. 1101-1122
- Barras, L., Scaillet, O., Wermers, R., 2010. "False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas." *The Journal of Finance*, Vol. 65, pp. 179-216
- Bauer, R., Koedijk, K., Otten, R., 2005, "International evidence on ethical mutual fund performance and investment style", *Journal of Banking & Finance*, Vol. 29, pp. 1751-1767
- Bello, Z., 2005, "Socially responsible investing and portfolio diversification", *The Journal of Financial Research*, Vol. 28, pp. 41-57
- Bello, Z., Janjigian, V., 1997, "A reexamination of the market timing and security-selection performance of mutual funds", *Financial Analysts Journal*, Vol. 53, pp. 24-30
- Benson, K., Brailsford, T., Humphrey, J., 2006, "Do socially responsible fund managers really invest differently?", *Journal of Business Ethics*, Vol. 65, pp. 337-357
- Bollen, N., 2007, "Mutual fund attributes and investor behaviour", *Journal of Financial and Quantitative Analysis*, Vol. 42, pp. 683-708
- Busse, J., Tong, Q., 2012, "Mutual Industry Selection and Persistence", *The Review of Asset Pricing Studies*, Vol. 2, pp. 245-274
- Carhart, M., 1997, "On persistence in mutual fund performance", *Journal of Finance*, Vol. 52, pp. 57-82
- Chang, E., Lewellen, W., 1987, "Market timing and mutual fund investment performance", *Journal of Business*, Vol. 57, pp. 57-72
- Cremers, M., Petajisto, A. and Zitzewitz, E., 2012, "Should Benchmark Indices Have Alpha? Revisiting Performance Evaluation", *Critical Finance Review* (forthcoming)
- Cremers, M., Petajisto, A., 2009, "How active is your fund manager? A new measure that predicts performance", *Review of Financial Studies*, Vol. 22, pp. 3329-3365
- Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997, "Measuring mutual fund performance with characteristic-based benchmarks", *Journal of Finance*, Vol. 52, pp. 1035-1058

Elton, E., Gruber, M., 2011, "Mutual funds", in Constantinides, G., Harris, M., Stulz, R., *Handbook of the economics of finance: financial markets and asset pricing volume 1b*, The Netherlands: Amsterdam

Fabozzi, F.J., Ma, K.C., Oliphant, B. J., 2008, "Sin Stock Returns", *Journal of Portfolio Management* 35(1), 82-94.

Gregory, A., Matako, J., Luther, R., 1997, "Ethical unit trust financial performance: small company effects and fund size effects", *Journal of Business Finance and Accounting*, Vol. 24, pp. 705-725

Grinblatt, M., Titman, S., 1994, "A study of monthly mutual fund returns and performance evaluation techniques", *Journal of Financial and Quantitative Analysis*, Vol. 29, pp. 419-444

Grinblatt, M., Titman, S., Wermers, R., 1995, "Momentum investment strategies, portfolio performance, and herding: a study of mutual fund behaviour", *The American Economic Review*, Vol. 85, pp. 1088-1105

Hamilton, S., Jo, H., Statman, M., 1993, "Doing well while doing good? The investment performance of socially responsible mutual funds", *Financial Analysts Journal*. Vol. 49, pp. 62-66

Henriksson, R., Merton, R., 1981, "On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills", *Journal of Business*, Vol. 54, pp. 513-533

Henriksson, R., 1984, "Market timing and mutual fund performance: an empirical investigation", *The Journal of Business*, Vol. 57, pp. 73-96

Hunter, D., 2012, "Unmapped holdings and the performance measurement of U.S. equity mutual funds", working paper, University of Hawai'i.

