Sustainable Development
Sust. Dev. **19**, 77–94 (2011)
Published online 16 February 2011 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/sd.509

Assessing SRI Fund Performance Research: Best Practices in Empirical Analysis

Andrea Chegut,[1]* Hans Schenk[2] and Bert Scholtens[3]
[1] *Department of Finance, Maastricht University, Maastricht, The Netherlands*
[2] *Department of Economics, Utrecht University, Utrecht, The Netherlands*
[3] *Department of Economics, Econometrics and Finance, University of Groningen, Groningen, The Netherlands*

ABSTRACT
We review the socially responsible investment (SRI) mutual fund performance literature to provide best practices in SRI performance attribution analysis. Based on meta-ethnography and content analysis, five themes in this literature require specific attention: data quality, social responsibility verification, survivorship bias, benchmarking, and sensitivity and robustness checks. For each of these themes, we develop best practices. Specifically, for sound SRI fund performance analysis, it is important that research pays attention to dividend yields and fees, incorporates independent and third party social responsibility verification, corrects for survivorship bias and tests multiple benchmarks, as well as analyzing the impact of fund composition, management influences and SRI strategies through sensitivity and robustness analysis. These best practices aim to enhance the robustness of SRI financial performance analysis. Copyright © 2011 John Wiley & Sons, Ltd and ERP Environment.

Received 1 September 2009; revised 2 December 2009; accepted 4 January 2010
Keywords: mutual funds; socially responsible investing; performance evaluation; best practices

Introduction

In this paper, we investigate performance attribution analysis with respect to socially responsible investment (SRI). This analysis is relevant in the decision making process of financial institutions in constructing and offering SRI portfolios. Financial portfolio theory and the classical theory of the firm suggest that including non-financial restrictions will not benefit financial performance. Portfolio theory implies that criteria that constrain an investor's investment possibilities result in lower diversification and greater risk exposure or additional costs. The classical theory of the firm implies that SRI will be less financially efficient than non-restricted investments, since the firms that responsible investors do invest in may incur higher costs. This would make these firms less profitable. In contrast, the social theory of the firm suggests that the financial performance of responsible investments is superior to that of conventional investing because the former incorporates information that is more relevant and, thereby, allows better decision making.

*Correspondence to: Andrea Chegut, Department of Finance, Maastricht University, Tongersestraat 53, 6211LM Maastricht, The Netherlands.
E-mail: a.chegut@maastrichtuniversity.nl

To find out how screening for responsibility impacts portfolio performance, empirical studies are useful. Empirical research generally does not arrive at significant differences in the financial performance of responsible and conventional investing (see for example Goldreyer and Diltz, 1999; Statman, 2000; Bauer *et al.*, 2005; Galema *et al.*, 2008). However, SRI empirical research faces several problems, and inconsistent results may have important consequences for mainstreaming SRI investment.

There are three main arguments against mainstreaming SRI funds, which directly relate to how SRI funds are empirically measured. First, there is a suspicion that these portfolios have increased costs and risk due to reduced diversification (Geczy *et al.*, 2005; Renneboog *et al.*, 2006; Cortez *et al.*, 2008). Second, there is a suspicion of increased monitoring costs from SRI managers (Bauer *et al.*, 2007). Third, SRI may lead to decreased returns, leading financial managers to a breach of their fiduciary duty to provide the highest possible return with the lowest possible risk (Schröder, 2004; Bauer *et al.*, 2005). To investigate the impact of these issues, SRI studies employ multiple methods of risk and return analysis, derived mainly from modern portfolio theory. Empirical evaluation techniques employed include capital asset pricing models (CAPMs), multi-index models, multi-factor models and arbitrage pricing theory. As such, SRI studies rely on conventional portfolio evaluation, a body of empirical literature that has taken 50 years to develop and test (for a collection of criticisms see Elton *et al.*, 2006).

The motivation of many SRI studies is to develop estimates of the average returns of a population of SRI funds with low bias and estimation errors (e.g. Bauer *et al.*, 2005). This implies that the SRI fund's empirical average returns must be consistent, i.e. a good estimate of the SRI population's returns, and efficient, i.e. with the smallest possible variance (Greene, 2008). In this respect, accounting for measurement error and misspecification is crucial (Kennedy, 2008).

In the past 15 years, many empirical studies of SRI fund performance have been conducted (see Renneboog *et al.*, 2007, and Hoepner and McMillan, 2008, for an overview). In particular, changes in SRI verification and specification procedures have influenced the development of the SRI research domain.[1] As these changes occurred, researchers incorporated new methodologies, data and specific social responsibility features into their performance assessments. However, there is little explicit knowledge about the best practices within the domain of SRI performance attribution analysis. Renneboog *et al.* (2007) provide an extensive overview of the usage of risk-adjusted performance measures and performance evaluation models in SRI fund performance analysis. Their principal contribution is in appropriate model selection. Our study aims to complement this contribution of Renneboog *et al.* (2007) and to provide an assessment of the best practices that influence SRI fund empirical analysis. More specifically, we investigate non-model specific empirical issues in SRI research. Our study reviews SRI fund performance studies to arrive at recommendations for best practices in empirical analysis, especially practices that aim at minimizing measurement error and misspecification.

To this extent, we use two meta-approaches on 41 SRI fund performance studies. The first meta-approach is content analysis, a quantitative method used to discern common practices in the literature. The second is a meta-ethnographic approach, which is a qualitative method to reveal analogies and demarcations in the literature. From the latter approach, five themes result that repeatedly surface in the SRI literature: (1) data quality; (2) social responsibility verification; (3) survivorship bias; (4) benchmarking and (5) sensitivity and robustness checks. Apart from the second theme, these issues do play a role in conventional financial performance attribution analysis (see Elton *et al.*, 2006). We argue that careful consideration of data quality, social responsibility verification and survivorship bias helps to minimize measurement errors in SRI studies too. Benchmarking as well as sensitivity and robustness analysis are tools that help minimize misspecification. Measurement error can arise in several areas, but in SRI it mainly results from poor data collection and the integrity of responsibility information received from producers and verifiers. In SRI, the accurate measurement of income and fees is critical for having a proper comparison with conventional funds. Furthermore, what constitutes an SRI fund is a categorical issue. Survivorship bias is critical for accounting for surviving and dead income streams. Misspecification may arise from poor matching with conventional funds and inadequate SRI fund specific data controls.

[1] In the special issue (Cerin and Scholtens, 2011), several papers relate responsible investment to different agents. For example, Manescu (2011) investigates the role of financial markets, Scholtens (2011) investigates CSR with insurance companies, Hedesström *et al.* (2011) analyze how information specialists arrive at information about responsible conduct and policies of firms, and Jansson and Biel (2011) look into motives of private and institutional investors to engage with SRI.

DOI: 10.1002/sd

Our study relates to the approaches by Margolis and Walsh (2001, 2003) and Orlitzky *et al.* (2003), who critically investigate the literature about the relationship between corporate social and financial performance. Our study also relates to the work of Hoepner and McMillan (2008), who examine the SRI literature in general, but specifically look into the journals in which SRI studies appear. However, we investigate the SRI research processes and practices and shall not focus on the actual results. As such, we aim to complement the Renneboog *et al.* (2007) study, which reviews various models to assess SRI fund performance.

Based on our analysis, we find that much of the SRI literature is inconsistent in its treatment of data quality, social responsibility verification, survivorship bias, benchmark treatment and robustness analysis. We suggest that future research includes and treats dividend yield and fees in the analysis, incorporates independent and third party social responsibility verification, corrects for survivorship bias, tests multiple benchmarks and analyzes the impact of fund composition, management influences and SRI strategies through sensitivity and robustness checks.

The structure of this paper is as follows. The following section provides the motivation for the specific themes reviewed in this paper. The next section discusses the methodology used to conduct our analysis and the selection of SRI studies. Following this, we present and discuss our results in the fourth section and conclude with their implications in the last section.

