J Bus Ethics (2013) 118:349–363
DOI 10.1007/s10551-012-1584-z

The Influence of Primary Study Characteristics on the Performance Differential Between Socially Responsible and Conventional Investment Funds: A Meta-Analysis

Sebastian Rathner

Received: 14 March 2012 / Accepted: 21 November 2012 / Published online: 7 December 2012

Abstract Empirical studies, which analyze the performance of socially responsible investment (SRI) funds relative to conventional funds, find contradictory results. The aim of this paper is to investigate, with the help of a meta-analysis, how selected primary study characteristics influence the probability of a significant under- or outperformance of SRI funds compared with conventional funds. 25 studies with more than 500 observations are included in the meta-analysis. The results of this paper suggest that the consideration of the survivorship bias in a study increases (decreases) the probability of a significant outperformance (underperformance) of SRI funds relative to conventional funds. The focus on United States (US) SRI funds increases (decreases) the probability of a significant outperformance (underperformance) too. The time period influences the probability of a significant under- and outperformance of SRI funds as well, but based on the results of this paper, it is not possible to draw general conclusions on this variable.

Keywords Corporate social responsibility (CSR) · Ethical investment · Fund performance · Socially responsible investment (SRI) · Sustainability

Introduction

Socially responsible investment (SRI) is an investment process that combines an investor's financial objectives with environmental, social, or ethical considerations (Renneboog et al. 2008a; European Sustainable Investment Forum (Eurosif) 2010). Thus, SRI stock funds, for example, use financial screens as well as environmental, social, or ethical screens to select their stocks.

The SRI assets under management in Europe increased from €2.7 trillion in 2007 to €5 trillion in 2009 which is a growth of 87 % (Eurosif 2010). Eurosif divides the SRI market into two segments, a stricter "core" SRI segment (investments have to apply sophisticated SRI techniques), and a "broad" SRI segment with less strict requirements.[1] The "core" segment (€1.2 trillion) is estimated to represent 10 % of the asset management industry in Europe in 2009 (Eurosif 2010). In addition, the number of European SRI retail funds increased from 280 in 2001 to 886 in 2011, which is an increase of 216 % (Vigeo 2011). Furthermore, Eurosif (2010) reports the compound annual growth rates of SRI and conventional funds by asset class between 2007 and 2009. The assets in SRI bond funds grew by 114 % and in SRI monetary funds by 33 %, while conventional bond funds experienced only a growth of 4 % and conventional monetary funds a decrease of 5 %. Owing to the financial crisis, assets in SRI equity funds decreased by 7 %, but assets in conventional equity funds decreased by 14 %. As can be seen from these percentages, SRI funds show superior growth rates compared with their conventional counterparts in all asset classes.

The question on SRI which gained most attention from academic scholars is whether the performance of SRIs differs from the one of conventional investments. This question is addressed in most studies by investigating SRI funds and conventional funds. From a theoretic perspective, there are three different hypotheses about performance

Sections for reviewing purposes: Corporate Responsibility (quantitative issues); Finance.

S. Rathner (✉)
Department of Economics and Social Sciences, University of Salzburg, Residenzplatz 9, 5010 Salzburg, Austria
e-mail: sebastian.rathner@sbg.ac.at

[1] For more information on the definition of "broad" and "core" SRI, see Eurosif (2010), p. 9.

comparisons of SRI and conventional funds. The "underperformance-hypothesis" suggests that SRI funds generate weaker financial performance than conventional funds. The main reason for the underperformance can be seen in the fact that the implementation of SRI screens limits the full diversification potential which "may shift the mean-variance frontier towards less favorable risk-return tradeoffs than those of conventional portfolios" (Renneboog et al. 2008b, p. 304). The "outperformance-hypothesis" claims superior returns of SRI funds. An outperformance of SRI funds may occur if the SRI screening process generates value-relevant information (about the Corporate Social Performance of companies) which would not be available to fund managers otherwise. This "additional" information may help fund managers to select securities, respectively, companies with higher risk-adjusted returns (Renneboog et al. 2008b).[2] The "no-effect-hypothesis" suggests that there is no significant difference between the returns of SRI and conventional funds. This hypothesis proposes that the SRI screening process has neither a positive nor a negative influence on the financial performance (Hamilton et al. 1993; Renneboog et al. 2008b).

Narrative literature reviews of the empirical studies of this extensive body of literature corroborate the "no-effect-hypothesis" (e.g., Cortez et al. 2009) but there is some evidence for the other two hypotheses as well (e.g., Renneboog et al. 2008b; Gil-Bazo et al. 2010). The reasons for the contradictory evidence are largely unexplored. One possibility is that primary study characteristics influence the results. It is reasonable to assume that, for instance, the chosen sample period may influence the results. This could be true if, for example, the performance of the SRI fund industry during the early period of its development was worse than in later periods (possibly due to learning effects).[3]

Therefore, the aim of this paper is to investigate, with the help of a meta-regression, how selected primary study characteristics (the domicile of the investigated funds, the survivorship bias consideration in a study, the sample period) influence the probability of a significant under- or outperformance of SRI funds compared with conventional funds. These three characteristics are selected because they belong to the most important characteristics in primary studies and it is possible to derive hypotheses on them based on the literature.

The remainder of this paper is organized as follows: Sect. Literature overview and hypotheses presents a literature overview of prior studies, which compare the performance of SRI and conventional funds and develops the hypothesis of the study at hand. Section Methods and data describes the data and methods. Section Results and discussion presents and discusses the empirical results. Section Robustness conducts several "robustness checks". Section Conclusion provides a conclusion and various suggestions for future research.

Literature Overview and Hypotheses

The results of empirical studies that compare SRI and conventional fund performance are contradictory. Both, a significant out- or underperformance of SRI funds as well as no significant performance difference at all can be observed by investigating, for example, the studies mentioned in the next paragraphs. The following overview reveals several primary study characteristics (survivorship bias consideration, fund domicile, time period, performance evaluation measure, number of investigated funds, and matching procedure) that might influence the probability of a significant under- or outperformance of SRI funds compared with conventional funds. Therefore, all of them will be included in the meta-regression framework. The paper at hand focuses on the first three primary study characteristics (survivorship-bias consideration, domicile of the investigated funds, sample period) as they belong to the most important characteristics in primary studies and it is possible to derive hypotheses on them based on the literature.

An interesting characteristic, which distinguishes relevant studies, is whether a study considers survivorship bias or if it does not. Several authors show that the consideration of survivorship bias influences the average fund performance (e.g., Brown et al. 1992). A survivorship bias appears if fund samples (in a study) contain currently active funds only and do not include "dead" funds. This bias leads to an overestimation of the average fund performance because the average "dead" fund performs poorly. Hence, a systematic difference in the attrition rate between SRI and conventional funds would influence the performance comparisons in all studies which ignore the survivorship bias. Interestingly, there is some empirical evidence which suggests that the attrition rates of SRI and conventional funds are dissimilar and therefore, fund samples suffer from survivorship bias to a different degree. Gregory and Whittaker (2007) find that 29.93 % of their conventional fund sample died before the end of the sample

[2] Thus, the most pressing question is if there are any reasons why a "good" company may be a successful company as well? Heal (2008) provides an overview of theoretic reasons. This topic is investigated empirically by a vast amount of studies. For example, the often-cited meta-analysis of Orlitzky et al. (2003) finds a positive relationship between Corporate Social Performance and Corporate Financial Performance. Furthermore, a recent literature review was conducted by Van Beurden and Goessling (2008).