Jagannathan, R., Korajczyk, R., 1986, "Assessing the market timing performance of managed portfolios", *Journal of Business*, Vol. 59, pp. 217-235

Jegadeesh, N., Titman, S., 1993, "Returns to buying winners and selling losers: implications for stock market efficiency", *Journal of Finance*, Vol. 48, pp. 65-91

Jegadeesh, N., Titman, S., 2001, "Profitability of momentum strategies: An evaluation of alternative explanations", *Journal of Finance*, Vol. 56, pp. 699-720

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

Jiang, G., Yao, T., Yu, T., 2007, "Do mutual funds time the market? Evidence from portfolio holdings", *Journal of Financial Economics*, Vol. 86, pp. 724-758

- Kempf, A., Osthoff, P., 2008, "SRI Funds: Nomen est Omen", *Journal of Business Finance & Accounting*, Vol. 35, pp. 1276-1294
- Kothari, S., and Warner, J., 2001, "Evaluating mutual fund performance", *Journal of Finance*, Vol. 56, pp. 1985-2010
- Kreander, N., Gray, R., Power, D., Sinclair, C., 2005, "Evaluating the performance of ethical and non-ethical funds: a matched pair analysis", *Journal of Business Finance & Accounting*, Vol. 32, pp. 1465-1493
- Mallin, C., Saadouni, B., Briston, R., 1995, "The financial performance of ethical investment funds", *Journal of Business Finance & Accounting*, Vol. 22, pp. 483-496
- Otten, R., Bams, D., 2002, "European mutual fund performance", *European Financial Management*, Vol. 8, pp. 75-101
- Renneboog, L., Horst, J., Zhang, C., 2008, "The price of ethics and stakeholder governance: the performance of socially responsible mutual funds", *Journal of Corporate Finance*, Vol. 14, pp. 302-322
- Sauer, D., 1997, "The impact of social-responsibility screens on investment performance: Evidence from the Domini 400 social index and Domini Equity Mutual Fund", *Review of Financial Economics*, Vol. 6, pp. 137-149
- Schroder, M., 2004, "The performance of socially responsible investments: investment funds and indices", *Financial Markets and Portfolio Management*, Vol. 18, pp. 122-142
- Statman, M., 2000, "Socially responsible mutual funds", *Financial Analysts Journal*, Vol. 56, pp. 30-39
- Treynor, J., Mazuy, K., 1966, "Can mutual funds outguess the market?", *Harvard Business Review*, Vol. 45, pp. 131-136
- Ullmann, A.A., 1985. "Data in search of a theory: a critical examination of the relationships among social performance, social disclosure, and economic performance of U.S. firms." *Academy of Management Review* 10, 540-557.
- U.S. Social Investment Forum, 1999, "1999 Trends Report on Responsible Investing Trends in the U.S. 1999", *U.S. SIF*, Washington, DC
- U.S. Social Investment Forum, 2007, "2007 Trends Report on Responsible Investing Trends in the U.S. 2007", *U.S. SIF*, Washington, DC
- U.S. Social Investment Forum, 2010, "2010 Trends Report on Responsible Investing Trends in the U.S. 2010", *U.S. SIF*, Washington, DC
- U.S. SIF, 2012, "2012 Trends Report on Responsible Investing Trends in the U.S. 2012", *U.S. SIF*, Washington, DC

Wermers, R., 2000, "Mutual fund performance: an empirical decomposition into stock-picking talent, style, transaction costs, and expenses", *Journal of Finance*, Vol. 55, pp. 1655-1703

Wermers, R., 2011, "Performance measurement of mutual funds, hedge funds, and institutional accounts", *Annual Review of Financial Economics*, pp. 537-574

Table 1: SRI and conventional fund characteristics

This table presents the scores for SRI and conventional funds across market value (size), book-to-market ratio and prior-year returns. Each stock is assigned a score of 1 (low) to 5 (high). For each mutual fund, the portfolio-weighted average characteristic score is computed each year for each characteristic. The characteristic scores are then averaged across all SRI or conventional funds each year. The overall score is the time-series average over the entire sample period. The sample period is 2000 to 2010.