Themes

We investigate five themes that are relevant with respect to eliminating measurement bias and estimation error. The categories are data quality, social responsibility verification, survivorship bias, benchmarks and robustness checks. Apart from the verification issue, they are applicable in a more general mutual fund performance analysis context as well (see Elton *et al.*, 2006). We base the selection of the five themes on a meta-ethnographic analysis of the literature. In fact, this analysis yielded six relevant themes. Apart from the five mentioned, it also pointed at model specification. However, as model specification is very well addressed in the study by Renneboog *et al.* (2007) and as it is much more related to modeling than to research processes and practices, we refrain from reviewing this theme in our paper. Next, we motivate the examination of each empirical practice in connection with SRI analysis.

The measurement of income returns and fees is the primary data input for SRI fund performance evaluation models. These data components are at the heart of the SRI managers' fiduciary duty debate and require explicit consideration when conducting performance analysis (Sauer, 1997). Data quality refers to the construction of the data, especially the inclusion or exclusion of fees, dividends or cash payments. Furthermore, it relates to whether these factors are dealt with in an explicit manner. Some papers suggest that SRI funds experience higher fees (Renneboog *et al.*, 2008), while others stress the occurrence of decreased dividends (Stone *et al.*, 2001; Gregory and Whittaker, 2007). Transaction costs outside management fees, such as load fees,[2] are difficult to account for in performance assessments (Bauer *et al.*, 2005; Geczy *et al.*, 2005; Renneboog *et al.*, 2008). However, if and how these accounting items are measured might matter for the SRI funds' bottom line performance.

The verification of socially responsibility relates to whether SRI funds are genuine or just labeled as SRI, and whether they are converging to conventional funds (Benson *et al.*, 2006; Bauer *et al.*, 2007; Kempf *et al.*, 2007; Renneboog *et al.*, 2007; Cortez *et al.*, 2008). This verification issue is very specific to SRI funds. It concerns the confirmation of ethical, environmental and social standards by independent assessment or third party verification.

Failing to account for survivorship bias may result in an overestimation of the mean average returns (Brown *et al.*, 1992; Elton *et al.*, 1996). For instance, Bauer *et al.* (2006) found, in their study of Australian ethical and conventional open-end mutual funds, that restricting the sample to surviving funds alone leads to an overestimation of average returns for domestic funds by 0.20% and for international funds by 1.13% per year.

Grinblatt and Titman (1994) point out the importance of benchmark efficiency. They argue that the choice of the benchmark can have a large and significant impact on conclusions about investment portfolio performance.

[2] According to the SEC, load fees are the commission the shareholder pays to the broker for the acquisition of new assets, which can be deferred until the end of the client–broker relationship or charged directly at each purchase (http://www.sec.gov/answers/mffees.htm 17 July 2008).

DOI: 10.1002/sd

Thus, the specific index chosen, whether SRI or conventional, may affect the evaluation of these funds. Furthermore, when conducting a matched pair analysis, the choice for specific factors to match conventional and SRI portfolios to one another needs careful consideration (Luther and Matatko, 1994).

Sensitivity and robustness checks are quite common in quantitative testing, but within SRI research they have developed a distinctive perspective due to the nature of SRI funds. Considering how style factors change under different models is pertinent to decide on the most accurate specification of SRI performance comparisons.

Methodology

In our review of the SRI fund performance literature, we use two different methods. The first method is content analysis (see, e.g., Kothari, 2004). To demonstrate each empirical practice's systemic reoccurrence and importance, we provide the results of the number of times these practices occur. We opt for content analysis to display basic descriptive statistics on the empirical practices in the literature. Orlitzky *et al.* (2003), among others, have criticized this method. They argue it is prone to bias as the descriptive statistic depends on the size of the sample produced. We use content analysis to categorize the underlying literature into common and varying empirical practices. To account for the criticism of Orlitzky *et al.* (2003), we complement this analysis with the so-called meta-ethnography method (Noblit and Hare, 1988). This method focuses on themes to reveal the analogies or demarcations between the studies we include in the analysis. Like other meta-approaches, meta-ethnography requires that the synthesis of the literature focus on a comparable research question. The objective is to decipher, synthesize and report the relevant themes. We report how often these themes appear in the literature. Furthermore, we utilize the themes to arrive at best practices.

Together, the content analysis and meta-ethnography yield a quantitative and qualitative assessment of the SRI mutual fund performance literature. From the content approach, we report empirical practices used to minimize measurement error and to conduct specification analysis. From meta-ethnography, we arrive at which empirical practices have sustained attention in the literature (see also the previous section).

To eliminate publication bias as much as possible, we searched along the following lines. To begin, we consulted references in the literature. Then, we searched the Google Scholar database on 'ethical investment performance' and 'social responsibility investment performance'. We searched for both terms until all papers containing the topic were exhausted. In addition, we did an internet search to exhaust possible online publications. The studies selected for cataloging rely on the following two criteria. First, we select empirical studies investigating performance of SRI funds[3] or a form of trust. Second, the fund's performance must be available. Following these criteria, we arrived at 41 studies. They are highlighted in the reference list with asterisks (**) next to the author(s). We are aware of the fact that these studies do not span all the SRI literature. However, we feel that they are representative for the literature as a whole because of our selection process.

Of the 41 studies, 33 were in journals, six were working papers and two were in printed sources. In total, they covered periods from July 1963 to February 2007. The longest study period was 39 years and the shortest was 3 years, with an average of 10.4 years. The literature predominantly studies the period from 1990 to 2004 (each year appears at a minimum 15 and at a maximum 24 times.) Thus, about half the studies concentrate on this period. A distribution of the study period by year is in Appendix A. There are 21 different countries included in the studies, as listed in Appendix B. The US is studied the most (25 times), followed by the UK (13 times) and the Netherlands (eight times). There were 22 different data sources used, with the most used data-source CRSP Survivorbias Free US Mutual Fund Database (nine). A distribution of the studies by data source is in Appendix C. As this study is primarily interested in best practices in the SRI fund performance literature and not in individual studies, it does not report the detailed characteristics of all 41 studies. This would result in far too many additional tables and would considerably increase the length of this paper.

[3] Shariah funds were not included in the sample as their portfolio characteristics are more restrictive, i.e. Shariah law compliant. Consequently, their unique form of SRI performance assessment would require specific treatment in the literature.

DOI: 10.1002/sd

Results

This section reports, first, on the results regarding the five key issues: data quality, social responsibility verification, survivor bias, benchmarking and robustness (first five subsections). Then, the last subsection suggests best practices based on these results.

Data Quality

The literature does not universally account for considerations regarding the income and fee data. All studies give the gross or net returns. Twenty studies (49%) provide an explicit description of the return contents, 12 studies (29%) give an explicit consideration of the fund's dividend yields and 15 studies (37%) explicitly mention the transaction costs and management fee. We find that explicit mentioning of load fees occurs in six studies (15%).

Thus, it appears that the inclusion and treatment of the dividend yield and fees have not been very systematic in SRI research so far. The dividend yield has been marginally considered, under the small cap effect and when utilizing conditional strategy models. Regarding fees, the infrequent treatment may result from the focus on US mutual funds. However, load fees require specific treatment as they may be included as front-end fees, or they are not included because they have yet to be charged to the customer, as back-end fees. This is admittedly a quite complex data issue.[4]

Some recent studies consider how fees may vary between investments in different countries. For example, Bauer *et al.* (2006) discern in their study of Australian ethical and conventional open-end mutual funds that domestic ethical fund fees are higher than their domestic conventional peers, but not fees for international funds. Renneboog *et al.* (2008) also conduct a global analysis of funds and discover that fees vary from country to country. They find that total fees are at their lowest in Belgium and The Netherlands (both at 1.3%), and at their highest in Malaysia (at 2.4%).[5] Geczy *et al.* (2005) report the arithmetic average of maximum fund loads between US domestic SRI, which charge a maximum of 4.26%, and conventional funds' load fees, which charge on average a maximum of 3.63%. Renneboog *et al.* (2008) and Geczy *et al.* (2005) also find that fund management fees and load fees, respectively, significantly reduce the risk-adjusted returns of both SRI and conventional funds. However, Gil-Bazo *et al.* (2010) provide evidence that suggests that fees do not significantly affect the performance of US SRI funds.