[3] An in-depth discussion of the potential influence of selected primary study characteristics on the observed results is provided in Sect. Literature overview and hypotheses.

period. In contrast, only 12.5 % of the SRI fund sample did so. Similarly, Kempf and Osthoff (2008) report an attrition rate of 36 % for conventional and 17 % for SRI funds. Accordingly, Renneboog et al. (2008b) discover a lower attrition rate for SRI than for conventional funds. If a study does not consider survivorship bias and the attrition rate of conventional funds is higher than the attrition rate of SRI funds (and therefore, the average performance of conventional funds is biased more upwards than the average performance of SRI funds), there should be a higher (lower) probability of a significant underperformance (outperformance) of SRI funds. In contrast, a study which accounts for survivorship bias (includes dead funds in the samples) should on average have a higher (lower) probability of a significant outperformance (underperformance) of SRI funds [hypothesis 1 (H1)].

One criterion, which distinguishes funds from each other, is their domicile. Renneboog et al. (2008b) do not find any significant performance difference for funds of thirteen countries, but report that SRI funds of France, Ireland, Sweden, and Japan significantly underperformed their conventional peers by 4–7 % per annum during 1991–2003 (using a 4-factor-model).[4] This suggests that the conclusion about the performance of SRI funds relative to conventional funds may be sensitive to the domicile of the investigated funds. Most studies focus on the SRI fund industry of the US which is claimed to be the oldest and most developed SRI fund industry in the world. Louche and Lydenberg (2006) report that the "Pioneer Fund," established in 1928 in the US, was the first SRI fund. Several other authors claim that the "PAX World Fund," established in 1971 in the US, was the first "modern" SRI fund (e.g., Renneboog et al. 2008a). Owing to the age and development of the SRI fund industry, I hypothesize that studies which investigate US SRI funds only tend to have, on average, a higher (lower) probability of a significant outperformance (underperformance) of SRI funds compared with studies which focus on funds of other countries (H2).

Bauer et al. (2006) discuss possible performance differences between Australian SRI and conventional funds during 1992–2003. Domestic SRI funds underperformed their conventional peers in the first 3.5 years of the study's time period, outperformed conventional funds in the second 3.5 years, and did not show any significant performance difference in the last 3.5 years. As Bauer et al. (2006), several authors divide their sample period into subperiods to investigate the influence of study subperiods on the results (e.g. Bauer et al. 2006; Renneboog et al. 2008b; Gil-Bazo et al. 2010). The findings of these studies "suggest that different sample periods may lead to different conclusions about the performance of SRI funds relative to that of conventional funds" (Gil-Bazo et al. 2010, p. 253). In particular, several studies find a "catching up phase" of SRI funds, which means that studies with a newer sample period show better results for SRI funds (Bauer 2005; Bauer et al. 2006). The main reason may be seen in the fact that the SRI fund industry is a relatively young industry and "may have experienced a learning phase during the early period of its development" (Renneboog et al. 2008b, p. 311). In accordance with the mentioned studies, I hypothesise that studies with a(n) newer (older) sample period have, on average, a higher (lower) probability of a significant outperformance and a lower (higher) probability of a significant underperformance of SRI funds (H3).

The study of Renneboog et al. (2008b) uses the one-factor-model based on Jensen (1968) as well as Carhart's (1997) four-factor-model to evaluate fund performance. They find a significant underperformance of SRI funds of Singapore using Jensen's Alpha, but no significant performance difference when they use Carhart's Alpha. Gil-Bazo et al. (2010) find a significant outperformance of US SRI funds when they use a four-factor-model but they find no significant performance difference when they use a one-factor-model or unadjusted returns. As Renneboog et al. (2008b) and Gil-Bazo et al. (2010), many studies use several measures to evaluate fund performance and measures vary from study to study as well. Hence, it is reasonable to include the performance evaluation measures as control variables in the meta-analysis. Jensen's Alpha and Carhart's Alpha are the most prominent performance measures in primary studies, and therefore I include a dummy variable for both measures in the meta-regression. The third category contains all other performance measures.

For example, the study of Renneboog et al. (2008b) estimates conditional regression models (cf Ferson and Schadt 1996) to evaluate fund performance and the use of this procedure influences their results. Thus, I include a control variable *conditional performance evaluation* in the meta-regression framework.

The study of Bauer et al. (2007) includes only 8 SRI funds in their analysis, whereas Chang and Witte (2010), for example, include 164 SRI funds. Obviously, the number of funds varies strongly throughout the studies, respectively, the effects, and therefore a control variable which accounts for this fact will be included in the meta-analysis.

Gil-Bazo et al. (2010) and several other studies (e.g., Kreander et al. 2005; Statman 2000) apply a matching procedure to select comparable funds whose main difference is the SRI characteristic. A fundamental distinction is if only one (1 vs. 1 matching) or several (1 vs. many) conventional funds are matched to a SRI fund. In addition to the results

[4] Renneboog et al. (2008b) do not find significant performance differences for the following countries: Belgium, Germany, Italy, Luxembourg, Netherlands, Norway, Switzerland, UK, US, Canada, Australia, Malaysia, and Singapore.

which are based on a matching procedure (1 vs. 1 and 1 vs. 4 matching), the study of Gil-Bazo et al. (2010) presents results which are not based on a matching procedure. In several cases, the results are diverse. For example, they do not find a significant performance difference when they compare the gross one-factor Alphas of SRI and conventional funds. They do find a significant outperformance of SRI funds when they use a 1 versus 1 matching procedure, but there is no significant performance difference when they use 1 versus 4 matching. Hence, two control variables which account for the use of a matching procedure (1 vs. 1 or 1 vs. many matching) in a study should be integrated in the meta-analysis.[5] In addition, the number of matching criteria varies throughout the various studies, and is therefore integrated in the meta-regression framework as well.

Methods and Data

Variable Description and Empirical Specification of the Meta-Analysis

Primary studies use different measures to compare the performance of SRI funds and conventional funds, and hence it is difficult to compare them directly. Thus, I create the two dummy variables *Outperformance* and *Underperformance*:

Outperformance = 1 if the SRI funds in a study significantly outperform conventional funds; *Outperformance* = 0 in all other cases

Underperformance = 1 if the SRI funds in a study significantly underperform conventional funds; *Underperformance* = 0 in all other cases

These two variables are used as dependent variables in the first approach, which estimates binary logit models.[6]

The second approach, which may be seen as a "robustness check," uses the categorical variable *Performance comparison* (dependent variable). It takes value 0 if the SRI funds significantly underperform the conventional funds. Value 1 is taken if there is no significant performance difference and value 2 if the SRI funds outperform their conventional peers significantly. The second approach estimates multinominal logit models.

The effects are coded as significant in the meta-analysis if they are significant at least at the 10 % level in the primary studies.