Characteristic	Category	Overall	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Size	SRI	4.09	4.31	4.23	4.13	4.12	4.03	4.03	4.02	3.98	4.05	4.14	4.07
	Conventional	3.56	3.71	3.86	3.79	3.67	3.54	3.52	3.50	3.49	3.42	3.46	3.43
	Difference	0.53	0.60	0.37	0.33	0.45	0.49	0.51	0.52	0.49	0.63	0.68	0.64
Book-to-market	SRI	2.56	2.40	2.52	2.45	2.47	2.55	2.55	2.62	2.61	2.56	2.64	2.66
	Conventional	2.31	2.15	2.37	2.27	2.36	2.24	2.29	2.40	2.30	2.30	2.30	2.39
	Difference	0.25	0.25	0.15	0.18	0.11	0.31	0.26	0.22	0.31	0.26	0.35	0.27
Prior year return	SRI	3.10	3.48	2.97	2.95	3.26	3.14	3.05	3.03	3.05	3.23	3.08	2.97
	Conventional	2.67	3.05	2.58	2.62	2.80	2.63	2.60	2.58	2.68	2.69	2.67	2.56
	Difference	0.43	0.43	0.39	0.33	0.46	0.52	0.45	0.45	0.37	0.54	0.41	0.40

Table 2: Total returns (per quarter)

This table presents SRI and conventional funds' total returns. The total return for a fund at quarter t is the aggregation of the quarter t CS, CT and AS measures. Total returns are equally weighted across SRI/conventional funds each year. Time-series t -statistics are shown in parentheses. Difference in means is computed with paired t -tests. The Z-statistic is the Wilcoxon rank-sum test for differences in medians. The Welch's test is a t -test for differences in means for samples with possibly unequal variances. The sample period is 2000 to 2010.

Year	SRI funds	Conventional funds	Difference in means	Z-statistic	Welch's Test t-stat
Average (2000-2010)	1.56%*** (11.65)	1.21%*** (12.58)	0.35%** (2.36)	3.62***	2.20**
2000	1.66%** (2.37)	1.89%*** (5.44)	-0.23% (-0.31)	-0.53	-0.30
2001	0.49% (0.89)	1.31%*** (3.45)	-0.82% (-1.33)	-1.35	-1.22
2002	-1.18%*** (-3.14)	0.51%** (1.99)	-1.69%*** (-3.81)	-3.45***	-3.72***
2003	4.36%*** (11.33)	1.09%*** (3.37)	3.27%*** (7.14)	5.60***	6.52***
2004	2.52%*** (10.12)	0.75%*** (3.91)	1.77%*** (5.47)	4.80***	5.63***
2005	1.55%*** (6.66)	1.08%*** (7.05)	0.46% (1.81)	1.81**	1.66
2006	2.02%*** (11.85)	0.77%*** (4.77)	1.25%*** (5.40)	4.67***	5.30***
2007	1.45%*** (7.26)	1.21%*** (5.49)	0.24% (0.93)	0.69	0.80
2008	-3.24%*** (-8.12)	0.21% (0.54)	-3.45%*** (-6.11)	-5.23***	-6.25***
2009	4.98%*** (14.75)	2.47%*** (7.02)	2.51%*** (4.87)	4.95***	5.15***
2010	2.50%*** (10.33)	2.13%*** (4.93)	0.37% (0.73)	2.61***	0.74

*, ** and *** denote significance at 10%, 5% and 1% respectively.

Table 3: Stock selection ability (CS measure, per quarter)

This table provides results on Characteristic Selectivity measures. Each stock's CS measure is the difference between the buy-and-hold return of the stock minus the buy-and-hold return of the quarter $t-1$ benchmark portfolio. The CS measure for each fund is calculated by value-weighting the CS measures of all stocks in its portfolio in every quarter. The average CS measure is then taken of all SRI or conventional funds in each year. Time-series t -statistics are shown in parentheses. Difference in means is computed with paired t -tests. The Z-statistic is the Wilcoxon rank-sum test for differences in medians. The Welch's test is a t -test for differences in means for samples with possibly unequal variances. The sample period is 2000 to 2010.