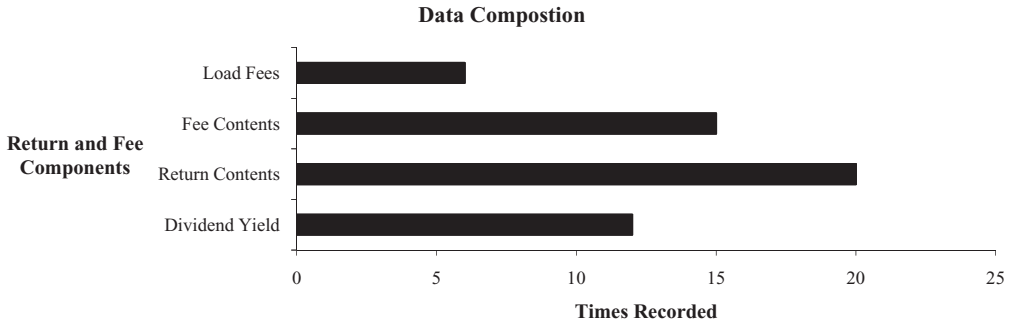


Figure 1. Return and fee components by number of times discussed in the literature

[4]To eliminate the fee issue, Schröder conducted studies on the performance of SRI performance indices relative to a variety of benchmark indices. Performance indices generally express the total return to the investor and include dividend payments, but exclude the need to incorporate fee data, as they are not actively managed (Schröder, 2004). As a result, this has been one method to get around the fee issue. However, this does not resolve the problem for SRI retail mutual funds.

[5]This high rate may be attributable to Malaysia's' Shariah compliant funds. They require considerable monitoring and Shariah law expertise. Considerable attention to the cost of this expertise should be given when drawing conclusions for this specific asset class.

DOI: 10.1002/sd

Social Responsibility Verification

Thirty-three of the 41 studies (81%) take account of social responsibility verification. Verification takes place in one or both of two manners, namely independent verification by the author(s) or verification by a third party source. Verification by the author may occur by interviewing the individual fund managers, reviewing fund websites and reading individual fund prospectuses. This type of verification takes place in seven studies (17%). Verification by a third party source occurs by importing a flag into the dataset, which indicates that the fund is an SRI fund. Rating agency services, research organizations or an independent financial organization that gives an independent brief on what constitutes ethical investment may provide this type of verification. Twenty-one studies used this type of verification (51%). Both independent and third party verification did occur in three studies (7%). For a list of third party verification sources used, see Appendix D.

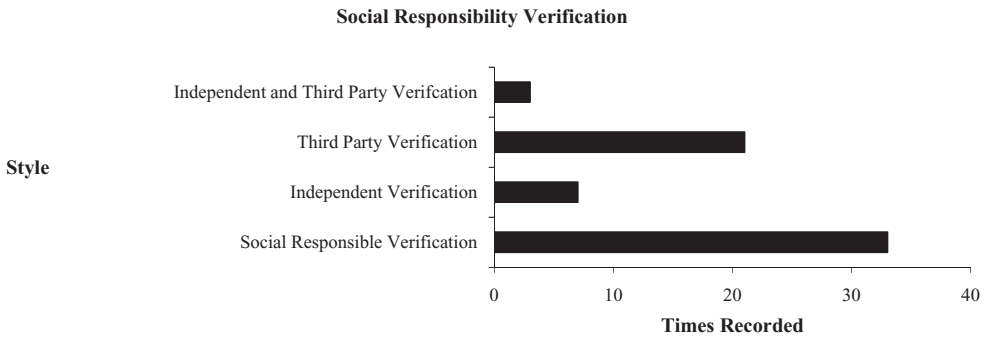


Figure 2. Social responsibility verification styles by number of times discussed in the literature

We find that there is no consensus about social responsibility verification in the literature. Some studies give considerable effort to justify the existence of social responsibility verification while the use of a flag from a third party source suffices in others. Some studies do not appear to recognize this issue at all. Yet, studies that are more recent give considerable weight to this matter in their data discovery, utilizing both independent investigation and third party institutions to verify the integrity with respect to social responsibility of the data (Renneboog *et al.*, 2008).

Mutual funds without socially responsible components are conventional mutual funds, but it is difficult to discern the difference with SRI funds without a qualifying label. Furthermore, it is difficult to trust a label without a guarantee. Consequently, over the past 20 years, there have been significant developments in ethical investment research. A large part of this research is about certifying that SRI funds invest in socially responsible companies. Some research suggests that SRI funds are not as different from conventional funds as investors may have assumed (Benson *et al.*, 2006; Bauer *et al.*, 2007; Kempf *et al.*, 2007). Furthermore, Kreander (2001) puts forward that there are bate SRI funds for attracting new customers. He argues that SRI funds are 'genuine' when there is an in-house research authority associated with the fund (Kreander, 2001). Renneboog *et al.* (2008) find that this can result in increased expenses. But Gil-Bazo *et al.* (2010) do not detect differences in the research expenses between in-house and external information provision.

Furthermore, there is confusion on whether the various rating agencies agree what actually is socially responsible investing. As an example, we refer to the debate between funds, NGOs, rating agencies and investment or fund analysts in the US and Europe (see Louche and Lydenberg, 2006). In addition, there is no overarching SRI governing board to discuss these principles. Illustratively, Scholtens (2005, p. 67) writes in reference to SRI indices that 'A problem is that institutions that constitute these indices may have very different views about what actually is ethically or socially responsible behavior'. Thus, it appears that there is not a standard set of guidelines either for the funds or for the verifiers.

Survivorship Bias

Overall, 20 of the 41 studies (49%) recognize the existence of survivorship bias in their research. We find four distinct ways in which the literature deals with survivorship. First, four studies (10%) regard the survivorship bias

DOI: 10.1002/sd

as insignificant and do not deal with it. Second, one study (2%) discerns the bias from independent SRI knowledge and experience. Third, 15 studies (37%) confirm that there is a bias based on the database. Fourth, 21 studies (51%) do not treat it at all.

We find that there is neither universal survivorship bias recognition nor treatment of this bias in the SRI fund performance literature. However, recent studies are more likely to consider survivorship bias or to recognize their limitations in not doing so. For example, the study by Bauer *et al*. (2005) comprehensively deals with the survivorship bias. However, in their 2007 study on Canadian SRI funds, they are limited in doing so, due to data restrictions (Bauer *et al*., 2007). The topic of survivorship bias is worthy of vigilance. This is mainly because not all data sources incorporate 'dead funds' into their data archives and because survivorship bias is not yet universally recognized around the globe. Thus, with the development of SRI funds, exchanges and databases have to keep systemic records of fund returns, even after their failure, to be able to eliminate errors in the estimation of returns.

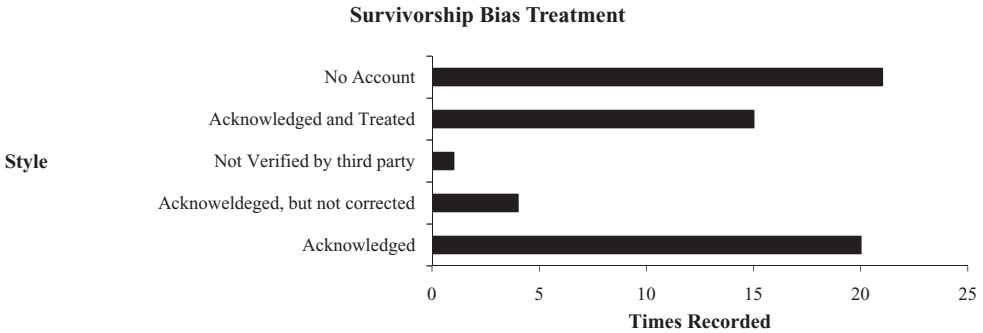


Figure 3. Survivorship bias treatment style by number of times discussed in the literature

Benchmarks

Grinblatt and Titman (1994) point out that the choice of the benchmark can have a substantial impact on conclusions about investment portfolio performance. In SRI fund performance analysis, researchers appear to use three categories of benchmarks to measure against the performance of SRI funds, namely conventional indices, matched pair analysis and sustainability indices.[6] The 41 studies commonly have conventional indices, both major global and regional, prior to the creation of the first sustainability indices. Fifteen studies (37%) use major indices, and six (15%) used regional indices. Another 15 studies (37%) use matched pair analysis between SRI and conventional funds of similar composition. Seven studies (17%) use major sustainability indices, and five (12%) incorporate regional sustainability indices. For a complete list of indices used in the studies, we refer to Table 1.