The independent variables are the three previously discussed primary study characteristics and the additional control variables as shown in Table 1. As discussed in Sect. Literature overview and hypotheses, these characteristics emerged from the literature and may moderate the significance of the results reported in primary studies. As some authors contribute more than one study to the meta-analysis which may bias the results if the studies from the same author are correlated, I include the additional independent variable *Highest number of authorships* in the meta-regression framework. This variable indicates the highest number of studies one of the authors of a certain study has contributed to. For example, the study of Bauer et al. (2007) has the value 3 because Bauer (Otten as well) contributed to 3 included studies. If the results of the studies of an author are correlated and tend to find, for example, an outperformance of SRI funds, the variable controls for this correlation.[7]

Study Selection Process and Descriptive Statistics

The starting points for this research were several narrative literature reviews (Chegut et al. 2011; Capelle-Blancard and Monjon 2010; Hoepner and McMillan 2009; Renneboog et al. 2008a). In addition, a computer search in "ScienceDirect" and "google scholar" using the keywords "SRI" and "performance" was conducted and the references of included studies were explored. For being included in the meta-analysis, a study had to meet the following criteria: First, the study investigated the performance of "real" SRI funds relative to conventional funds quantitatively. A study which focused on SRI funds only or SRI indices was not included. Second, a study needed to provide information on the significance of the observed effects.

A limitation of this study is that it is not possible to guarantee that all relevant studies were found during the searching process, as there is an enormous amount of journals and other web-sources where studies may be published. Nonetheless, from my point of view, the selected studies are representative for this body of literature.

Another limitation of this study is that it may suffer from publication bias, which suggests that journals tend to publish studies with significant results rather than publishing studies with insignificant results. Only four included studies are unpublished papers (two master theses and two working papers).[8]

[5] The base group of these two dummy variables in the regression framework is "no matching procedure."

[6] A similar approach was used by García-Quevedo (2004).

[7] A better approach would be to include a dummy variable for every single study ("study" fixed effects). Unfortunately, this is not possible in my specification as lots of studies do not report an out- and an underperformance of SRI funds. Hence, it would not be possible to estimate the underperformance and the outperformance logit model (study dummy variables would predict the dependent variables perfectly).

[8] I suppose the influence of the publication bias on this body of literature to be smaller than on other subjects because lots of studies with insignificant results were published. Table 2 reports that almost 75 % of the primary studies' results are insignificant. Nevertheless, a publication bias may be present. Unfortunately, it is not possible to control for publication bias with a dummy variable in my specification as there is no unpublished effect which reports an outperformance of SRI funds. As a conclusion, a dummy variable *Published paper* would predict the dependent variable (*Outperformance*) perfectly.

Table 1 Independent variables

Survivorship bias consideration	=	1 if a study considers survivorship bias
US funds	=	1 if a study investigates US SRI funds only
Time period 1981–1990	=	1 if the biggest part of a study's sample period is between 1981–1990
Time period 1991–2000	=	1 if the biggest part of a study's sample period is between 1991–2000
Time period 2001–2008	=	1 if the biggest part of a study's sample period is between 2001–2008
Performance evaluation Jensen's Alpha	=	1 if a study uses a one-factor regression model to evaluate fund performance (Jensen's Alpha)
Performance evaluation Carhart's Alpha	=	1 if a study uses a multi-factor regression model to evaluate fund performance (e.g. Carhart's four factor Alpha)
Other performance evaluation	=	1 if a study uses a fund performance evaluation measure, which cannot be assigned to the other two groups
Conditional performance evaluation	=	1 if a study uses a conditional regression approach to evaluate fund performance (e.g. Ferson and Schadt 1996)
1 vs. 1 Matching	=	1 if a study uses a matching procedure to match one conventional fund to one SRI funds (based on e.g. fund size and age)
1 vs. Many matching	=	1 if a study uses a matching procedure to match a certain number (higher than 1) of conventional funds to one SRI fund (based on e.g. fund size and age)
Number of matching criteria	=	Number of matching criteria used in the 1 vs. 1 or the 1 vs. many matching procedure (as stated by the authors of the studies)
Highest number of authorships	=	Some authors contribute to more than one study. This variable indicates the highest number of studies one of the authors has contributed to
Number of SRI funds	=	Number of studied SRI funds
Number of conventional funds	=	Number of studied conventional funds

25 studies with 517 effects (= comparisons between SRI and conventional fund performance in primary studies) are included in the meta-analysis. Single studies contain several performance comparisons between SRI and conventional funds; e.g., for funds of different countries. Basic information on the included studies and their results can be found in Table 2.[9] Detailed information on the included studies can be found in Appendix Tables 9 and 10.

Table 2 shows the distribution of the primary studies' results. 372 effects (almost 72 %) do not show any significant performance difference between SRI and conventional funds. A significant under- and outperformance of SRI funds is found by approximately 15 and 13 % of the effects, respectively. The descriptive results of Table 2 must be treated with caution and should not be interpreted as a "vote-counting" approach which could often be misleading. "Vote-counting" approaches count the number of significant and insignificant results in primary studies and pick the category with the largest number of "votes" as winner. The problem is that these approaches treat nonsignificant results of studies as evidence that a "true" effect is absent and ignore the possibility that the nonsignificant results occur because of low statistical power (Borenstein et al. 2009; Hunter and Schmidt 1990).

Panel A of Table 3 provides information on the sample periods of the effects of primary studies.[10] I create three dummy variables which divide the sample period throughout all 25 primary studies, lasting from 1981 to 2008, into the following three subperiods (almost decades) 1981–1990, 1991–2000, and 2001–2008.[11] A dummy variable takes value 1 if the biggest part of the sample period of an effect is in this subperiod (cf Table 1). The first period reflects the beginning of the SRI movement. Eleven effects investigate funds in this period. The small number seems reasonable because in this early period only some SRI funds existed. All over the world the SRI fund industry started to expand in the early 1990s (Renneboog et al. 2008a). Since the early 2000s, the growth of the SRI industry has accelerated as large institutional investors, in particular pension funds, increasingly entered the market. The adoption of SRI techniques by large institutional investors is regarded as a kind of "mainstreaming" of SRI as well as an important step in the maturity of SRI (Sparkes and Cotwon 2004; Bengtsson 2008). As a result, most

[9] All included studies are marked in the reference list with an asterisk.

[10] Unfortunately, not every study provides information on the sample period of all effects.

[11] A similar procedure to divide the sample period is used, for example, by *Bauer (2005) and Bauer et al. (2006) who divide their sample periods into three equal and non-overlapping subperiods.