Year	SRI funds	Conventional funds	Difference in means	Z-statistic	Welch's Test t-stat
Average (2000-2010)	0.84%*** (10.54)	0.85%*** (11.95)	-0.01% (-0.08)	-0.07	-0.08
2000	1.64%** (2.57)	1.32%*** (4.77)	0.32% (0.48)	0.75	0.46
2001	1.10%*** (3.19)	0.90%*** (2.87)	0.19% (0.43)	0.38	0.42
2002	0.69%*** (3.22)	0.35%** (1.97)	0.34% (1.19)	1.27	1.22
2003	0.91%** (2.56)	0.89%*** (4.34)	0.02% (0.05)	-0.75	0.04
2004	0.65%*** (4.00)	0.55%*** (4.52)	0.10% (0.44)	0.09	0.21
2005	0.72%*** (3.80)	0.73%*** (7.73)	-0.01% (-0.06)	-0.66	-0.03
2006	0.46%*** (3.77)	0.52%*** (5.06)	-0.06% (-0.40)	0.66	-0.06
2007	0.78%*** (4.30)	0.82%*** (5.87)	-0.04% (-0.20)	-0.93	-0.55
2008	1.28%**** (5.47)	0.03% (0.14)	1.24%*** (3.78)	3.19***	3.36***
2009	0.99%*** (4.23)	1.53%*** (7.26)	-0.54% (-1.60)	-2.10**	-1.50
2010	0.24% (1.63)	1.69%*** (4.23)	-1.45%*** (-3.46)	-3.63***	-3.74***

*, ** and *** denote significance at 10%, 5% and 1% respectively.

Table 4: Characteristic timing ability (CT measure, % per quarter)

This table provides SRI and conventional funds' CT measures. The CT measure for a stock is the difference between the buy-and-hold benchmark return of stock j at quarter $t-1$ minus the buy-and-hold benchmark return at quarter $t-5$. The CT measure for a fund is calculated on a value-weighted basis and all portfolios are rebalanced quarterly. Finally, the CT measures are averaged across SRI/conventional funds each year. Time-series t-statistics are shown in parentheses. Difference in means is computed with paired t-tests. The Z-statistic is the Wilcoxon rank-sum test for differences in medians. The Welch's test is a t-test for differences in means for samples with possibly unequal variances. The sample period is 2000 to 2010.

Year	SRI funds	Conventional funds	Difference in means	Z-statistic	Welch's Test t-stat
Average (2000-2010)	0.13%*** (3.06)	-0.39%*** (-2.97)	0.52%*** (3.48)	6.65***	3.76***
2000	0.53%* (1.88)	1.07%*** (3.74)	-0.54% (-1.38)	-1.73**	-1.34
2001	-0.09% (-0.53)	1.29%*** (2.81)	-1.38%*** (-3.19)	-2.85**	-2.82***
2002	-0.09% (-0.69)	2.79%*** (9.54)	-2.89%*** (-9.04)	-5.75***	-8.96***
2003	0.40%*** (3.17)	-3.91%*** (-12.87)	4.31%*** (13.76)	6.85***	13.11***
2004	0.17%** (2.33)	-1.45%*** (-12.81)	1.62%*** (13.29)	6.88***	12.04***
2005	-0.06% (-0.97)	-0.74%*** (-8.92)	0.68%*** (6.68)	5.25***	6.80***
2006	0.20%*** (3.50)	-1.90%*** (-19.27)	2.09%*** (17.61)	7.37***	18.49***
2007	0.35%*** (4.67)	-0.34%*** (-3.50)	0.69%*** (5.13)	4.50***	5.61***
2008	-1.03%*** (-6.32)	6.34%*** (22.13)	-7.37%*** (-22.04)	-7.57***	-22.36***
2009	0.70%*** (4.52)	-3.50%*** (-14.68)	4.20%*** (14.92)	7.46***	14.78***
2010	0.36%*** (4.91)	-2.25%*** (-13.04)	2.61%*** (14.27)	7.59***	13.92***

*, ** and *** denotes significance at 10%, 5% and 1% respectively.