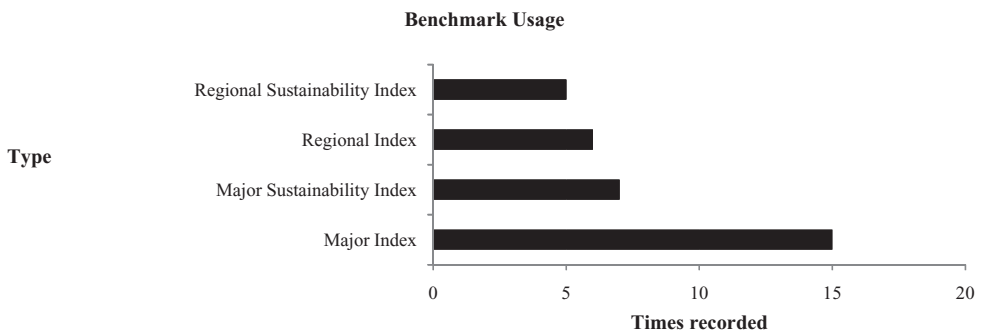


Figure 4. Benchmark usage type by number of times discussed in the literature

[6] Matched pair analysis in the context of SRI fund evaluation is the matching of SRI funds with conventional funds commonly of similar company size, age, fund size, region, industry or fee composition.

DOI: 10.1002/sd

Thus, there is broad usage of benchmarks, both conventional and SRI. In addition, considerable use is made of matched pair analysis. Bauer *et al.* (2006) argue that the construction of ethical investments using social, environmental and ethical factors screens may preclude them from the adequate assessment by broad market indices. Consequently, more studies use multiple benchmarks, conventional, matched pairs and SRI, to put fund performance into perspective. Luther *et al.* (1992) and Luther and Matatko (1994) deem conventional indices unable to meet the needs of SRI as they comprise socially irresponsible companies as well. When SRI benchmarks are nonexistent, they regard matched pair analysis as a solution. Thus, matched pairs were the main benchmark in the early literature and they are still widely used for comparisons today. The primary advantage of using matched pairs is that the researcher can decide the match based on a series of pre-determined properties, such as age, size, diversification and capitalization (see, e.g., Luther and Matatko, 1994; Bauer *et al.*, 2005; Schröder, 2004). However, there are caveats regarding SRI funds that may not make them a suitable match against conventional funds, especially in the case of cross-country studies. For example, matching US or British conventional funds against various pools of SRI funds in Europe may not prove fair, as the specific SRI strategies have shown themselves to be culturally motivated (Schröder, 2004; Louche and Lydenberg, 2006). This may distort the comparison of financial returns and risks.

Developments within the product offerings of the SRI domain resulted in new metrics to test SRI funds. For example, it was questioned whether conventional benchmarks, either matched pairs or published indices, were suitable for SRI funds as they did not incorporate the same scrutiny in their equity selection process as an SRI fund did (Bauer *et al.*, 2006). SRI benchmark indices started small, but then developed global indices and further still generated individual country indices and were incorporated into the analysis.[7] However, even here, concerns arose as to which SRI benchmarks or other specialized benchmarks were required for an unbiased analysis (Plantinga and Scholtens, 2002; Schröder, 2004). Furthermore, some evidence suggests that standard equity indexes are better capable of explaining SRI fund performance than an SRI index is (Bauer *et al.*, 2007, 2005).

Major indices	*Major sustainability indices*
AEX	Dow Jones Sustainability Index STOXX
Dow Jones World	FTSE4Good Global
Dow Jones World Tech/Energy	Dow Jones Sustainability Index World
DJ STOXX	Ethical Investment Research Service
Financial Times All Share Actuaries Index	FTSE4Good Global
Financial Times World Index	ImpaxET50
Hoare Govett Smaller Companies Index	*Regional indices*
Morgan Stanley Capital Int. Perspective World Index	All Ordinaries Accumulation Index
MSCI AC Europe	Australia Index
MSCI AC World	*Regional sustainability indices*
MSCI EMU	DJSG Europe, America
MSCI European Capital Markets Index	Domini 400 Social Index
MSCI Indices	FTSE4Good Europe
MSCI Pacific, Europe, North America	Jantzi Social Index
MSCIIWI	Westpac Monash Eco Index
S&P 500	
Wilshire 5000 Equity Index	
Worldscope	

Table 1. Indices in SRI fund performance studies

[7] The indices were developed as a product to serve customers who wanted a passive investment strategy (Geczy *et al.*, 2005).

DOI: 10.1002/sd

Sensitivity and Robustness Analysis

Sensitivity and robustness analysis help to assess the soundness of the estimates reported. Examples are the impact of fund style and composition, the impact of management skills and SRI strategies. Eight studies (20%) assess fund composition through growth versus value investment styles. Six studies (15%) go into asset class diversification, 15 studies (37%) investigate asset size, nine studies (22%) asset age, 18 studies (44%) capitalization of underlying assets, five studies (12%) assess sector composition and ten studies (24%) investigate international versus domestic diversification. Other sensitivity checks discern the influence of management skill in procuring returns.

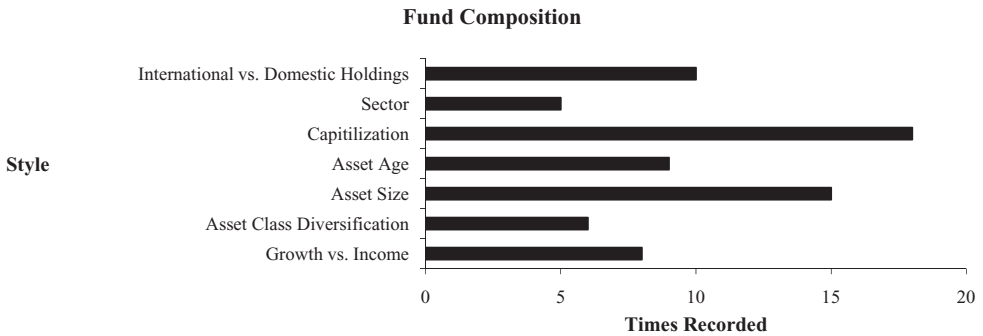


Figure 5. Fund composition evaluation style by number of times discussed in the literature

Primarily, the focus is on the skill of the manager in acquiring returns. The literature reports controls for market timing ability (in six studies or 15%) and manager skill level, i.e. evolutionary learning effects or management changes (five studies or 12%). In addition, we assess different SRI strategies, predominantly screening, monitoring and engagement (Scholtens, 2006). These three strategies discern fund performance based on screens, i.e. style of screen (e.g. positive, negative and best in class), type of screen (e.g. corporate governance, environment and social) and the number of screens that may influence fund performance. Twenty-three studies (56%) test screening strategies and their influence on performance. Two studies (5%) investigate and test the extent of monitoring and engagement using in-house research providers. Twelve studies (29%) estimating multiple models. Last, there are 18 studies (44%) that test against multiple benchmarks to discern counterevidence or further support or rejection of the hypothesis.