Table 2 Information on the included studies

Authors	Publication year	Significant underperformance of SRI funds	No significant performance difference	Significant outperformance of SRI funds	Total
Bauer et al.	2007	0	6	0	6
Bauer et al.	2005	4	22	4	30
Bauer et al.	2006	1	8	2	11
Bello	2005	0	6	1	7
Benson et al.	2006	10	32	0	42
Bollen	2007	2	8	5	15
Chang and Witte	2010	10	20	4	34
Derwall and Koedijk	2009	0	23	9	32
Gil-Bazo et al.	2010	6	52	39	97
Goldreyer et al.	1999	3	9	0	12
Gregory et al.	1997	1	5	0	6
Gregory and Whitaker	2007	0	4	2	6
Hamilton et al.	1993	0	2	0	2
Humphrey and Lee	2011	0	8	0	8
Kempf and Osthoff	2008	0	2	0	2
Koellner et al.	2007	0	5	1	6
Kreander et al.	2005	0	7	0	7
Kryzanowski et al.	2011	0	36	0	36
Liedekerke et al.	2007	0	5	1	6
Mueller	1991	3	0	0	3
Renneboog et al.	2008b	25	107	0	132
Sánchez and Sotorrío	2009	6	2	0	8
Spekl	2009	5	1	0	6
Statman	2000	0	2	0	2
Stenström and Thorell	2007	1	0	0	1
Total		77	372	68	517
Total in %		14.90	71.95	13.15	

effects study SRI funds in the periods 1991–2000 and 2001–2008.

Panel B of Table 3 reports the number of effects which considers survivorship bias and the number which ignores it.[12] 76 % of the effects consider survivorship bias, while 24 % do not. This is consistent with Chegut et al. (2011) who find substantial differences between studies concerning the treatment of survivorship bias too.

Panel C of Table 3 shows how often individual countries/regions are investigated. US funds are by far studied the most. This is consistent with, for example, Cortez et al. (2009) who suggest that most studies were conducted in the US market. It is remarkable that four Anglo-Saxon countries, namely, the US, Canada, the UK, and Australia are considered most in this research although Europe has the largest share of the global SRI market today (Eurosif 2010).

Panel D to F of Table 3 show further descriptive statistics.

[12] Some studies do not provide information on the consideration of survivorship bias.

Results and Discussion

Tables 4 and 5 present the results of the logit models with *Underperformance* (of SRI funds) and *Outperformance* (of SRI funds) as dependent variables and the independent variables, respectively, as stated in Table 1. The coefficients represent average marginal effects.[13] The standard errors are clustered by study so I am adjusting for the fact that effects of the same study may be correlated.[14] In the following tables, the first models do not include the variables on the number of funds in the primary studies because their inclusion reduces the number of meta-regression

[13] Average marginal effects are calculated by computing individual marginal effects at every observation and by averaging these individual marginal effects across the sample.

[14] For instance, some studies use several models to evaluate the performance of their fund samples. The results of the models of one study may be correlated to a certain degree because all models use the identical dataset.

Table 3 Descriptive statistics

	Freq.	Percent	Cum.
A: Frequency of effects (sample period)			
1981–1990	11	2.29	2.29
1991–2000	287	59.79	62.08
2001–2008	182	37.92	100.00
Total	480	100.00	
B: Frequency of effects (consideration of survivorship bias)			
Survivorship bias considered	381	75.90	75.90
Survivorship bias not considered	121	24.10	100.00
Total	502	100.00	
C: Frequency of effects (domicile of the funds)			
Australia	27	5.22	5.22
Belgium	14	2.71	7.93
Canada	49	9.48	17.41
Europe	14	2.71	20.12
France	8	1.55	21.66
Germany	14	2.71	24.37
Germany/Austria/Switzerland	6	1.16	25.53
International	3	0.58	26.11
Ireland	8	1.55	27.66
Italy	7	1.35	29.01
Japan	8	1.55	30.56
Luxembourg	7	1.35	31.91
Malaysia	8	1.55	33.46
Netherlands	8	1.55	35.01
Norway	7	1.35	36.36
Singapore	7	1.35	37.72
Sweden	9	1.74	39.46
Switzerland	8	1.55	41.01
UK	33	6.38	47.39
UK/Sweden/Germany/Netherlands	4	0.77	48.16
US	268	51.84	100.00
Total	517	100.00	
D: Frequency of effects (performance evaluation measure)			
Performance evaluation Jensen's Alpha	72	13.93	13.93
Performance evaluation Carhart's Alpha	287	55.51	69.44
Other performance evaluation	158	30.56	100.00
Total	517	100.00	
E: Frequency of effects (conditional performance evaluation)			
No Conditional performance evaluation	455	88.01	88.01
Conditional performance evaluation	62	11.99	100.00
Total	517	100.00	
F: Frequency of effects (matching procedure)			
1 vs. 1 Matching	84	16.25	16.25
1 vs. Many matching	120	23.21	39.46
No matching procedure	313	60.54	100.00
Total	517	100.00	

observations strongly. The second models include all independent variables.

Concerning the consideration of survivorship bias, the results of Table 4 are consistent with H1. Model (1) and (2) find a (significant) lower probability of a significant underperformance of SRI funds if a study accounts for survivorship bias. The probability of a significant underperformance is on average approximately 14–25 % [model (1) and (2)] smaller if a study considers survivorship bias in comparison to not considering this bias (everything else being equal). Accordingly, Table 5 shows a (significant) higher probability of a significant outperformance of SRI funds if a study accounts for survivorship bias (model 1). Strictly explaining, based on these models, the consideration of survivorship bias influences the probability of an out- or underperformance of SRI funds. From the author's perspective the most important implication of these findings is that all future studies should give at least an explicit statement on how they deal with the survivorship bias. The best option would be to eliminate survivorship bias by means of survivorship bias free data or by adding back closed funds to the sample. Moreover, the evidence of this paper may help interpreting the results of existing studies.

The results of Table 4 [model (2)] and 5 support H2 as well. Effects, which investigate US SRI funds only, have, on average, a 15 % lower probability of a significant underperformance [model (2)] and a 16 %, respectively, 29 %, higher probability of an outperformance of SRI funds compared with effects that focus on funds of other countries. As approximately half of the primary study effects focuses on SRI funds of the US and their results appear to be sample specific, it seems to be necessary to investigate SRI funds of single non-US countries in more detail. In addition, an interesting topic for future research may be the empirical investigation of possible differences between US and non-US SRI funds.[15] Differences may exist as far as performance, screening type and intensity, fund size, fund age, etc., are concerned.

Regarding H3, mixed evidence is found. The variable *Time period 2001–2008* was chosen to be the benchmark category.[16] As can be observed from Table 4, model (1) and (2) report a higher probability of an underperformance of SRI funds if an effect has the biggest part of the sample period between 1981 and 2000 compared with effects that

[15] Louche and Lydenberg (2006) investigate this issue from a historic perspective.

[16] For the empirical estimation, the dummy variables *Time period 1981–1990* and *time period 1991–2000* are taken together because there are only eight observations in the first subperiod with information on all variables of the logit models. All of these observations have the identical outcome in the dependent variable and hence, *Time period 1981–1990* would predict the dependent variable perfectly.