Table 5: Average style (AS measure, % per quarter)

This table provides a comparison of fund styles, using AS measures for SRI and conventional funds. The AS measure for a stock is the difference between the quarter $t - 5$ buy-and-hold benchmark return of stock j minus the buy-and-hold quarter $t - 5$ benchmark return. The AS measure for a fund is calculated on a value-weighted basis and all portfolios are rebalanced quarterly. Finally, the CT measures are averaged across SRI/conventional funds each year and presented as follows. Time-series t -statistics are shown in parentheses. Difference in means is computed with paired t -tests. The Z-statistic is the Wilcoxon rank-sum test for differences in medians. The Welch's test is a t -test for differences in means for samples with possibly unequal variances. The sample period is 2000 to 2010.

Years	SRI funds	Conventional funds	Difference in means	Z-statistic	Welch's Test t-stat
Average (2000-2010)	0.64%*** (6.01)	0.79%*** (6.08)	-0.15% (-1.60)	-2.67***	-0.91
2000	-0.51% (-1.39)	-0.29% (-0.92)	-0.22% (-0.44)	-0.18	-0.46
2001	-0.64% (-1.35)	-0.81%* (-1.77)	0.17% (0.29)	-0.29	0.26
2002	-2.09%*** (-7.20)	-2.60%*** (-8.81)	0.51% (1.41)	1.62	1.23
2003	3.16%*** (10.63)	4.11%*** (15.36)	-0.95%** (-2.26)	-1.85*	-2.38***
2004	1.76%*** (9.01)	1.65%*** (21.25)	0.11% (0.49)	-0.97	0.52
2005	0.92%*** (10.84)	1.11%*** (19.95)	-0.19% (-1.85)	-1.68*	-1.87*
2006	1.41%*** (13.08)	2.16%*** (24.92)	-0.75%*** (-6.33)	-5.18***	-5.39***
2007	0.37%*** (5.69)	0.77%*** (11.21)	-0.40%*** (-4.11)	-3.86***	-4.19***
2008	-3.96%*** (14.95)	-6.15%*** (-27.19)	2.19%*** (6.25)	5.37***	6.27***
2009	3.30%*** (16.32)	4.44%*** (20.73)	-1.15%*** (-4.42)	-4.24***	-3.90***
2010	1.89%*** (10.96)	2.69%*** (17.18)	-0.79%*** (-3.40)	-3.00***	-3.41***

*, ** and *** denotes significance at 10%, 5% and 1% respectively.

Table 6: DGTW characteristic annual returns

This table outlines the average yearly return for the size, book-to-market and momentum portfolios over the sample period 2000-2010.

Portfolio	Size	Book-to-market	Momentum
1	17.29% (Small cap)	7.77% (Growth)	10.42% (Low prior year return)
2	13.77%	10.29%	12.71%
3	11.95%	12.07%	12.75%
4	11.11%	13.11%	13.52%
5	4.52% (Large cap)	15.40% (Value)	9.24% (High prior year return)

Table 7: Changing matching criteria

This table summarises the results from using different criteria to match SRI and conventional funds. Results are from paired *t* tests between means, Wilcoxon rank-sum tests for differences in medians and Welch's *t*-test for differences in means and are at the 5 percent significance level. 0 indicates no statistically significant difference between SRI and conventional funds, (+) indicates that SRI funds significantly outperformed matched conventional funds and (-) indicates that SRI funds significantly underperformed matched conventional funds.

Methodology	CS			CT			AS		
	t	Wilcoxon	Welch	t	Wilcoxon	Welch	t	Wilcoxon	Welch
Initial Specification	0	0	0	+	+	+	0	-	0
1:1 (yearly basis)	0	0	0	+	+	+	0	-	0
1:1 (matched at 2000)	-	-	-	+	+	+	-	-	-
1:1 (matched at 2004)	-	-	-	+	+	+	0	-	0
1:1 (matched 2010)	0	0	0	+	+	+	-	-	0