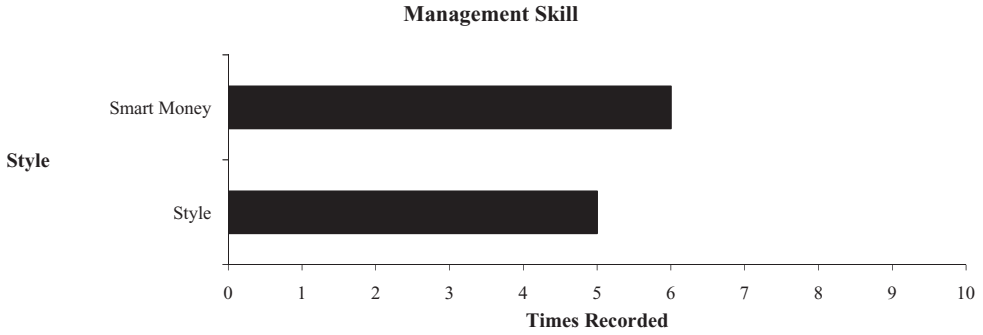


Figure 6. Management skill evaluation style by number of times discussed in the literature

Sensitivity and robustness analysis are important when discerning the funds' composition, influence of management and extent of SRI strategy incorporation to arrive at the correct specification of the model. Our study finds three areas where sensitivity and robustness checks are used to understand fund composition, i.e. asset class diversification, capitalization, and value and growth attributes. Asset class diversification is based on the composi-

DOI: 10.1002/sd

tion of the fund, via equity, cash or fixed income securities (for example Plantinga and Scholtens, 2002; Bauer *et al.*, 2006). Asmundson and Foerster (2001) suggest that the extent of cash or fixed income investment actually influences the returns on SRI portfolios. Likewise, capitalization is sometimes controlled for with an index or through multifactor models. Luther and Matatko (1994) use a small cap benchmark index to control for the small company effect on returns. Schröder (2004) suggests that using a small cap index is not appropriate, but that instead the Fama-French multifactor model is to be preferred.[8]

To evaluate the influence of management skill, market timing ability is the main determinant that influences fund performance (Bollen and Busse, 2001). Kreander *et al.* (2005) discern that it is not the stock selecting ability of managers that is problematic, but their market timing ability. Managers in both SRI and non-SRI funds are unable to sell high and buy low, thus diminishing their portfolio returns. Renneboog *et al.* (2007) and Bauer *et al.* (2007) also found this result. Thus, an adequate interpretation of fund performance style requires an assessment of the managers' market timing ability. In this respect too, SRI fund managers do not seem to deviate from conventional fund managers.

The role of SRI strategies is at the heart of the SRI debate, as the number of screens, style and type influence the returns of SRI funds. There is mixed evidence on the number of screens employed; some support a linear positive relationship (Renneboog *et al.*, 2007), where others see a curvilinear relationship with a maximum number of screens before losses occur (Barnett and Salomon, 2006). Evidence suggests that negative screening leads to exclusion and potentially smaller profits (Lozano *et al.*, 2006; Barnett and Salomon, 2006), whereas positive screens and best in class approaches may result in increased returns (Goldreyer and Diltz, 1999; Derwall *et al.*, 2005). Renneboog *et al.* (2008) observe that decreased returns result from corporate governance and social screen use. However, Derwall *et al.* (2005) do not arrive at this conclusion. Accordingly, we infer that screening may influence returns.

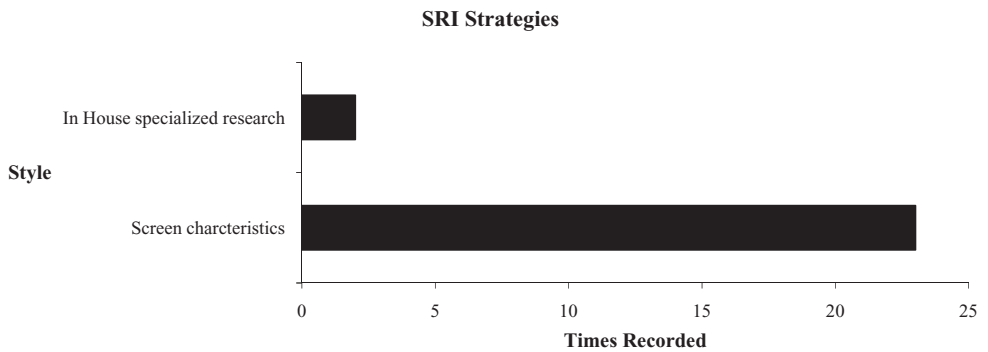


Figure 7. Social responsibility investment strategies by number of times discussed in the literature

Various model specifications can discern how performance and risk measures adjust (see also Renneboog *et al.*, 2007). Cortez *et al.* (2008) show that performance changes from specification to specification. This is sometimes contingent upon a static (e.g. Fama and French or Carhart multifactor) or a dynamic (conditional strategy model) specification. Cortez *et al.* (2008) also establish that there is a performance increase when there is a conditional strategy specification relative to a static multifactor model (see also Gregory and Whittaker, 2007; Bauer *et al.*, 2007; Renneboog *et al.*, 2008).

[8]The multifactor model by Fama and French controls for two additional style factors beyond market risk: (1) the risk premium associated with small or large capitalization companies; (2) the risk associated with value or growth weighted companies.

DOI: 10.1002/sd

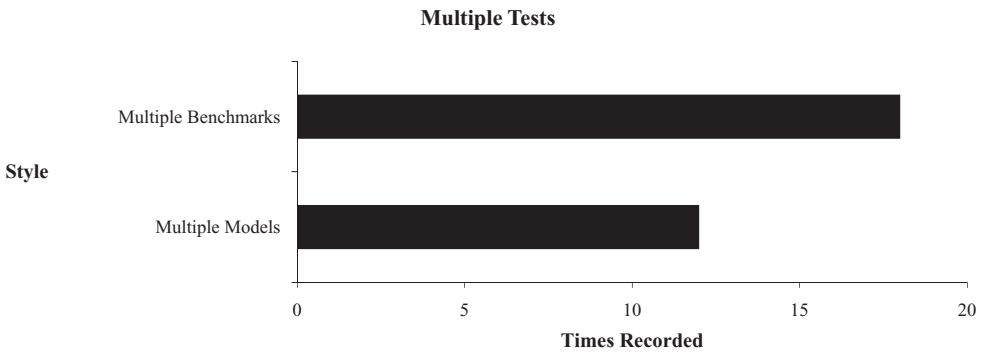


Figure 8. Robustness checks by number of times discussed in the literature

Best Practices

Based on the assessment of the SRI mutual fund performance literature, we come up with a list of best practices for performance attribution research for socially responsible investments. Table 2 lists these recommendations. These best practices are congruent with those that were derived for fund performance studies in general. However, we relate the recommendations to SRI. As such, we also are able to provide some very specific recommendations for best practices, especially in the case of benchmarking and social responsibility verification. As discussed, the SRI empirical research reports investment practices over an extended period. In some cases, they report practices that minimize measurement error or realize better-fitted specification. By synthesizing these reports, we intend to provide a systematic checklist for conducting empirical analysis with respect to SRI mutual fund performance, to consolidate the empirical issues for performance analysis, and to isolate the main arguments for proper empirical SRI performance analysis.

To wrap up, the best practices for SRI mutual fund performance analysis with respect to the five themes are the following.

1. *Data quality*. The treatment of dividend yields and fees should be included and, more in particular, studies have to reveal how the dividend yield and fees affect the income and costs of operating a SRI fund. This is the case when performance analyses show that dividends and fees are accountable for a substantial difference in the returns of SRI and conventional funds. They connect with management factors and fee systems, which in turn are related to the jurisdiction where the mutual fund is domiciled.
2. *Social responsibility verification*. We recommend resolving the trust issue and upholding a best practice of independent research of funds, considering both their prospectuses and fund managers' information, and verification through independent third party sources. Since the standards vary between continents, it may be inappropriate to apply US standards to EU funds and vice versa, and this of course holds for other cultures as well. A key issue is the definition, measurement and assessment of responsibility.
3. *Survivorship bias*. SRI fund performance evaluations need to account for dead funds (stocks) and for their impact on the results.
4. *Benchmarking*. The choice of the benchmark (or the matched pair) must be well motivated. Alternative benchmarks must show the sensitivity of the results. We suggest utilizing multiple benchmarks to evaluate the performance of SRI. Ranking an SRI fund against socially responsible and conventional indices gives more insight about the fund. With matched pairs, multiple conventional pools per country are helpful. Furthermore, matched pairs must include multiple criteria, beyond age, size and capitalization.
5. *Sensitivity and robustness*. Numerous sensitivity and robustness tests can be undertaken. However, four important considerations are important to arrive at best practices. The first is to consider the fund's composition. The second is to consider the impact of fund management. The third is to consider the role of specific SRI strategies used by the fund. The last is to engage in alternative model specifications.