Table 4 Results of the meta-regression with the dependent variable *underperformance* (logit model)

	(1)		(2)	
	Coef.	Std. Err.	Coef.	Std. Err.
Performance evaluation Jensen's Alpha	0.004	0.056	0.017	0.040
Performance evaluation Carhart's Alpha	0.009	0.049	0.023	0.038
Conditional performance evaluation	−0.065***	0.018	−0.054***	0.007
1 vs. 1 Matching	−0.042	0.091	−0.158***	0.022
1 vs. Many matching	−0.120***	0.036	−0.157***	0.039
Number of matching criteria	0.009	0.025	0.037**	0.018
Highest number of authorships	−0.023	0.047	−0.134***	0.038
Survivorship bias consideration	−0.140***	0.051	−0.249***	0.033
US funds	−0.049	0.041	−0.154***	0.037
Time period 1981–2000	0.003	0.068	0.011	0.049
Number of SRI funds			0.002***	0.001
Number of conventional funds			−0.000	0.000
Obs	477		376	
Log pseudolikelihood	−179.941		−109.568	
Pseudo R2	0.070		0.247	

This table shows the average marginal effects of the independent variables in decimal notation and standard errors (clustered by study). The dependent variable is *underperformance*, which takes the value 1 if the SRI funds in a study significantly underperform the conventional funds, *underperformance* = 0 in all other cases

* Coefficient is statistically significant at the 10 % level

** Coefficient is statistically significant at the 5 % level

*** Coefficient is statistically significant at the 1 % level

Table 5 Results of the meta-regression with the dependent variable *outperformance* (logit model)

	(1)		(2)	
	Coef.	Std. Err.	Coef.	Std. Err.
Performance evaluation Jensen's Alpha	−0.038	0.064	−0.035	0.064
Performance evaluation Carhart's Alpha	−0.021	0.045	0.009	0.091
Conditional performance evaluation	0.095	0.118	0.130	0.173
1 vs. 1 Matching	0.227	0.206	0.211	0.323
1 vs. Many matching	0.102	0.169	0.047	0.224
Number of matching criteria	−0.011	0.021	−0.003	0.035
Highest number of authorships	0.031	0.047	−0.061	0.045
Survivorship bias consideration	0.130*	0.072	0.099	0.081
US funds	0.157*	0.083	0.290*	0.155
Time period 1981–2000	−0.104***	0.027	0.004	0.105
Number of SRI funds			0.001	0.001
Number of conventional funds			−0.000	0.000
Obs	477		376	
Log pseudolikelihood	−150.663		−104.375	
Pseudo R^2	0.222		0.340	

This table shows the average marginal effects of the independent variables in decimal notation and standard errors (clustered by study). The dependent variable is *outperformance*, which takes the value 1 if the SRI funds in a study significantly outperform the conventional funds, *outperformance* = 0 in all other cases

* Coefficient is statistically significant at the 10 % level

** Coefficient is statistically significant at the 5 % level

*** Coefficient is statistically significant at the 1 % level

investigate the period 2001–2008. Although, these coefficients are not significant. Model (1) of Table 5 shows significant differences. The average probability of an outperformance of SRI funds is 10 % lower for effects that have the biggest part of their sample period in 1981–2000 compared with effects which have the biggest part of their sample period in 2001–2008. The results of Table 5 [model (1)] are consistent with H3. However, the results of Table 4 and model (2) of Table 5 are not. In order to support H3, Table 4 should show a significant higher probability of an underperformance of SRI funds for effects with an older sample period.

Additional interesting results concerning the variables *1 versus 1 matching* and *1 versus many matching* are found in the binary logit models. If an effect uses one of these matching procedures to match conventional funds to the SRI fund sample (based on criteria such as fund age or fund size), there is, on average, a 12–16 % lower probability of an underperformance of SRI funds (Table 4) compared with effects which do not use a matching procedure. Possibly, the underperformance of SRI funds in studies, which do not use a matching procedure, is not caused primarily by the SRI characteristics, but by other fund characteristics (like fund size or fund age).

Another result is that there is, on average, a significant lower probability of an underperformance of SRI funds if a conditional regression model is used to evaluate fund performance. By means of a conditional approach, it can be assumed that the risk exposure of funds may be systematically changed by fund managers according to macroeconomic conditions. The most prominent approach in SRI fund literature is the conditional performance evaluation model introduced by Ferson and Schadt (1996). It suggests the inclusion of several lagged macroeconomic variables into single- or multi-factor regression models.

Robustness

The second approach, which can be seen as "robustness check," uses the dependent variable *Performance comparison*. Value 0 is taken if the SRI funds significantly underperform the conventional funds. Value 1 is taken if there is no significant performance difference and value 2 if the SRI funds outperform their conventional peers significantly. Table 6 shows the results of the multinominal logit model for the outcomes "significant under- and outperformance of SRI funds" and "no significant performance difference." Once again, the first model does not include the variables on the number of funds in the primary studies because their inclusion reduces the number of the meta-regression observations strongly. The second model includes all independent variables. The results regarding the survivorship bias consideration (H1) and domicile of the funds (H2) are in accordance with the results of the logit models. Again, a lower probability of an underperformance and a higher probability of an outperformance of SRI funds occur if a study considers survivorship bias or focuses on US funds only. The magnitudes of all coefficients are comparable to the ones found in the binary logit models. There is mixed evidence in the binary logit models concerning H3. The "robustness check" does not reveal any clear evidence in favor of H3. The probability of an underperformance of SRI funds for effects with a sample period between 1981 and 2000 is statistically not different from effects with a sample period between 2001 and 2008 in model (1) and (2). A lower probability of an outperformance of SRI funds is found in model (1) for effects with an earlier sample period. This latter result is in accordance with the evidence of the binary logit model and H3.

A concern about the reported results may be that the findings are dominated by two studies which contribute almost 45 % of the 517 effects (Renneboog et al. 2008b; Gil-Bazo et al. 2010). To address this issue I exclude the effects of these two studies and estimate the logit models again.[17] Two independent variables (*1 vs. many matching* and *Conditional performance evaluation*) have to be removed from the model as they would predict the dependent variable perfectly. Two major insights can be gained from this analysis. First, the results about the variable *Survivorship bias consideration* remain unchanged. If an effect (of the reduced sample) accounts for survivorship bias there is on average a higher (lower) probability of an outperformance (underperformance) of SRI funds. Second, the exclusion of the effects of these two studies leads to a disappearance of the effect of the variable *US funds*. This suggests that the reported effect of the variable *US funds* is largely driven by the excluded studies (in particular by Gil-Bazo et al. 2010).[18]

Alternatively, a more sophisticated approach of dealing with the problem that most studies contribute more than one effect to the meta-analysis is to estimate weighted regression models, which is suggested, for example, by Nelson and Kennedy (2009). I follow the approach of Horvathova (2010), who weighs the observations of a certain study by 1/number of effects of this study.[19] Tables 7 and 8 show the results of the weighted logit models. The results on the variables *Survivorship bias consideration* and *Us funds* are in accordance with H1, H2 and with the results of the logit models in Tables 4 and 5. Some evidence in favor of H3 (concerning the time period) can be found in these additional results as well. There is on average a lower probability of an outperformance if an effect uses and older sample period (from 1981 to 2000) in model 2. Although, it is not possible to draw a general conclusion on this variable because there is not any higher probability of an underperformance of SRI funds if an effect has the biggest part of the sample period between 1981 and 2000. Additional results from Tables 7 and 8 show that the use of a matching procedure (1 vs. 1 and 1 vs. many) increases (decreases) the probability of an outperformance (underperformance) of SRI funds in comparison to effects which do not use a matching procedure. Therefore, future studies should conduct "robustness checks" concerning the use of a matching procedure.