DOI: 10.1002/sd

Data quality	
1.	Explain the returns on each portfolio, with specific attention to dividend yields, cash payments and the reinvestment of these returns. Control for the dividend yield (and stock splits).
2.	Explain the transaction costs on each portfolio, with specific attention to specific components such as management fees, load fees and other transaction costs charged by the funds.
Social responsibility verification	
3.	Clarify how the social responsibility of the fund was established and how responsibility information translates into actions by the fund.
Survivorship bias	
4.	Incorporate dead funds into the analysis or explain how refraining from dead funds and stocks might influence the results.
Benchmarking	
5.	Test against several benchmarks (conventional and social responsibility benchmarks) and motivate benchmark choices.
6.	Utilizing a match pair analysis with SRI funds requires the consideration of conventional funds that are of comparable age, size, sector, country/culture, asset diversification.
Sensitivity & robustness	
7.	Show how changes in fund composition (asset class diversification, capitalization, value or growth diversification, age, size and international vs domestic diversification) impact the results.
8.	Examine potential management influences (market timing ability, evolutionary learning effect, expenses etc.).
9.	Test the influence of different social responsibility strategies, for example screening characteristics or the existence of in-house vis-à-vis outsourced research and its effect on performance.
10.	Examine the result of different models and model specifications to confirm the robustness of results.

Table 2. Recommendations for best practices in SRI mutual fund performance analysis

Conclusion

The purpose of this paper is to point out how current research of socially responsible investment (SRI) mutual fund performance is conducted and what can be regarded as best practices. To this extent, we use content analysis and meta-ethnographic analysis on 41 studies. These studies encompass a study period of 45 years, research funds in 21 countries, and the use of more than 20 different data sources. We consider five different research themes, based on their relationship with eliminating bias and estimation error in the performance estimates, namely data quality, social responsibility verification, survivorship bias, benchmarking, and sensitivity and robustness checks. We find that within these themes several issues warrant attention and require proper treatment in order to arrive at proper and sound analysis. Our recommendations for best practices are in line with those from the general performance measurement literature. However, especially in the case of benchmarking and social responsibility verification, we address issues that are very specific for performance measurement with SRI mutual funds.

References

Asmundson P, Foerster SR. 2001. Socially responsible investing: better for your soul or your bottom line? *Canadian Investment Review* **14: 62–66.

Barnett ME, Salomon R. 2006. Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal* **27: 1101–1122.

DOI: 10.1002/sd

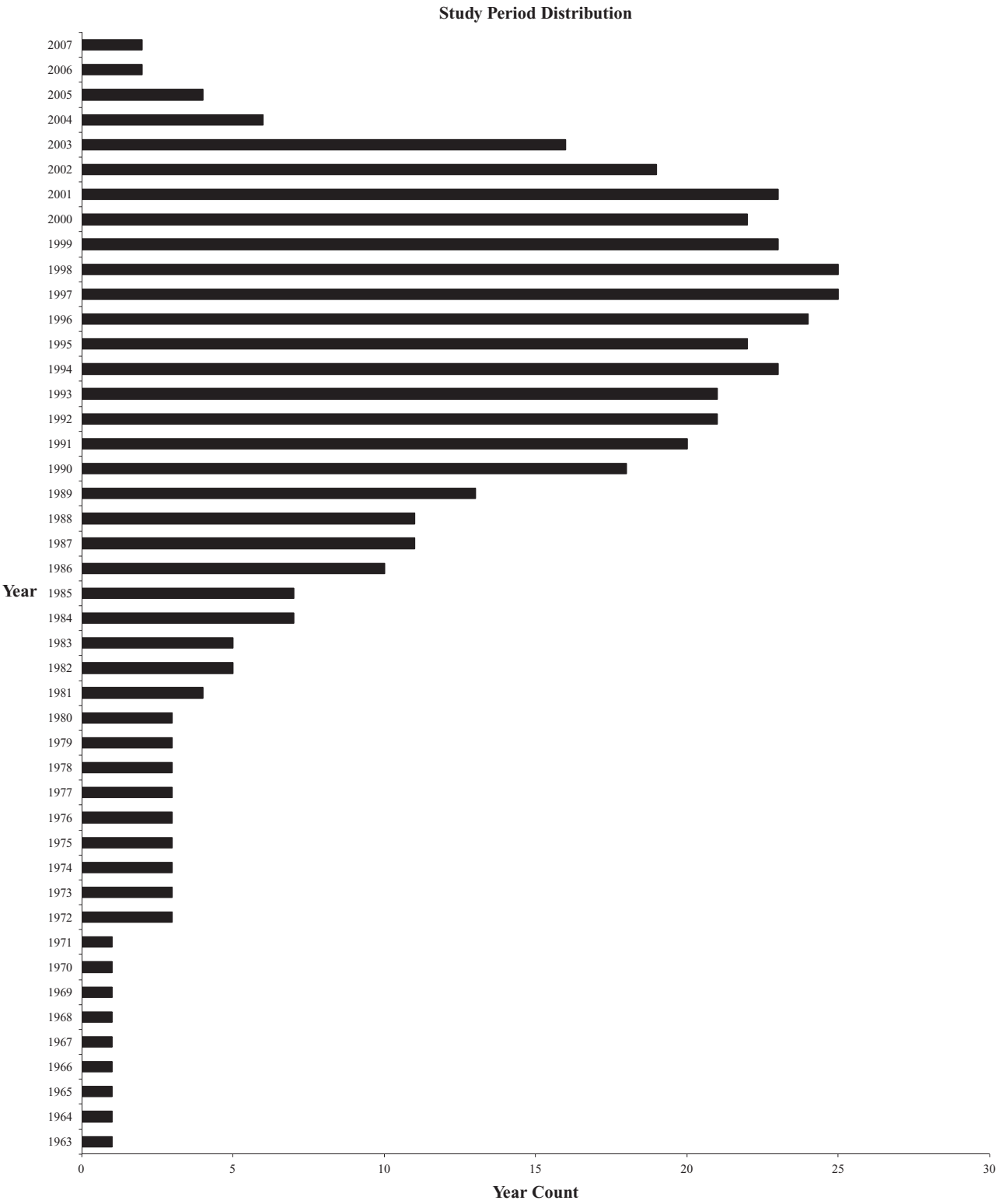
Basso A, Funari S. 2003. Measuring the performance of ethical mutual funds: a DEA approach. *Journal of the Operational Research Society* **54: 521–531.
Bauer R, Derwall J, Otten R. 2007. The ethical mutual fund performance debate: new evidence from Canada. *Journal of Business Ethics* **70: 111–124.
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.
Bauer R, Otten R, Rad AT. 2006. Ethical investing in Australia: is there a financial penalty? *Pacific-Basin Finance Journal* **14: 33–48.
Bello ZY. 2005. Socially responsible investing and portfolio diversification. *The Journal of Financial Research* **28: 41–57.
Benson KL, Brailsford TJ, Humphrey JE. 2006. Do socially responsible fund managers really invest differently? *Journal of Business Ethics* **65**: 337–357.
Bollen NPB, Busse JA. 2001. On the timing ability of mutual fund managers. *Journal of Finance* **56**: 1075–1094.
Brown S, Goetzmann W, Ibbotson R. 1992. Survivorship bias in performance studies. *Review of Financial Studies* **5**: 553–580.
Cerin P, Scholtens B. 2011. Linking responsible investments to societal influence: motives, assessments and risks. *Sustainable Development* in press.
Cortez MC, Silva F, Areal N. 2008. The performance of European socially responsible funds. *Journal of Business Ethics* **87: 1–16.
Cummings LS. 2000. The financial performance of ethical investment trusts: an Australian perspective. *Journal of Business Ethics* **25: 79–92.
Derwall J, Guenster N, Bauer R, Koedijk K. 2005. The eco-efficiency premium puzzle. *Financial Analysts Journal* **61: 51–63.
**DiBartolomeo D, Kurtz L. 1999. *Managing Risk Exposures of Socially Screened Portfolios*. Northfield Information Services: Boston, MA. http://www.northinfo.com/documents/63.pdf [Accessed 5 July 2009].
Diltz JD. 1995. The private cost of socially responsible investing. *Applied Financial Economics* **5: 69–77.
Elton EJ, Gruber MJ, Blake CR. 1996. Survivorship bias and mutual fund performance. *Review of Financial Studies* **8**: 1097–1120.
Elton EJ, Gruber MJ, Brown SJ, Goetzmann WJ. 2006. *Modern Portfolio Theory and Investment Analysis*, 7th edn. Wiley: United States of America.
Galema R, Plantinga A, Scholtens B. 2008. The stocks at stake: return and risk in socially responsible investment. *Journal of Banking and Finance* **32**: 2446–2454.
**Garz H, Volk C, Gilles M. 2002. *More Gain Than Pain. SRI Sustainability Pays Off*. WestLB Panmure: Cologne.
** Geczy CC, Stambaugh RF, Levin D. 2005. Investing in socially responsible mutual funds, working paper. Wharton School.
Gil-Bazo J, Portela A, Ruiz-Verdu P. 2010. The performance of socially responsible mutual funds: the role of fees and management companies. *Journal of Business Ethics* **94: 243–263.
**Gluck K, Becker Y. *The Impact of Eco-Efficiency Alphas on an Actively Managed US Equity Portfolio Performance*. State Street Global Advisors: Boston, MA.
Goldreyer EF, Diltz JD. 1999. The performance of socially responsible mutual funds: Incorporating sociopolitical information in portfolio selection. *Managerial Finance* **25: 23–36.
Greene WH. 2008. *Econometric Analysis*, 6th edn. Prentice-Hall: Englewood Cliffs, NJ.
Gregory A, Whittaker JM. 2007. Performance and performance persistence of ethical unit trusts in the UK. *Journal of Business Finance and Accounting* **34: 1327–1344.
Grinblatt M, Titman S. 1994. A study of monthly mutual fund returns and performance evaluation techniques. *Journal of Financial and Quantitative Analysis* September: 419–444.
Guerard JB Jr. 1997. Is there a cost to being socially responsible in investing? *The Journal of Investing* **6: 11–18.
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.
**Havemann R, Webster P. 1999. *Does Ethical Investment Pay*, EIRIS Report.
Hedesström M, Lundqvist U, Biel A. 2011. Investigating consistency of judgement across sustainability analyst organisations. *Sustainable Development* in press.
Hoepner AGF, McMillan DG. 2008. The 52 most influential studies on 'Responsible Investment': an Influential Literature Analysis (ILA) comprising a rating, characterisation, categorisation & review, working paper. St. Andrews.
Jansson M, Biel A. 2011. Motives to engage in sustainable investment: a comparison between institutional and private investors. *Sustainable Development* in press.
Kennedy P. 2008. *A Guide to Econometrics*, 6th edn. Wiley-Blackwell: India.
Kempf A, Osthoff PC, Platz AM. 2007. The effect of socially responsible investing on portfolio performance. *European Financial Management* **13: 908–922.
Kothari CR. 2004. *Research Methodology. Methods & Techniques*, 4th edn. New Age: India.
Kreander N, Association of Chartered Certified Accountants (Great Britain). 2001. *An Analysis of European Ethical Funds* Certified Accountants Educational Trust: London.
Kreander N, Gray R, Power D, Sinclair C. 2002. The financial performance of European ethical funds 1996–1998. *Journal of Accounting and Finance* **1: 3–22.
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 and Accounting* **32: 1465–1493.
Louche C, Lydenberg S. 2006. Socially responsible investment: differences between Europe and United States, *Vlerick Leuven Gent Management School Working Papers*.

DOI: 10.1002/sd

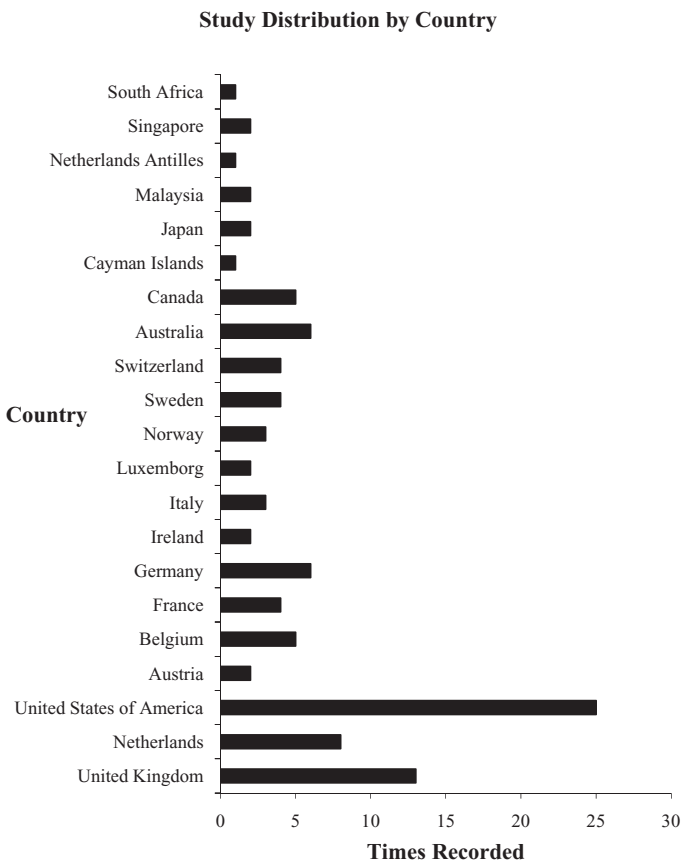
Lozano JM, Albareda L, Balaguer MR. 2006. Socially responsible investment in the Spanish financial market. *Journal of Business Ethics* **69**: 305–316.
Luther R, Matatko J. 1994. The performance of ethical unit trusts: choosing an appropriate benchmark. *British Accounting Review* **26: 77–77.
Luther RG, Matatko J, Corner D. 1992. The investment performance of UK ethical unit trusts. *Accounting, Auditing and Accountability Journal* **5: 57–70.
Mallin C, Saadouni B, Briston R. 1995. The financial performance of ethical investment funds. *Journal of Business Finance and Accounting* **22: 483–496.
Manescu C. 2011. Stock returns in relation to environmental, social, and governance performance: mispricing or compensation for risk? *Sustainable Development* in press.
Margolis JD, Walsh JP. 2001. *People and Profits?: the Search for a Link Between a Company's Social and Financial Performance*. Erlbaum: Mahwah, New Jersey.
Margolis JD, Walsh JP. 2003. Misery loves companies: rethinking social initiatives by business. *Administrative Science Quarterly* **48**: 268–305.
Matatko GA, Luther R. 1997. Ethical unit trust financial performance: small company effects and fund size effects. *Journal of Business Finance and Accounting* **2: 705–725.
Mill GA. 2006. The financial performance of a socially responsible investment over time and a possible link with corporate social responsibility. *Journal of Business Ethics* **63: 131–148.
**M'Zali B, Turcotte MF. 1998. The financial performance of Canadian and American environmental and social mutual funds. Paper presented at *Partnership and Leadership: Building Alliances for a Sustainable Future 7th International Conference The Greening Of Industry Network*, Rome.
Noblit GW, Hare RD. 1988. *Meta-Ethnography: Synthesizing Qualitative Studies*. Sage: Newbury Park, California.
Orlitzky M, Schmidt FL, Rynes SL. 2003. Corporate social and financial performance: a meta-analysis. *Organization Studies* **24**: 403.
Plantinga A, Scholtens B. 2002. Socially responsible investing and management style of mutual funds in the Euronext stock markets. *Revue Bancaire et Financière/Bank- en Financiewezen* **66: 248–254.
**Renneboog L, Ter Horst J, Zhang C. 2006. The price of ethics: evidence from socially responsible mutual funds. *Discussion Papers* 2007-012.
Renneboog L, Ter Horst J, Zhang C. 2007. Socially responsible investments: methodology, risk exposure and performance. *Discussion Papers* 2007-013.
Renneboog L, Ter Horst J, Zhang C. 2008. 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 mutual fund. *Review of Financial Economics* **6**: 137–149.
Scholtens B. 2005. Style and performance of Dutch socially responsible investment funds. *The Journal of Investing* **14: 63–72.
Scholtens B. 2006. Finance as a driver of corporate social responsibility. *Journal of Business Ethics* **68**: 19–33.
Scholtens B. 2011. Corporate social responsibility in the international insurance industry. *Sustainable Development* in press.
Scholtens B, Dam L. 2007. Cultural values and international differences in business ethics. *Journal of Business Ethics* **75: 273–284.
Schröder M. 2004. The performance of socially responsible investments: investment funds and indices. *Financial Markets and Portfolio Management* **18: 122–142.
Statman M. 2000. Socially responsible mutual funds. *Financial Analysts Journal* **56: 30–39.
**Stone BK, Guerard JB, Gultekin MN, Adams G. 2001. *Socially Responsible Investment Screening: Strong Empirical Evidence of No Significant Cost for Actively Managed Value-Focused Portfolios*. Department of Finance, Provo, Brigham Young University.
Tippet J. 2001. Performance of Australia's ethical funds. *The Australian Economic Review* **34: 170–178.
Van de Velde E, Vermeir W, Corten F. 2005. Corporate social responsibility and financial performance. *Corporate Governance* **5: 129–138.
**Wheat D. 2002. *Performance of Socially and Environmentally Screened Mutual Funds*. SRI World Group. http://www.institutionalshareowner.com/pdf/Performance-of-Screened-Mutual-Funds.pdf [Access 10 June, 2009].