Conclusion

The aim of this paper is to investigate, with the help of a meta-regression, how selected primary study characteristics

[17] It is not possible to include a dummy variable for the study of Renneboog et al. (2008b) because they do not find an outperformance of SRI funds, and hence it would not be possible to estimate the logit model (dependent variable *Outperformance*). This is the reason why I exclude the effects of the two studies.

[18] Detailed results on this robustness check are not reported in the paper, but can be obtained from the author upon request.

[19] I want to acknowledge that the weighting by 1/n is a little less precise than weighing the observations by their true independence in light of the study sample (e.g., if a study has two subsample periods of equal size and one main sample regression period, the first two observations should carry a weight of 0.25 and the latter a weight of 0.5). Based on the fact that studies use more characteristics than the sample period to create subsamples the weighting scheme would get really complex and such a weighting would provide only a marginal gain. Therefore, I decided to follow Horvathova (2010) and weigh the observations by 1/n. I want to thank the unknown 'PRI Academic Fellow' for raising the aspects mentioned in this endnote and the suggestion of the weighted regression approach in general.

Table 6 Results of the meta-regression with the dependent variable *performance comparison* (multinominal logit model)

	(1)		(2)	
	Coef.	Std. Err.	Coef.	Std. Err.
Underperformance				
Performance evaluation Jensen's Alpha	0.014	0.058	0.024	0.040
Performance evaluation Carhart's Alpha	0.019	0.049	0.029	0.038
Conditional performance evaluation	−0.069***	0.018	−0.059***	0.011
1 vs. 1 matching	−0.043	0.103	−0.161***	0.024
1 vs. many matching	−0.130***	0.032	−0.161***	0.041
Number of matching criteria	0.011	0.026	0.039**	0.018
Highest number of authorships	−0.022	0.047	−0.128***	0.036
Survivorship bias consideration	−0.130***	0.049	−0.229***	0.033
US funds	−0.039	0.040	−0.144***	0.038
Time period 1981–2000	0.006	0.068	0.006	0.048
Number of SRI funds			0.002***	0.001
Number of conventional funds			−0.000	0.000
No performance difference				
Performance evaluation Jensen's Alpha	0.025	0.080	0.016	0.064
Performance evaluation Carhart's Alpha	0.003	0.071	−0.035	0.083
Conditional performance evaluation	−0.026	0.113	−0.061	0.168
1 vs. 1 matching	−0.191	0.155	−0.068	0.300
1 vs. many matching	0.025	0.160	0.100	0.215
Number of matching criteria	0.000	0.025	−0.035	0.031
Highest number of authorships	−0.010	0.055	0.174***	0.054
Survivorship bias consideration	−0.001	0.088	0.125	0.091
US funds	−0.118	0.085	−0.134	0.126
Time period 1981–2000	0.098	0.075	−0.004	0.123
Number of SRI funds			−0.002**	0.001
Number of conventional funds			0.000	0.000
Outperformance				
Performance evaluation Jensen's Alpha	−0.039	0.064	−0.039	0.062
Performance evaluation Carhart's Alpha	−0.022	0.044	0.006	0.090
Conditional performance evaluation	0.095	0.120	0.119	0.177
1 vs. 1 matching	0.234	0.207	0.229	0.315
1 vs. many matching	0.105	0.167	0.061	0.226
Number of matching criteria	−0.012	0.021	−0.005	0.033
Highest number of authorships	0.032	0.047	−0.047	0.042
Survivorship bias consideration	0.131*	0.073	0.104	0.086
US funds	0.157*	0.083	0.279**	0.136
Time period 1981–2000	−0.104***	0.027	−0.002	0.104
Number of SRI funds			0.001	0.001
Number of conventional funds			−0.000	0.000
Obs	477		376	
Log pseudolikelihood	−322.012		−211.104	
Pseudo R2	0.144		0.285	

This table shows the average marginal effects of the independent variables in decimal notation and standard errors (clustered by study). The dependent variable *performance comparison* is used

* Coefficient is statistically significant at the 10 % level

** Coefficient is statistically significant at the 5 % level

*** Coefficient is statistically significant at the 1 % level

influence the probability of a significant under- or outperformance of SRI funds compared with conventional funds.

Almost 75 % of the performance comparisons (SRI with conventional funds) do not find any significant performance difference. A significant out- and underperformance is virtually found to the same degree (13 and 15 %, respectively). Furthermore, the most studied time period in primary studies is 1991–2000. In addition, approximately 50 % of the effects investigate funds of the US.

Significant evidence is found that the consideration of survivorship bias increases (decreases) the probability of a significant outperformance (underperformance) of SRI funds. Therefore, on the one hand, it is necessary for future studies to report on the treatment of the survivorship bias in detail. On the other hand, the evidence of this study can be used to interpret the results of existing studies. Further evidence reveals that effects, which investigate US SRI funds only, have a higher (lower) probability of an outperformance (underperformance) compared with effects which focus on funds of other countries. The most important implication of this evidence is that if the results of the US studies are sample specific, it is reasonable to investigate SRI funds of other countries in more detail. Some studies started to investigate SRI funds around the globe (e.g., Renneboog et al. 2008b), but further evidence is needed to cope with special circumstances of national SRI markets. This could be particularly interesting for European countries, as they have the largest share of the

Table 7 Results of the weighted meta-regression with the dependent variable *underperformance* (logit model)

	(1)		(2)	
	Coef.	Std. Err.	Coef.	Std. Err.
Performance evaluation Jensen's Alpha	0.085	0.063	0.030	0.032
Performance evaluation Carhart's Alpha	0.026	0.055	−0.001	0.037
Conditional performance evaluation	−0.135***	0.037	−0.072*	0.039
1 vs. 1 matching	−0.198***	0.036	−0.215***	0.041
1 vs. many matching	−0.174***	0.032	−0.218***	0.037
Number of matching criteria	0.053***	0.015	0.051***	0.019
Highest number of authorships	−0.080***	0.029	−0.136***	0.033
Survivorship bias consideration	−0.200***	0.032	−0.238***	0.039
US funds	−0.110***	0.032	−0.111***	0.022
Time period 1981–2000	−0.166***	0.028	−0.053**	0.026
Number of SRI funds			0.002***	0.001
Number of conventional funds			−0.000**	0.000
Obs	477		376	
Log pseudolikelihood	−6.682		−4.067	
Pseudo R2	0.378		0.606	

This table shows the average marginal effects of the independent variables in decimal notation and standard errors. The dependent variable is *underperformance*, which takes the value 1 if the SRI funds in a study significantly underperform the conventional funds, *underperformance* = 0 in all other cases

* Coefficient is statistically significant at the 10 % level

** Coefficient is statistically significant at the 5 % level

*** Coefficient is statistically significant at the 1 % level

Table 8 Results of the weighted meta-regression with the dependent variable *outperformance* (logit model)

	(1)		(2)	
	Coef.	Std. Err.	Coef.	Std. Err.
Performance evaluation Jensen's Alpha	−0.101***	0.017	−0.079***	0.016
Performance evaluation Carhart's Alpha	−0.055*	0.030	−0.074**	0.031
Conditional performance evaluation	0.190*	0.110	0.356***	0.138
1 vs. 1 matching	0.311*	0.182	0.425***	0.134
1 vs. many matching	0.195	0.130	0.294**	0.148
Number of matching criteria	−0.028	0.019	−0.026	0.021
Highest number of authorships	0.004	0.033	−0.054*	0.032
Survivorship bias consideration	0.163*	0.087	0.169**	0.084
US funds	0.114**	0.048	0.195**	0.087
Time period 1981–2000	−0.048	0.032	−0.090*	0.047
Number of SRI funds			−0.000	0.000
Number of conventional funds			0.000	0.000
Obs	477		376	
Log pseudolikelihood	−5.970		−4.369	
Pseudo R^2	0.188		0.331	

This table shows the average marginal effects of the independent variables in decimal notation and standard errors. The dependent variable is *outperformance*, which takes the value 1 if the SRI funds in a study significantly outperform the conventional funds, *outperformance* = 0 in all other cases

* Coefficient is statistically significant at the 10 % level

** Coefficient is statistically significant at the 5 % level

*** Coefficient is statistically significant at the 1 % level

global SRI market (Eurosif 2010). It is important to note that the findings on this variable seem to be driven largely by the results of two primary studies (Renneboog et al. 2008b; Gil-Bazo et al. 2010). The results of primary studies are sensitive to the time period of an effect as well, but based on the results of this study it is difficult to draw general conclusions on this variable. Additional results from the weighted meta-regression suggest that the use of a matching procedure (1 vs. 1 and 1 vs. many matching) increases (decreases) the probability of an outperformance (underperformance) of SRI funds.

Regarding the meta-level, future research might explore the influence of additional study characteristics. On the level of primary studies, it may be reasonable to investigate differences between US and non-US SRI funds empirically. A further interesting topic could be the dissimilar attrition rates of SRI and conventional funds.

Acknowledgments I want to thank Hannes Winner for his valuable suggestions throughout the process of conducting this paper and Claudia B. Woehle for her general support and feedback. Furthermore, I appreciate the very helpful comments of the two anonymous reviewers. Several (former) colleagues at the University of Salzburg helped to improve the paper significantly – namely, Harald Oberhofer, Benjamin Furlan, Andreas Pacher, Matthias Stöckl, Philipp Weigl, Klaus Nowotny and Alex Avedikjan. I am deeply grateful to Magdalena Braendle for our fruitful discussions and her suggestions for improvement.

Appendix

See Tables 9 and 10

Table 9 Detailed information on the included studies

Study	Survivorship bias consideration	US funds	Time period 1981–1990	Time period 1991–2000	Time period 2001–2008	Performance evaluation Jensen's Alpha	Performance evaluation Carhart's Alpha	Other performance evaluation	Conditional performance evaluation	1 vs. 1 Matching	1 vs. Many matching	Number of matching criteria
Bauer et al. (2007)	0	0	0	1	0	1	1	0	0/1	0	0	0
Bauer et al. (2005)	1	0/1	0	1	0	1	1	0	0	0	1	2
Bauer et al. (2006)	1	0	0	1	1	0	1	0	0/1	0	0	0
Bello (2005)	0	1	0	1	0	1	0	1	0	0	1	1
Benson et al. (2006)	0	1	0	1	1	0	1	1	0	0	0	0
Bollen (2007)	1	1	1	1	1	1	1	1	0	0	0	0
Chang and Witte (2010)	0	1	0	0	1	1	0	1	0	0	0	0
Derwall and Koedijk (2009)	1	1	0	1	0	1	1	1	0	0	1	3
Gil-Bazo et al. (2010)	1	1	0	1	1	1	1	1	0	0/1	0/1	0/3/5
Goldreyer et al. (1999)		1				0	0	1	0	0	0	1/2
Gregory et al. (1997)	1	0	1	1	0	0	1	1	0	0/1	0	4
Gregory and Whitaker (2007)	1	0	0	1	0	0	1	0	0/1	0	1	2
Hamilton et al. (1993)	0	1	1	0	0	1	0	0	0	0	1	1
Humphrey and Lee (2011)	1	0	0	0	1	1	1	0	0	0/1	0	0/5
Kempf and Osthoff (2008)	1	1	0	1	0	0	1	0	0	0	0	0
Koellner et al. (2007)	0	0	0	0	1	0	0	1	0	1	0	1

Table 9 continued

Study	Survivorship bias consideration	US funds	Time period 1981–1990	Time period 1991–2000	Time period 2001–2008	Performance evaluation Jensen's Alpha	Performance evaluation Carhart's Alpha	Other performance evaluation	Conditional performance evaluation	1 vs. 1 Matching	1 vs. Many matching	Number of matching criteria
Kreander et al. (2005)	0	0	0	1	0	1	1	1	0	1	0	4
Kryzanowski et al. (2011)	1	0	0	1	0	1	1	1	0/1	0	0	0
Liedekerke et al. (2007)	1	0	0	1	1	0	1	0	1	0	0	0
Mueller (1991)		1	1	0	0	0	0	1	0	0	0	0
Renneboog et al. (2008b)	1	0/1	0	1	1	1	1	1	0/1	0	0/1	0/5
Sánchez and Sotorrío (2009)	0	0	0	0	1	0	0	1	0	1	0	3
Spekl (2009)	0	0	0	0	1	1	1	1	0	1	0	4
Statman (2000)	0	1	0	1	0	1	0	1	0	0	1	1
Stenström and Thorell (2007)	0	0	0	0	1	1	0	0	0	0	0	0

This table presents detailed information on the independent variables of the meta-regression, respectively, on the included studies. Value 1 is taken if the effects of a study, for example, consider survivorship bias (second column). Value 0 is taken if the effects of a study do not consider survivorship bias

Table 10 Detailed information on the included studies

Study	Highest number of authorships	Number of SRI funds	Number of conventional funds
Bauer et al. (2007)	3	8	267
Bauer et al. (2005)	3	50	150
Bauer et al. (2006)	3	15	195
Bello (2005)	1	42	84
Benson et al. (2006)	2	184	6074
Bollen (2007)	1	187	9189
Chang and Witte (2010)	1	164	11913
Derwall and Koedijk (Derwall 2009)	2	15	75
Gil-Bazo et al. (2010)	1	86	1761
Goldreyer et al. (1999)	1	29	20
Gregory et al. (1997)	2	16	92
Gregory and Whitaker (2007)	2	20	100
Hamilton et al. (1993)	2	17	170
Humphrey and Lee (2011)	2	27	514
Kempf and Osthoff (2008)	1	72	3906
Koellner et al. (2007)	1	13	13
Kreander et al. (2005)	1	30	30
Kryzanowski et al. (2011)	1	67	517
Liedekerke et al. (2007)	1	28	725
Mueller (1991)	1	10	
Renneboog et al. (2008b)	1	340	680
Sánchez and Sotorrío (2009)	1	103	103
Spekl (2009)	1	133	133
Statman (2000)	2	31	62
Stenström and Thorell (2007)	1	23	42

This table presents detailed information on the independent variables of the meta-regression, respectively, on the included studies

References

*Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking and Finance, 29*(7), 1751–1767.