DOI: 10.1002/sd

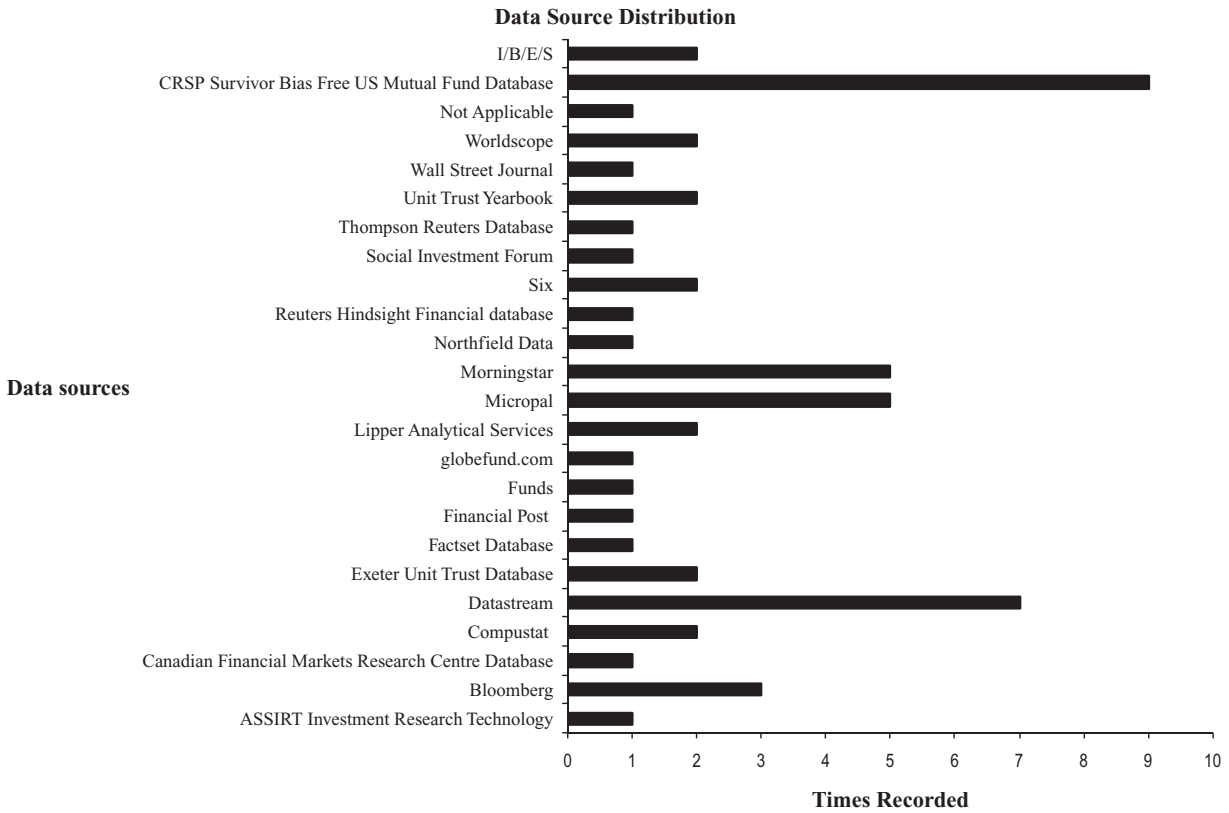
Appendix A: SRI Performance Analyses Periods Covered in The Literature



Appendix B: SRI Performance Literature Analyses by Country and Number of Times Discussed in The Literature

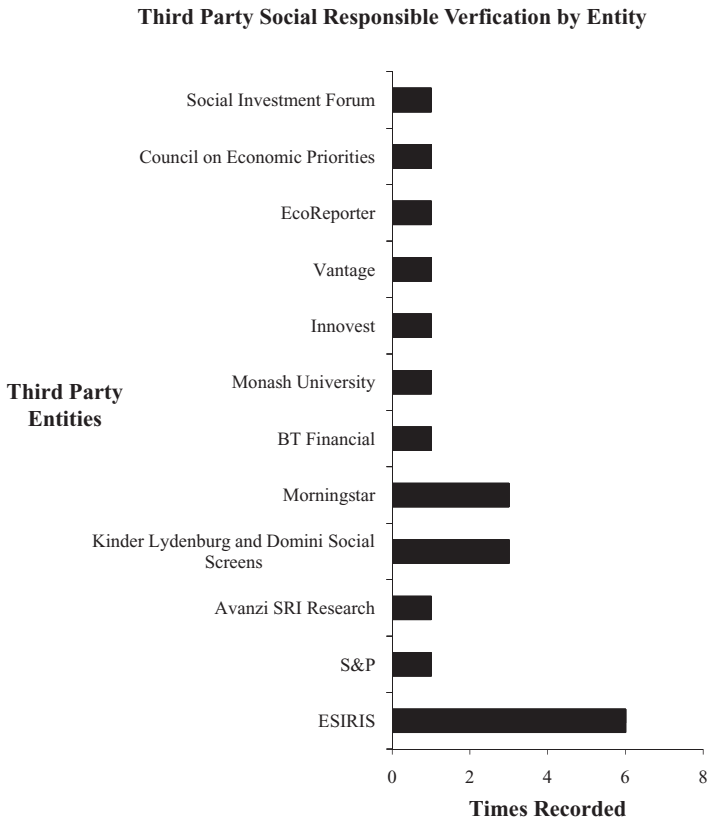


Appendix C: SRI Performance Literature Analyses by Data Sources and Number of Times Discussed in The Literature



DOI: 10.1002/sd

Appendix D: SRI Performance Literature Analyses by Third Party SRI Research Organizations and Number of Times Discussed in The Literature



DOI: 10.1002/sd