*Bauer, R., Otten, R., & Rad, A. T. (2006). Ethical investing in Australia, is there a financial penalty? *Pacific-Basin Finance Journal, 14*(1), 33–48.

*Bauer, R., Derwall, J., & Otten, R. (2007). The ethical mutual fund performance debate: New evidence from Canada. *Journal of Business Ethics, 70*(2), 111–124.

*Bello, Z. Y. (2005). Socially responsible investing and portfolio diversification. *Journal of Financial Research*, *28*(1), 41–57.

Bengtsson, E. (2008). A history of scandinavian socially responsible investing. *Journal of Business Ethics, 82*(4), 969–983.

*Benson, K. L., Brailsford, T. J., & Humphrey, J. E. (2006). Do socially responsible fund managers really invest differently?. *Journal of Business Ethics, 65*(4), 337–57.

*Bollen, N. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis*, *42*(3), 683–708.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex: Wiley.

Brown, S. J., Goetzmann, W., Ibbotson, R. G., & Ross, S. A. (1992). Survivorship bias in performance studies. *The Review of Financial Studies, 5*(4), 553–580.

Capelle-Blancard, G., & Monjon, S. (2010). *The performance of socially responsible funds: Does the screening process matter?* Paris: Working paper, Université Paris 1 Panthéon-Sorbonne.

Carhart, M. M. (1997). On the persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57–82.

*Chang, C. E., & Witte, H. D. (2010). Performance evaluation of U.S. socially responsible mutual funds: Revisiting doing good and doing well. *American Journal of Business*, *25*(1), 9–22.

Chegut, A., Schenk, H., & Scholtens, B. (2011). Assessing SRI fund performance research: Best practices in empirical analysis. *Sustainable Development, 19*(2), 77–94.

Cortez, M. C., Silva, F., & Areal, N. (2009). The performance of european socially responsible funds. *Journal of Business Ethics, 87*(4), 573–588.

*Derwall, J., & Koedijk, K. (2009). Socially responsible fixed-income funds. *Journal of Business Finance & Accounting*, *36*(1/2), 210–229.

Eurosif. (2010). *European SRI Study 2010*. Paris.

Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *The Journal of Finance, 51*(2), 425–461.

García-Quevedo, J. (2004). Do public subsidies complement business R&D? A meta-analysis of the econometric evidencec. *Kyklos, 57*(1), 87–102.

*Gil-Bazo, J., Ruiz-Verdu, P., & Portela, A. A. P. (2010). The performance of socially responsible mutual funds: the role of fees and management companies. *Journal of Business Ethics*, *94*(2), 243–263.

*Goldreyer, E. F., Ahmed, P., & Diltz, J. D. (1999). The performance of socially responsible mutual funds: incorporating sociopolitical information in portfolio selection. *Managerial Finance, 25*(1), 23–36.

*Gregory, A., & Whittaker, J. (2007). Performance and performance persistence of ethical unit trusts in the UK. *Journal of Business Finance & Accounting, 34*(7–8), 1327–1344.

*Gregory, A., Matatko, J., & Luther, R. (1997). Ethical unit trust financial performance: Small company effects and fund size effects. *Journal of Business Finance & Accounting, 24*(5), 705–725.

*Hamilton, S., Jo, H., & Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal, 49*(6), 62–66.

Heal, G. (2008). *When principles pay. Corporate social responsibility and the bottom line*. New York: Columbia University Press.

Hoepner, A. G. F., & McMillan, D. G. (2009). *Research on 'responsible investment': An influential literature analysis comprising a rating, characterisation, categorisation & investigation*. St. Andrews: Working paper, University of St. Andrews.

Horvathova, E. (2010). Does environmental performance affect financial performance? A meta-analysis. *Ecological Economics, 70*(1), 52–59.

*Humphrey, J. E., & Lee, D. D. (2011). Australian socially responsible funds: Performance, risk and screening intensity. *Journal of Business Ethics, 102*(4), 519–533.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis. Correcting error and bias in research findings*. Newbury Park: Sage Publications.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance, 23*(2), 389–416.

*Kempf, A., & Osthoff, P. (2008). SRI funds: nomen est omen. *Journal of Business Finance and Accounting, 35*(9/10), 1276–1294.

*Koellner, T., Suh, S., Weber, O., Moser, C., & Scholz, R. W. (2007). Environmental impacts of conventional and sustainable investment funds compared using input-output life-cycle assessment. *Journal of Industrial Ecology, 11*(3), 41–60.

*Kreander, N., Gray, R. H., Power, D. M., & Sinclair, C. D. (2005). Evaluating the performance of ethical and non-ethical funds: A matched pair analysis. *Journal of Business Finance & Accounting, 32*(7–8), 1465–1493.

*Kryzanowski, L., Ayadi, M., & Ben-Ameur, H. (2011). *Luck versus skill in the cross-section of ethical mutual funds*. Montreal: Working paper, Concordia University.

*Liedekerke, L. V., De Moor, L., & Walleghem, D. V. (2007). Risk-return of Belgian SRI funds. *Tijdschrift voor Economie en Management, 52*(4), 673–685

Louche, C., & Lydenberg, S. (2006). *Socially responsible investment: Differences between Europe and United States*. Gent: Working paper, Vlerick Leuven Gent Management School.

*Mueller, S. (1991). The opportunity cost of discipleship: Ethical mutual funds and their returns. *Sociological Analysis*, *52*(1), 111–124.

Nelson, J. P., & Kennedy, P. E. (2009). The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment. *Environmental & Resource Economics, 43*(3), 345–377.

Orlitzky, M., Schmidt, F. L., & Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization Studies, 24*(3), 403–441.

Renneboog, L., Horst, J. T., & Zhang, C. (2008a). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance, 32*(9), 1723–1742.

*Renneboog, L., Horst, J. T., & Zhang, C. (2008b). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance, 14*(3), 302–322.

*Sánchez, J. L. F., & Sotorrío, L. L. (2009). *Performance of European SRI funds vs. conventional funds*. Santander: Working paper, University of Cantabria.

Sparkes, R., & Cowton, C. J. (2004). The maturing of socially responsible investment. A review of the developing link with corporate social responsibility. *Journal of Business Ethics, 52*(1), 45–57.

*Spekl, A. (2009). *European evidence on SRI mutual fund performance*. Maastricht: Unpublished master thesis, Universiteit Maastricht.

*Statman, M. (2000). Socially responsible mutual funds. *Financial Analysts Journal*, *56*(3), 30–39.

*Stenström, C., & Thorell, J. J. (2007). *Evaluating the performance of socially responsible investment funds: A holding data analysis*. Stockholm: Unpublished master thesis, Stockholm School of Economics.

Van Beurden, P., & Goessling, T. (2008). The worth of values. A literature review on the relation between corporate social and financial performance. *Journal of Business Ethics, 82*(2), 407–424.

Vigeo. (2011). *Green, social and ethical funds in Europe*. 2011 Review', Paris.